Noun-phrase production as a window to language selection: an ERP study

Von Grebmer Zu Wolfsthurn, S.; Pablos Robles, L.; Schiller, N.O.

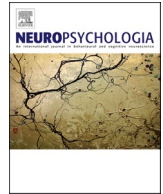**Note:** To cite this publication please use the final published version (if applicable).

Noun-phrase production as a window to language selection: An ERP study

Sarah von Grebmer zu Wolfsthurn [a,b,*], Leticia Pablos [a,b], Niels O. Schiller [a,b]

[a] *Leiden University Centre for Linguistics (LUCL), Reuvensplaats 3-4, 2311, BE Leiden, the Netherlands*
[b] *Leiden Institute for Brain and Cognition (LIBC), LUMC, PO Box 9600, 2300, RC Leiden, the Netherlands*

ABSTRACT

Characterising the time course of non-native language production is critical in understanding the mechanisms behind successful communication. Yet, little is known about the modulating role of cross-linguistic influence (CLI) on the temporal unfolding of non-native production and the locus of target language selection. In this study, we explored CLI effects on non-native noun phrase production with behavioural and neural methods. We were particularly interested in the modulation of the P300 as an index for inhibitory control, and the N400 as an index for co-activation and CLI. German late learners of Spanish overtly named pictures while their EEG was monitored. Our results indicate traceable CLI effects at the behavioural and neural level in both early and late production stages. This suggests that speakers faced competition between the target and non-target language until advanced production stages. Our findings add important behavioural and neural evidence to the underpinnings of non-native production processes, in particular for late learners.

1. Introduction

From the speaker's perspective, producing a determiner-noun phrase (NP) e.g., [*the flower*] seems an effortless operation. However, according to current models, word production is a complex, multi-stage process. For example, the influential LRM model (Levelt et al., 1999) describes three primary stages of single word production. In a picture naming task, first, the depicted object is conceptualised; second, the concept is lexicalized, i.e., it is given a grammatical, phonological and phonetic form; and finally, the name of the object is articulated. Current research has increasingly focused on characterising the time course of word production. For example, Indefrey and Levelt (2004) and Indefrey (2011) combined behavioural, chronometric and electrophysiological evidence to estimate the time course of each stage in native language word production (Abdel Rahman and Sommer, 2008; Camen et al., 2010; Cheng et al., 2010; Hanulová et al., 2011; Laganaro et al., 2009; Rodriguez-Fornells et al., 2002; Schiller et al., 2003; Schiller, 2006; Schmitt et al., 2001; Van Turennout et al., 1997; Zhang and Damian, 2009), see Fig. 1.

Electroencephalography (EEG) and event-related potentials (ERPs) are particularly valuable tools to explore the neuro-cognitive processes of native language production. More specifically, the EEG signal yields an implicit measure of the neural signature and the time course of each individual production stage (Aristei et al., 2011; Bürki and Laganaro,

2014; Costa et al., 2009; Eulitz et al., 2000; Habets et al., 2008; Hoshino and Thierry, 2011; Valente et al., 2014). For example, Bürki and Laganaro (2014) found that producing French NPs, such as [*le chat*] "the cat" or NPs including an adjective [*le grand chat*] "the big cat", in comparison to a bare noun [*chat*] "cat" was linked to the topographic stability of the EEG signal between 190 ms and 300 ms and following 530 ms post-stimulus onset. These findings were interpreted as a longer duration of *lexical retrieval* (lemma retrieval in LRM terms) and *phonological encoding* for NPs compared to bare nouns. They were corroborated by longer naming latencies for bare nouns and NPs compared to NPs including an adjective (Bürki and Laganaro, 2014; Lange et al., 2015; Schriefers et al., 1999). *Lexical retrieval* has been previously associated with lexical access and grammatical gender processing (Alario and Caramazza, 2002; Badecker et al., 1995; Bürki and Laganaro, 2014; Levelt et al., 1999; Strijkers et al., 2010). In contrast, *phonological encoding* was described as the processing of the phonological code of the word and its subsequent syllabification (Levelt et al., 1999). Grammatical gender, hereafter gender, is a noun classification system (Corbett, 1991). More specific to gender processing in Romance languages, research suggested that the activation and selection of determiners in NPs occurred both during *lexical retrieval* and at the early part of the consecutive *phonological encoding* (Alario and Caramazza, 2002; Bürki et al., 2016; Miozzo and Caramazza, 1999; Sá-Leite et al., 2020). As shown in Bürki et al. (2014), in these languages producing the

phonological forms of determiners and adjectives is partially dependent on the phonological form of the noun, e.g., Spanish [la_F taza_F roja_F] vs. English [the red mug] (Miozzo and Caramazza, 1999; Sá-Leite et al., 2020; Schriefers, 1992, 1993).

The work by Bürki and Laganaro (2014) and similar studies (Eulitz et al., 2000; Habets et al., 2008; Koester and Schiller, 2008; Lange et al., 2015) characterised the time course of native language word production. However, the time course of non-native production continues to be a complex issue in multilingualism research, especially with respect to the locus of target language selection (Costa et al., 2009; Hanulová et al., 2011; Hoshino and Thierry, 2011; Strijkers et al., 2010). In light of the increasing proportion of non-native speakers and multilingual communities (Berthele, 2021), the need to further characterise the individual production stages in non-native production has become more urgent.

In this study, we build upon the theoretical models of native speaker single word production and empirical findings on native NP production (Bürki and Laganaro, 2014; Indefrey, 2011; Indefrey and Levelt, 2004; Levelt et al., 1999). We specifically concentrated on the time course of those production stages preceding the articulation stage, namely *lexical retrieval* and *phonological encoding*. Collecting behavioural and EEG measures, we examined the overt production of determiner + noun NPs in the non-native language Spanish, e.g., [la flor] "the flower", by native speakers of German.

Producing utterances can demonstrably be more challenging in the non-native than in the native language (Pivneva et al., 2012; Runnqvist et al., 2011). Studies have found longer and more variable naming latencies in the non-native compared to the native language (Gollan et al., 2005; Hanulová et al., 2011; Ivanova and Costa, 2008; Kroll et al., 2006). These quantitative differences between native and non-native speech production were reported for various levels of language proficiency (Christoffels et al., 2007; Ivanova and Costa, 2008; Sholl et al., 1995), and for language pairs with varying phonological and orthographic similarity, e.g., intermediate German learners of Dutch (Christoffels et al., 2007) and highly proficient Greek learners of English (Parker-Jones et al., 2012). The question which arises at this point is: where does this delay in naming latencies originate from? Recent studies explored word frequency and age of acquisition (AoA) as modulating factors of the non-native production processes (see Hanulová et al., 2011; for discussion). In this study, we focus on another factor that could influence the time course of non-native short utterance production, namely *cross-linguistic influence* (CLI), as described in section 1.1. By extension, we aim to explore the following question crucial to non-native production research: during which production stage does the delay in naming latency occur? In section 1.2, we outline the electrophysiological correlates of CLI in more detail and we discuss how they offer us an insight into these issues.

1.1. Cross-linguistic influence

Non-native speakers face cross-linguistic influence (CLI) during language production and language comprehension (Cárdenas-Hagan et al., 2007; Ganushchak et al., 2011b; Lemhöfer et al., 2008; Morales et al., 2016; Müller and Hulk, 2001; Thierry and Wu, 2007; von Grebmer zu Wolfsthurn et al., 2021). Broadly speaking, CLI is the interaction of the languages within a multilingual system and its influence on the underlying cognitive processing mechanisms. CLI supports the notion that the native and non-native language are co-activated during language production (Guo and Peng, 2006; Hermans et al., 1998; Kroll et al., 2008, Lee and Williams, 2001). Co-activation and CLI are rooted into theoretical models. For example, the Revised Hierarchical Model (RHM; Kroll and Stewart, 1994) postulates a conceptual level and separate lexical levels for the native and non-native language with strong lexical connections between the two languages. Critically, the model also suggests that the strength of the connections is modulated by proficiency: as non-native proficiency increases, the connection strength between the non-native lexicon and the conceptual level increases and the involvement of the L1 becomes less prominent (Kroll and Stewart, 1994).

For a speaker to successfully complete a naming task in either the native or non-native language, it is crucial that co-activation and CLI are resolved prior to articulation. A robust finding in the CLI and language selection literature is the presence of an inhibitory control system which mitigates CLI effects and effectively inhibits the non-target language (Abutalebi and Green, 2007; Green, 1998), but see (Verdonschot et al., 2012). The mitigation of CLI effects and the associated increased cognitive effort is evident at the neural level. For example, increased activation of brain areas involved in language production for non-native compared to native language production was linked to increased error monitoring of competing representations during CLI (Parker-Jones et al., 2012; Rodriguez-Fornells et al., 2005; Rossi et al., 2018).

Previous studies on CLI have almost exclusively focused on early acquisition and intermediate to high proficiency levels in the non-native language, thereby leaving a systematic gap in the exploration of the effects of CLI on the time course of NP production in late language learners with lower proficiency levels (Costa et al., 2003; Hoshino and Thierry, 2011; Lemhöfer et al., 2008). Yet, this is a critical issue because studies suggested that proficiency impacted language-related neurocognitive mechanisms in multilinguals, shown in that CLI effects were more pronounced at lower proficiency levels (Bosch and Unsworth, 2020; Heidlmayr et al., 2021; Steinhauer et al., 2009; Van der Meij et al., 2011; White et al., 2017; Yip and Matthews, 2007). For example, Sunderman and Kroll (2006) found that compared to lower proficient learners, highly proficient English learners of Spanish were less susceptible to CLI from the native language, and performed better in a picture naming task. Costa and Santesteban (2004) previously proposed that during production, highly proficient speakers activated only the lexical entry from the target language, thereby effectively avoiding CLI during *lexical retrieval*. Therefore, in this study we directly focused on CLI effects in a group where first, CLI was found to be most prevalent; and second, is frequently understudied in the literature, namely late language learners with intermediate proficiency levels. We defined late language learners as having acquired an additional (non-native) language later in development (AoA >12 years), see Rossi et al. (2006). Moreover, and in contrast to highly proficient late language learners,
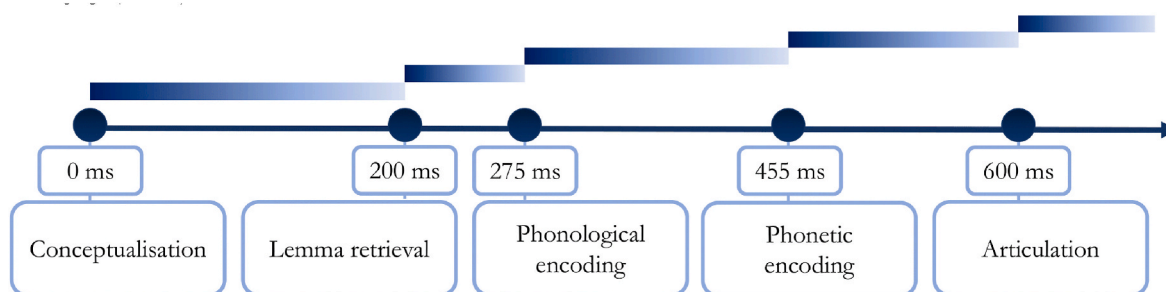

Fig. 1. *Estimated time course of single word production in the native language according to* Indefrey and Levelt (2004) *and* Indefrey (2011). [11].

our group was further characterised by less than three years of exposure to the non-native language and intermediate proficiency levels in the B1/B2 range according to the Common European Framework of Reference for Languages, CEFR (Council of Europe, 2001).

Immediately relevant to CLI effects is the question about when the target language is selected in non-native production. For example, is the target language selected (and CLI resolved), prior to *lexical retrieval*? Or instead, does CLI carry over to later production stages, such as *phonological encoding*? Current debates remain inconclusive with respect to two accounts of the locus of target language selection (Costa et al., 2006; Hanulová et al., 2011; Hoshino and Thierry, 2011; Sá-Leite et al., 2019). One account suggests that lexical entries from both the target and non-target language are activated, but only the lexical entry corresponding to the target language is selected for subsequent phonological processing (Gollan et al., 2005; Hermans et al., 1998; Lee and Williams, 2001). Under this account, CLI is resolved at *lexical retrieval*. On the other hand, a second account suggests that the lexical entries from the target and non-target language are both activated and selected for *phonological encoding* (Christoffels et al., 2007; Colomé, 2001; Costa et al., 2000; Green, 1998; Hoshino and Kroll, 2008; Pulvermüller, 2007; Rodriguez-Fornells et al., 2005). Within this perspective, CLI is not resolved at *lexical retrieval*, but continues into subsequent phonological processing.

In order to discriminate between these two accounts, in this study we focus on two linguistic phenomena representing CLI, the *gender congruency effect* and the *cognate facilitation effect*. These effects provide us with further insight into the underlying production stages and their inner mechanisms in non-native NP production (sections 1.1.1. and 1.1.2., respectively).

1.1.1. The gender congruency effect

The gender congruency effect is reflected in faster processing of *congruent* versus *incongruent* nouns, as reported in language production studies (Bordag and Pechmann, 2007; Lemhöfer et al., 2008; Morales et al., 2016; Paolieri et al., 2019, 2020; Schiller and Caramazza, 2003; Schiller, 2006; Schiller and Costa, 2006; Schriefers, 2003). Congruent nouns have similar grammatical gender values across languages, for example the lexical items for the concept "arm", which are masculine in German [der$_M$ Arm] and in Spanish [el$_M$ brazo]. In contrast, incongruent nouns have dissimilar gender values across languages, for example the lexical items for "key", which are masculine in German [der$_M$ Schlüssel] and feminine in Spanish [la$_F$ llave]. Gender systems can vary across languages. For example, German has three gender values, i.e. feminine, masculine and neuter, and their distribution is not equally distributed across all lexical items (Schiller and Caramazza, 2003). On the other hand, Spanish is characterized by a feminine-masculine gender value distinction with an approximately balanced distribution (Bull, 1965; Eddington, 2002).[2]

As previously discussed, gender processing in Romance languages was linked to *lexical retrieval* and *phonological encoding* (Alario and Caramazza, 2002; Badecker et al., 1995; Bürki and Laganaro, 2014; Miozzo and Caramazza, 1999). Subsequently, this links the gender congruency effect to these two production stages. Therefore, the gender congruency effect offers a gateway to the following three issues: first, it allows us to observe the mechanisms underlying CLI of the gender systems during *lexical retrieval*; second, it provides us with a way to study the implications of CLI of the gender systems on the time course of non-native NP production; and third, it allows us to explore the locus of target language selection with respect to the two accounts regarding target language selection (Hoshino and Thierry, 2011; Sá-Leite et al., 2019). If the target language was selected prior to *lexical retrieval* in multilingual language production, we would not observe a gender congruency effect during subsequent production stages. Alternatively, if the target language was not selected before *lexical retrieval*, we would observe a gender congruency effect because the activated lexical entries from both languages would be subject to CLI of the gender systems during gender processing. As a result, CLI would facilitate the processing of congruent nouns compared to incongruent nouns. Current behavioural evidence supports this notion in late language learners with intermediate proficiency compared to early learners with high proficiency (Bordag and Pechmann, 2007; Costa et al., 2003; Sá-Leite et al., 2020).

Yet, few studies have investigated the gender congruency effect from a neural perspective (Heim et al., 2009). One ERP component associated with this effect is the N400 (Paolieri et al., 2020; Wicha et al., 2003). It is reflected in a negative voltage amplitude peak around 400 ms post-stimulus onset and was previously linked to lexical-semantic integration and lexical co-activation (Chen et al., 2017; Hoshino and Thierry, 2011; Kutas and Federmeier, 2011; Lau et al., 2008; Leckey and Federmeier, 2019). In relation to the gender congruency effect, less negative N400 amplitudes were linked to congruent trials compared to incongruent trials in the time window between 300 ms and 500 ms post-stimulus onset in a translation-recognition task (Paolieri et al., 2020; see also Wicha et al., 2003). Given the temporal characteristics of these neural correlates of the gender congruency effect, this suggests that both co-activation and CLI between the languages may remain unresolved until around 500 ms post-stimulus onset. In LRM terms, this time window coincides with *phonological encoding*. Therefore, the behavioural and ERP findings related to the gender congruency effect suggest that the target language is not selected prior to *lexical retrieval*. However, from these findings it remains unclear whether CLI is resolved upon termination of gender processing or whether it indeed continues into *phonological encoding* stages.

Regarding the specific question about the locus of target language selection, there are two possible scenarios. In the first scenario, the target language is selected after the completion of gender processing. Subsequently, only the lexical entry from the target language carries over to *phonological encoding*. In contrast, the second scenario postulates that the target language is selected during or after *phonological encoding*. Here, the lexical entries from both the target and non-target language are processed for *phonological encoding*. This implies that both languages remain active after the completion of gender processing and that CLI potentially results in further delays during later phonological processing. Evidence by Hoshino and Thierry (2011) preliminarily supported the latter notion. In a picture-word interference (PWI) task with highly proficient Spanish learners of English, the EEG signal was modulated during *lexical retrieval* for semantically and phonologically related trials compared to unrelated trials. However, they found no further modulation of the signal after 400 ms post-stimulus onset. The authors interpreted these results as showing that the target language had not been selected at *lexical retrieval*, but that the selection had taken place by 400 ms post-stimulus onset. In LRM terms, this time window coincides with *phonological encoding*. Therefore, these results supported the second scenario whereby first, lexical entries from both the target and non-target language were selected for phonological processing, and second, CLI continued beyond *lexical retrieval* (Christoffels et al., 2007; Colomé, 2001; Costa et al., 2000; Hoshino and Kroll, 2008; Hoshino and Thierry, 2011; Pulvermüller et al., 2009; Rodriguez-Fornells et al., 2005). Building on the work by Hoshino and Thierry (2011) to add further evidence for discriminating between the two scenarios outlined above, we also explored the *cognate facilitation effect*. Via the exploration of this effect, we probed whether or not CLI continued after gender processing in late language learners.

---

[1] Note that Fig. 1 serves for visualisation purposes and does not claim that the production processing stages are discrete stages or follow a sequential pattern, in line with Levelt et al. (1999) and Indefrey (2011). This notion is subject to an open debate (Camen et al., 2010), which is beyond the scope of our study.

[2] Note that similar labels for gender values ("masculine", "feminine") can be found across different languages. However, we do not assume that these labels are conceptually identical (see also Lemhöfer et al., 2010) but merely utilise them for descriptive purposes.

1.1.2. The cognate facilitation effect

In a broad sense, cognates are words with a large degree of phonological and orthographic overlap (Li and Gollan, 2018). For example, [*Melone*] and [*melón*] "melon" are examples for *cognates* in German and Spanish, respectively, whereas [*Arm*] and [*brazo*] "arm" are *non-cognates*. It has previously been shown that cognates are processed faster compared to non-cognates (Bosma et al., 2019; Casaponsa et al., 2015; Li and Gollan, 2018). This is a critical finding as it suggests that this *cognate facilitation effect* can be linked to the CLI during *phonological encoding* in production (Christoffels et al., 2007; Costa et al., 2000, 2005; Hoshino and Kroll, 2008). In order for this effect to occur, the lexical entries from both the target and non-target language need to be subject to subsequent *phonological encoding*. Here, we used the cognate facilitation effect to first, study CLI during this particular production stage with respect to the overall time course of non-native production; and second, to add to the discussion of whether or not the target language is selected after gender processing.

While neural correlates of the cognate facilitation effect have been scarcely researched in non-native production, evidence from non-native comprehension links the N400 to this specific effect (Midgley et al., 2011; Peeters et al., 2013; Xiong et al., 2020). For example, Peeters and colleagues (2013) found faster latencies and smaller N400 amplitudes for cognates compared to non-cognates between 400 ms and 500 ms in French late learners of English in a lexical decision task. Contrastingly, Christoffels et al. (2007) found faster naming latencies and more negative amplitudes for cognates compared to non-cognates in fronto-central regions between 275 ms and 475 ms post-stimulus onset in unbalanced German-Dutch speakers in an overt picture naming task, without linking it to the N400. Further, studies showed that the size of the cognate facilitation effect decreased as non-native proficiency increased (Bultena et al., 2014; Casaponsa et al., 2015). Therefore, ERP evidence from the cognate facilitation effect suggests that first, the target and the non-target language are co-activated, which in turn leads to CLI of the phonological systems; second, that lexical entries from both languages are subject to *phonological encoding;* and finally, that the non-target language is not inhibited during *lexical retrieval*, particularly at lower proficiency levels. Instead, the cognate facilitation effect suggests that CLI continues beyond lexical retrieval into *phonological encoding* stages, as supported by the literature (Christoffels et al., 2007; Colomé, 2001; Costa et al., 2000; Hoshino and Kroll, 2008; Hoshino and Thierry, 2011).

Combining evidence from the gender congruency effect and the cognate facilitation effect, the findings presented above suggest a modulating influence of CLI on the time course of non-native production that continues beyond gender processing until phonological processing stages. However, these interpretations are debatable given the scarcity of research on CLI, in particular in terms of the neural correlates of the gender congruency effect and the cognate facilitation effect. Therefore, aside from exploring CLI effects in late language learners, we also focused on the neural underpinnings of CLI to characterise the time course of non-native production and the locus of target language selection.

1.2. Electrophysiological correlates of CLI

As discussed above, the N400 was linked to CLI effects such as the gender congruency effect and the cognate facilitation effect, but also the co-activation of languages (Chen et al., 2017; Paolieri et al., 2020; Peeters et al., 2013). In this study, we used the N400 to capture the co-activation and the linguistic aspects of CLI. Importantly, studies suggested that the N400 onset may be delayed in late language learners and that overall N400 amplitudes in these learners decrease compared to those of native speakers (Midgley et al., 2009; Weber-Fox and Neville, 1996). Some studies further suggested that N400 amplitudes were modulated by non-native proficiency, i.e., the N400 became more native-like with increasing proficiency (Midgley et al., 2009; Newman et al., 2012; White et al., 2017; but see Wood Bowden et al., 2013). For

example, in a phoneme discrimination task with French low and high proficient late language learners of English, Heidlmayr et al. (2021) found a smaller N400 effect for the low compared to the high proficient group of late language learners. Therefore, with respect to our group of late language learners, the N400 is likely to be delayed, or smaller in size, compared to that in Paolieri et al. (2020) and Peeters et al. (2013).

In this study, we also focused on the P300 component to capture the cognitive mechanisms underlying the successful mitigation of CLI and the selection of the target language. The P300 is a positive-going deflection of the EEG signal with a peak around 300 ms post-stimulus onset. Early studies found that the P300 was elicited at different topographical sites (Ritter and Vaughan, 1969), leading to suggest separate P300 subcomponents. These subcomponents include the P3a, and the P3b (Barry et al., 2020; Polich, 2007). The P3a component is a positive-going wave with a fronto-central distribution which occurs around 200 ms–300 ms post-stimulus onset. In contrast, the P3b component was found in later 300 ms–400 ms time windows at centro-parietal electrodes (Hruby and Marsalek, 2003; Squires et al., 1975). Relevant to this study, the P300 has been previously linked to cognitive processes such as cognitive interference, cognitive control, working memory load and inhibition (Barker and Bialystok, 2019; Luck, 1998; Neuhaus et al., 2010; Polich, 2007). More recently, the P300 was linked to the allocation of attentional resources (Barker and Bialystok, 2019; González Alonso et al., 2020). It is typically found with inhibitory paradigms such as the Flanker task or the Oddball paradigm (Eriksen and Eriksen, 1974; Soares Pereira et al., 2019). More relevant to our study, it was also reported in paradigms which included a Flanker task preceded by a linguistic task such as code-switching or picture naming, e.g., in Bosma and Pablos (2020) and in Jiao et al. (2020). These studies highlighted the critical role of the P300 in inhibition in regulating native and non-native language use. Our experimental paradigm relies on the successful mitigation of CLI effects and the inhibition of the non-target language, therefore the P300 is a critical component to consider in this study alongside the N400.

1.3. The current study: non-native NP production

In the current study, we explore the effect of CLI on the time course of non-native NP production from a behavioural and neural perspective. The goals of the present study are twofold. First, we investigate how CLI affects behavioural measures and the EEG signal in non-native NP production. On the basis of the LRM model, we use two CLI effects to characterise the unfolding and neural signatures of individual production stages: *lexical retrieval* via the *gender congruency effect,* and *phonological encoding* via the *cognate facilitation effect*. Our second goal is to gain further insight into the process of target language selection. More specifically, we probe the locus of target language selection by investigating the two CLI effects with respect to the individual production stages.

Therefore, our research questions are: first, are there traceable effects of gender congruency (congruent vs. incongruent) and cognate status (cognate vs. non-cognate) at the behavioural and neural processing level in non-native NP production? Second, what can the temporal unfolding of processing gender congruency and cognate status tell us about the time course of non-native NP production? Finally, during which processing stage is the target language selected during non-native NP production?

We study non-native NP production in German late intermediate learners of Spanish with a B1/B2 proficiency level by employing an overt picture naming task. We combine behavioural measures of naming accuracy and naming latencies with EEG recordings. We are particularly interested in the modulation of the P300 as an index for inhibitory control, and the N400 as an index for co-activation and CLI. To obtain information about the linguistic background of our late language learners, we combine the Language Experience and Proficiency Questionnaire (LEAP-Q, Marian et al., 2007) with a Spanish vocabulary size

test, the LexTALE-Esp (Izura et al., 2014). To formulate the hypotheses for this study, we rely both on the time estimates proposed by Indefrey and Levelt (2004), Indefrey (2011), Bürki et al. (2014) and Hoshino and Thierry (2011) and on the theoretical framework and discussion of ERP components outlined in the previous sections.

1.3.1. Hypotheses

Behavioural hypotheses. We predict effects of gender congruency and cognate status on behavioural measures of naming accuracy and naming latencies. For congruent and cognate nouns, we predict a facilitatory CLI effect, reflected in higher naming accuracy and shorter naming latencies compared to incongruent and non-cognate nouns. In turn, this has direct implications for the time course of non-native NP production.

For congruent non-cognates and incongruent cognates, we predict more subtle CLI effects. In concrete behavioural terms, we expect lower naming accuracy and longer naming latencies for congruent non-cognates compared to congruent cognates, and higher naming accuracy and shorter naming latencies compared to incongruent non-cognates. On the other hand, we anticipate the reverse pattern for incongruent cognates: CLI would hinder gender processing, but act as a facilitator during phonological processing and influence the time course of non-native NP production.

EEG hypotheses. We first probe the presence of a P300 or an N400 effect, as existing research remains inconclusive about whether or not we can expect both components to be elicited in our experimental paradigm. Further, we predict a modulation of P300 and N400 as a function of condition. We expect smaller P300 and N400 amplitudes for producing congruent cognates compared to incongruent non-cognates. This would reflect higher processing costs and more involvement of the inhibitory control system for the latter.

For congruent non-cognates and incongruent cognates, we predict a similar degree of processing costs. These trials are subject to both CLI facilitation and hindrance. Therefore, we do not expect significant differences between congruent non-cognates and incongruent cognates in terms of P300 and N400 amplitudes. However, we do expect the P300 and N400 amplitudes for these particular trials to be significantly larger compared to congruent cognates, and to be significantly smaller for incongruent non-cognates. Therefore, we expect the smallest P300 and N400 amplitudes for congruent cognates, followed by larger amplitudes for congruent non-cognates and incongruent cognates, and finally, the largest amplitudes for incongruent non-cognates.

2. Methods

2.1. Participants

Thirty-three healthy, right-handed native German speakers (twenty-seven females) with a B1/B2 level of Spanish were recruited from the campus of the University of Konstanz (M_{age} = 23.06 years, SD_{age} = 2.47 years). At the time of testing, participants did not report any psychological or language disorders, nor visual and hearing impairments. Prior to the experiment, we provided all participants with an information sheet. Next, they signed an informed consent form before the experiment in compliance with the Ethics Code for linguistic research in the Faculty of Humanities at Leiden University. Upon termination of all tasks, participants received a debrief form, signed the final consent form and received a monetary compensation.

2.1.1. LEAP-Q: linguistic profile of participants

Prior to the experimental session, the linguistic profile of participants and their experience with Spanish was assessed using LEAP-Q (Marian et al., 2007); see Appendix A for details. We opted for a home-based administration of the questionnaire to minimise any self-report biases often induced by laboratory environment (Rosenman et al., 2011). The majority of participants (n = 31) reported English as their first

non-native language (M_{AOA} = 8.90, SD_{AOA} = 1.90), while two participants learned French as their first foreign language (M_{AOA} = 8.5, SD_{AOA} = 2.5). A total of sixteen participants learned Spanish as second non-native language. Further, Spanish was disclosed as third non-native language by fifteen participants, and as fourth non-native language by two participants.

The mean AoA of Spanish was M_{AOA} = 16.29 years (SD_{AOA} = 2.39). Participants reported to be fluent in Spanish on average at M = 18.53 years of age (SD = 2.29). They started to read in Spanish at M = 17.27 years of age (SD = 3.03). Before the time of testing, almost all participants (n = 31) spent some time in a Spanish-speaking country (M = 0.96 years, SD = 0.69). On a scale from one to ten (ten being maximally proficient), participants reported a current speaking proficiency of M = 6.76 (SD = 1.00) in Spanish. Further, they classified their comprehension proficiency with M = 7.34 (SD = 0.92) and finally, reading proficiency with M = 7.18 (SD = 1.07). On a scale from zero to ten (ten being maximally exposed), participants quantified their exposure to Spanish at the time of testing with M = 5.20 (SD = 2.48). This compares to an exposure of M = 3.12 (SD = 2.31) to their first foreign language. On a daily basis, this corresponded to an exposure to Spanish of M = 10.03% (SD = 9.48) compared to the other languages. Exposure to Spanish occurred via the following contexts: interaction with Spanish native speakers, listening to Spanish radio shows, watching Spanish television, reading or self-instruction in Spanish. At the time of testing, six participants reported a self-perceived proficiency of Spanish as first non-native language, twenty-six participants as second non-native language, and one participant as third non-native language. We used this metric as a proxy for moderate confidence levels with Spanish.

Noting here that most participants acquired one or more foreign languages prior to acquiring Spanish is important. Research on CLI effects in L3 (L4, …Ln) language processing have demonstrated that all languages within a multilingual system might affect processing in the target language (Lago et al., 2021; Lemhöfer et al., 2004; Rothman, 2015). Here, language dominance was found to be a driving factor of CLI, in that more dominant languages are linked to stronger interference, compared to less dominant languages (Francis and Gallard, 2005; Lago et al., 2021). In our study, a total of eighteen participants reported that they had acquired French, of which fourteen acquired it prior to Spanish. AoA of French was M = 11.38 years of age (SD = 1.98). Accordingly, speakers of French reported a speaking proficiency of M = 2.85 (SD = 0.87), a comprehension proficiency of M = 4.15 (SD = 1.46) and finally, a reading proficiency of M = 4.85 (SD = 1.72) on a scale from one to ten. At the time of testing, exposure to French was reported as M = 0.56 (SD = 2.72) on a scale from zero to ten. Compared to the other languages, participants were exposed to French M = 1.11% (SD = 1.41) on a daily basis. All participants who had acquired French claimed a higher self-reported proficiency for Spanish compared to French, which was reported as third non-native language following Spanish. Therefore, on the basis of previous research (e.g., Lago et al., 2021), we predict only a limited influence of French on CLI effects due to the low dominance and proficiency of speakers in this language at the time of testing. Nevertheless, we included the acquisition of French as a co-variate in our analysis to see whether this had an effect on our results. As discussed in section 3, we did not find an effect on our outcome variables.

2.2. Materials and design

Prior to the experimental session, participants were asked to complete the LEAP-Q. In the laboratory, they completed the Lextale-Esp vocabulary size test and an overt picture naming task. We measured EEG during the picture naming task.

2.2.1. Tasks and stimuli

The Lextale-Esp and the picture naming task were programmed in E-prime 2 (Schneider et al., 2002) and administered on a Windows 10

computer.

Lextale-Esp. We generated an E-prime version of the original Lextale-Esp task with identical instructions and stimuli. This task was used to complement self-reported measures from the LEAP-Q and the vocabulary size score was added as a co-variate to subsequent statistical analyses.

Picture naming task. Our picture stimuli were obtained from the MultiPic picture database (Duñabeitia et al., 2018). We selected the picture stimuli according to two criteria: those with the highest percentage of valid responses given by participants and those with the highest percentage of participants giving the object's exact name. Then, each picture was assigned a gender congruency type (*congruent* versus *incongruent* across German and Spanish) and a cognate status (*cognate* versus *non-cognate* across German and Spanish). The latter was based on the degree of semantic, phonological and orthographic overlap.

We excluded identical cognates [*die Kiwi*] – [*el kiwi*] "the kiwi", plural forms [*die Brille*] – [*las gafas*] "the glasses", professions [*die Sängerin*] – [*la cantante*] "the (female) singer", English loanwords [*der Boomerang*] – [*el boomerang*] "the boomerang", and translation equivalents of opposing genders [*der Esel*] – [*la$_F$ mula$_F$/el$_M$ burro$_M$*] "the donkey". To increase ecological validity of our stimuli, we modelled the distribution of terminal morphemes in Spanish according to previous work by Clegg (2011). Further, we included terminal phoneme of the target noun as a co-variate and item (i.e., the individual picture) as a random effect in the statistical analyses (see sections 3.1.2. and 3.2.2. for more details).

2.2.2. EEG recordings

EEG data were collected via the BrainVision Recorder software (Version 1.23.0001) by Brain Products GmbH. We used an EasyCap electrode cap following a standard 10/20 montage (Appendix B). Data were measured at thirty-two channel locations via passive electrodes. We recorded the horizontal electrooculogram (HEOG) from two electrodes at the outer canthus of the left and right eye. We recorded the vertical electrooculogram (VEOG) from an electrode placed below the left eye. All electrodes were initially referenced to channel Cz, which we later reused as a data channel during re-referencing. The ground electrode was placed on the right cheek of the participant. Impedances of the electrodes were checked and configured using actiCAP Control Software (Version 1.2.5.3) by Brain Products GmbH. We kept impedances below 5 kΩ for the reference and ground electrode. For the remaining channels, impedances were below 10 kΩ. The sampling rate was 500 Hz.

2.3. Procedure

2.3.1. Lextale-Esp

Participants first completed the Lextale-Esp task. For this task, we presented them with a fixation cross at the centre of the screen for 1000 ms. This was followed by the visual presentation of a letter string on the horizontal midline of the screen which corresponded to either a Spanish *word*, or a pronounceable *pseudo-word*. Participants were then asked to indicate via a button-press whether or not the letter string corresponded to a Spanish word. The letter string remained on the screen until the participant responded. Each letter string was only shown once. The total number of trials was 87, because we excluded three trials due to an overlap with the experimental stimuli from the picture naming task prior to the experiment. Offline, we calculated the vocabulary size score by subtracting the percentage of incorrectly identified pseudo-words from the percentage of correctly identified words for each participant (Izura et al., 2014). The maximum score was 100, whereas the minimum score varied as a function of false positives.

2.3.2. Picture naming task

For the picture naming task, we followed a 2 × 2 fully factorial within-subjects design with two main manipulations: *gender congruency* and *cognate status*. Half of the trials were *congruent* in that the gender was

similar across German and Spanish. The other half of trials were *incongruent*, characterised by a dissimilarity in gender across German and Spanish. Further, half of congruent and incongruent pictures were *cognate* words, and the other half were *non-cognate* words (Table 1). There were 24 stimuli per condition, resulting in a total of 96 stimuli.

Following the standard procedure in the field of speech production, the task was divided into a familiarisation phase and an experimental phase, with a total duration of 30–40 min. The familiarisation phase consisted of three rounds. In each round, participants were exposed to all 96 stimuli pictures and were instructed to overtly name each picture in Spanish using an NP construction with the correct determiner and noun (e.g., [*el brazo*] "the arm"). During this phase, the experimenter provided oral feedback on the accuracy of the NP production by the participant whenever necessary. Specifically, the correct determiner or noun was provided in Spanish for cases where either the determiner or the noun, or both, were incorrectly produced by the participant. In the experimental phase, participants named the objects as fast and accurately as possible using a Spanish NP. Participants' EEG and voice were recorded exclusively during the experimental phase. A typical trial was initiated with the display of a fixation cross for 1000 ms, followed by the display of the picture for 2700 ms in the centre of the screen. Each picture was shown only once during the experimental phase, resulting in a total of 96 trials. Trial order was randomized. Participants were reminded throughout the experimental phase to name the object as fast and accurately as possible and to reduce all unnecessary movement. There was a short break after 50 trials to minimise participants' fatigue.

3. Results

3.1. Behavioural data exclusion

Naming latencies and EEG data for one participant were lost due to a malfunctioning microphone and a subsequent failure during the EEG recording. Further two participants were excluded to match the datasets included in the EEG analysis (see section 3.4. for details). In total, we included 30 datasets in this analysis.

3.2. Behavioural data analysis

We used Praat (Broersma and Weenink, 2019) to calculate naming accuracy and naming latencies for each trial for the picture naming task. Next, we analysed our behavioural data in RStudio Version 1.3.959 (R Core Team, 2019). We employed a single-trial modelling approach using the *lme4* package (Bates et al., 2020) to model our two behavioural outcome variables, *naming accuracy* and *naming latencies*. We modelled naming accuracy using generalized linear mixed models (GLMM) and the *glmer()* function with a binomial distribution. Next, we modelled positively skewed naming latencies using the *glmer()* function in combination with a gamma distribution and the identity link function. Only correct trials were included in our analysis of the naming latencies. For both outcome variables, we generated the most theoretically plausible maximal model on the basis of our hypotheses and our two main manipulations, *gender congruency* and *cognate status*. To preserve statistical power and to control for potential confounds, we added *familiarisation phase performance* as a co-variate to our statistical analysis, rather than excluding trials where errors were made before the experimental phase. Similarly, we added *Lextale-Esp score, target noun gender, word length*,

Table 1
Sample set of stimuli for the picture naming task.

Gender congruency			congruent	incongruent
Cognate status	**cognate**	German	die$_F$ Giraffe$_F$	die$_F$ Melone$_F$
		Spanish	la$_F$ jirafa$_F$	el$_M$ melón$_M$
	non-cognate	German	der$_M$ Arm$_M$	der$_M$ Schlüssel$_M$
		Spanish	el$_M$ brazo$_M$	la$_F$ llave$_F$

order of acquisition of Spanish, *acquisition of French* and *terminal phoneme* as co-variates. Further, we included *subject* and *item* as random effects. To establish the model of best fit for the picture naming task, we followed a top-down model selection procedure by testing for the significance of each factor (Barr, 2013; Bates et al., 2018). In order to balance Type I error and power, random effects were chosen as maximal as possible while avoiding overfitting (Matuschek et al., 2017). In the case of non-convergence or singular fit, we simplified our model structure by removing first interactions for random slopes; second, correlations between random slopes; and finally, interactions between fixed effects. We used treatment coding as our contrast, which defaulted to congruent trials and cognates as the reference level. Absolute t-values greater than 1.96 were interpreted as statistically significant at $\alpha = 0.05$ (Alday et al., 2017). Next, we performed model comparisons using the *anova()* function based on the Akaike's Information Criterion, AIC (Akaike, 1974), the Bayesian Information Criterion, BIC (Neath and Cavanaugh, 2012) and the log-likelihood ratio. To perform model diagnostics, we checked the model fit by plotting the model residuals against the predicted values.

3.3. Lextale-Esp

The mean Lextale-Esp score was $M = 18.45$ ($SD = 20.52$). Scores were highly variable and ranged between -23 and 60, with 100 being the maximum score. Vocabulary scores of 60–80 on this task were previously associated with C1–C2 proficiency levels (Lemhöfer and Broersma, 2012), therefore all of our speakers fell below the B2 proficiency range.

3.4. Picture naming task

We first calculated descriptive statistics for naming accuracy and naming latencies for each condition (Table 2).

3.4.1. Naming accuracy
For naming accuracy, our model of best fit included main effects for *gender congruency* and *cognate status*, with *subject* and *item* as random effects. Moreover, the co-variates *Lextale-Esp score* and *familiarisation phase performance* were included in the final model (Appendix C). The remaining co-variates *target noun gender, word length, order of acquisition, acquisition of French* and *terminal phoneme* resulted in non-convergence or singular fit and were therefore excluded from the model fitting procedure. As predicted, participants were marginally more accurate for *congruent* trials compared to *incongruent* trials with $\beta = -0.452$, $SE = 0.232$, $z = -1.95.$, $p = 0.052$. Despite being included in the model of best fit, *cognates* were not significantly different from *non-cognates* with $\beta = -0.279$, $SE = 0.233$, $z = -1.19$, $p = 0.232$ (Fig. 2).

3.4.2. Naming latencies
For naming latencies, our model of best fit included a main effect for *gender congruency* as well as a random effect for *subject* and *item* and a by-subject random slope for gender congruency. *Cognate status* did not significantly improve the model fit and was dropped from the model fitting procedure. Further, *familiarisation phase performance* was included as a co-variate (Appendix D). *Lextale-Esp score, target noun*

Table 2
Mean naming accuracy and naming latencies (only correct trials included) for each condition.

Condition	Naming accuracy (%)		Naming latency (ms)	
	Mean	SD	Mean	SD
congruent/cognate	92.1	27.0	891	237
congruent/non-cognate	86.4	34.3	933	294
incongruent/cognate	87.2	33.4	971	313
incongruent/non-cognate	75.8	42.8	978	303

gender, word length, order of acquisition, acquisition of French and *terminal phoneme* resulted in non-convergence or singular fit and were not included in the subsequent model fitting procedure. Participants were significantly faster in naming *congruent* items compared to *incongruent* items with $\beta = 0.059$, $SE = 0.028$, $z = 2.04$, $p = 0.041$ (Fig. 2).

3.5. EEG data results

3.5.1. EEG data exclusion
The EEG data from the same participant where we lost the voice recordings was also lost during the EEG data acquisition process. Further, we determined a set of criteria to include data in the EEG analyses. First, we only included trials where the correct NP was produced. Second, only correct trials not contaminated by artefacts (valid trials) were analysed. Finally, we set the inclusion threshold for correct and valid trials at 60%. As a result, two additional data sets were excluded due to excess artefact contamination. See Appendix E for rejection rates by condition.

3.5.2. EEG data analysis
Articulatory artefacts pose a challenge when examining EEG data from word production tasks since they may contaminate the signal (Ganushchak et al., 2011a; Grözinger et al., 1975; Porcaro et al., 2015). Therefore, we applied a vigorous pre-processing procedure to separate the signal from artefacts using BrainVision Analyser 2.2. The pre-processing procedure included the following steps: visual inspection of the raw data, re-referencing from Cz to the average mastoid electrodes (TP9 and TP10) and reusing Cz as a data channel, filtering between 0.1 Hz and 30 Hz, linear derivation of the two HEOG electrodes to form a combined channel for horizontal eye movements, interpolation of noisy channels, ocular correction ICA using VEOG and HEOG parameters, and finally, artefact rejection. After pre-processing our EEG data, we added a unique voice onset (VO) marker to every correct trial to mark the articulation onset for each participant. We then generated segments around the picture onset markers and the VO markers for each participant from -200 ms prior to picture onset to 1200 ms after picture onset. Following segmentation, we applied baseline correction using the 200 ms prior to picture onset until picture onset. A novelty of our statistical analysis was the implementation of single-trial linear mixed effects models (LMM) for our EEG data (Frömer et al., 2018). For this, we exported all available voltage samples from valid segments for statistical analysis in RStudio (R Core Team, 2019). In contrast to more traditional EEG analyses involving ANOVAs, the assumptions for single-trial LMM do not include equal number of observations for each participant or uniform effects for each participant. Instead, single-trial LMM capture by-subject and by-item variance and therefore have superior explanatory power over more traditional ANOVAs when modelling EEG data (Baayen et al., 2008; Fröber et al., 2017).

After exporting our EEG data, we performed a permutation test to tentatively explore the locus of the effect of *gender congruency* and *cognate status* (collapsed into the variable *condition*) on voltage amplitudes. We used the *permutes* package (Voeten, 2019) to calculate F-values across all electrodes and the entire available time window between -200 ms and 1200 ms with respect to stimulus onset. Visual inspection of the outcome of the permutation test revealed potential modulatory effects of condition in centro-parietal areas between 350 ms and 600 ms post stimulus onset (Fig. 3). Previous literature on the distribution and time correlates of both the P300 and the N400 support this outcome (Barry et al., 2020; Koester and Schiller, 2008; Paolieri et al., 2020; Peeters et al., 2013; Polich, 2007; Roelofs et al., 2016). Due to increased articulatory artefacts in EEG data closer to the participant's articulatory onset, we only explored the EEG data up to a maximum of 600 ms post-stimulus onset (Porcaro et al., 2015).

On the basis of the outcomes of the permutation test and previous literature, we defined nine topographic areas for our data channels. Along the anterior-posterior axis, we defined anterior, central and
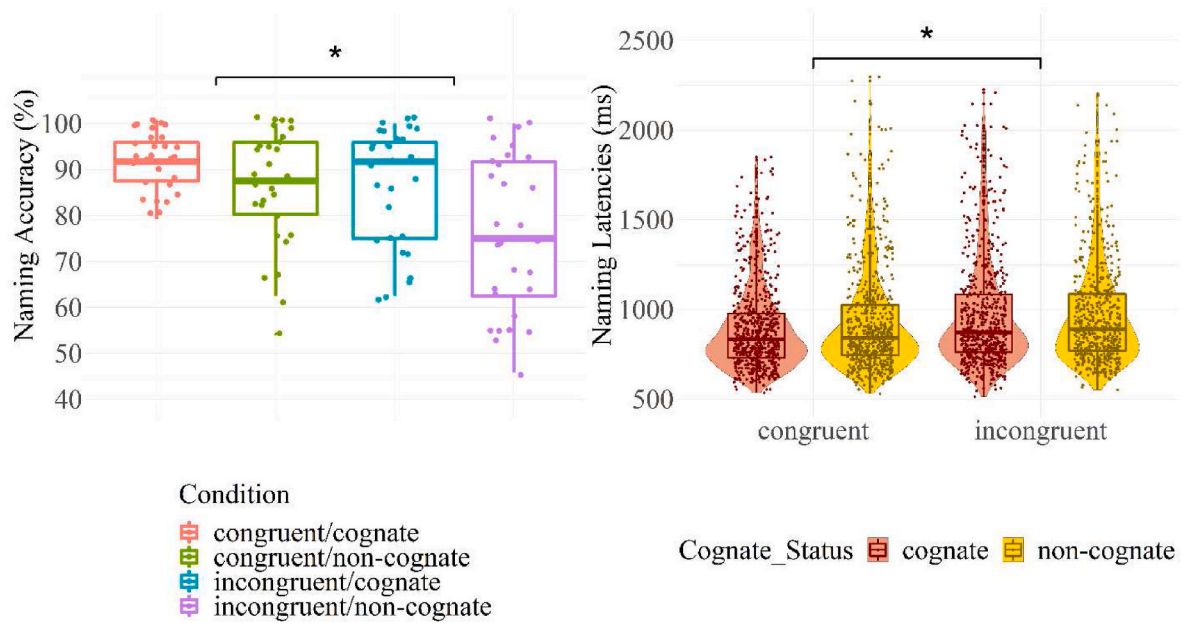
Fig. 2. Mean naming accuracy by subject and condition (left) and naming latencies by condition (right) for the picture naming task (n = 30).
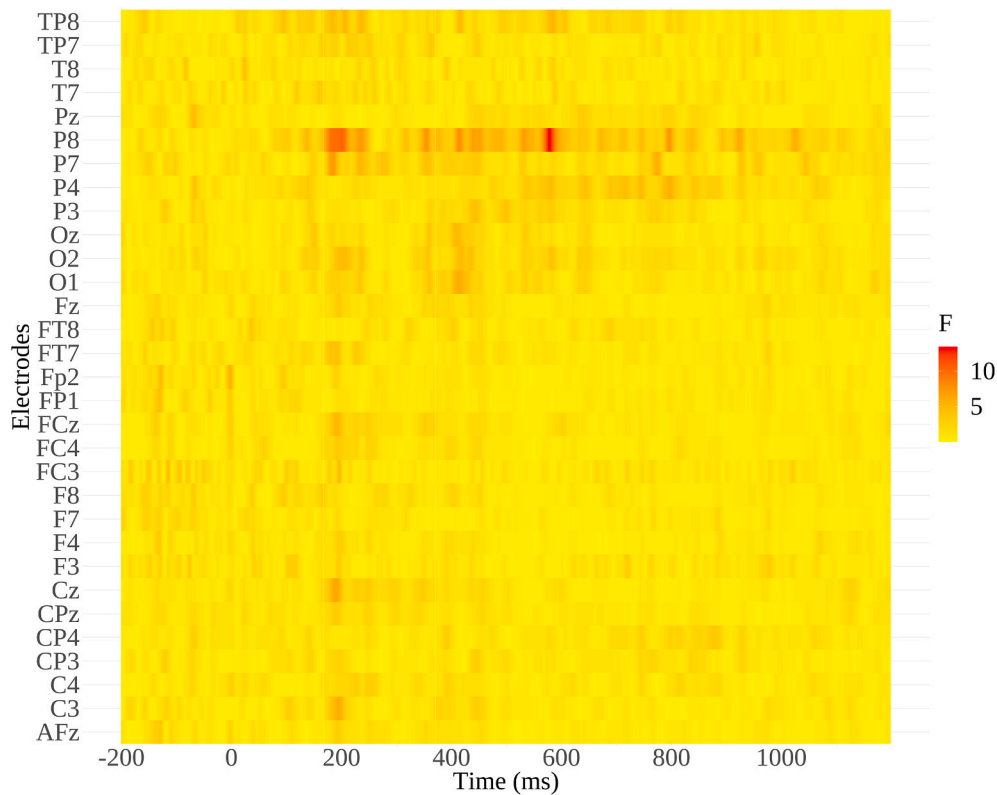

Fig. 3. *Permutation test across all data electrodes for the time window between -200 and 1200 ms post-stimulus onset.*

posterior regions. Each region was further divided into three smaller regions: anterior left, anterior midline and anterior right regions; central left, central midline and central right regions; and finally, posterior left, posterior midline and posterior right regions. In line with previous research on the P300 and the N400, we were particularly interested in a broader topography including the posterior left, posterior midline and posterior right regions between 350 ms and 600 ms post-picture onset. These channels included TP8, P8, O2, Oz, O1, P7, TP7, CP4, P4, Pz, P3, CP3 and CPz (Appendix B). In the statistical analysis, we modelled

modulating effects of *condition* on *voltage amplitudes* between 350 ms and 600 ms post-stimulus onset. Furthermore, we controlled for confounding effects of *hemisphere, Lextale-Esp score, familiarisation phase performance, target noun gender, word length, order of acquisition, acquisition of French* and *terminal phoneme.*

3.5.3. EEG results

Visual inspection of the voltage amplitudes for the selected channels revealed the characteristic P1/N2 complex for early visual processing

(Cheng et al., 2010; Eulitz et al., 2000; Misra et al., 2012; Schendan and Kutas, 2003). Further, visual inspection also revealed a positive-going wave between 350 ms and 600 ms, consistent with the topographic distribution of a P300 (Barry et al., 2020). Further, Fig. 4 shows a by-condition modulation of the EEG signal between 350 and 600 ms, as tentatively suggested in the permutation test (Fig. 3). Descriptively speaking, we saw the largest amplitudes for congruent cognates with *M* = 5.14 (*SD* = 9.29), followed by incongruent cognates with *M* = 5.05 (*SD* = 9.24), congruent non-cognates with *M* = 4.88 (*SD* = 8.94), and finally, incongruent non-cognates with *M* = 4.47 (*SD* = 9.04) in the 350 ms–600 ms time window. We found no indication for an N400 effect prior to the 600 ms, after which the signal becomes increasingly noisy due to the proximity to the articulatory onset. See Fig. 5 for a visualisation of the individual channels included in this analysis.

The model of best fit for voltage amplitudes included an interaction effect of *gender congruency* and *cognate status*. Further, *hemisphere* and *familiarisation phase performance* were included as co-variates (Appendix F). *Lextale-Esp score, target noun gender, word length, order of acquisition, acquisition of French* and *terminal phoneme* did not significantly improve the model fit or led to over-fitting. More specific to the co-variate of *acquisition of French*, the model comparison between the model with and without this particular co-variate yielded $\chi^2(1, 30) = 0.018$, $p = 0.893$. We therefore dropped *acquisition of French* from the model selection procedure. In the model of best fit, i*tem* and *subject* emerged as random effects, with a by-subject random slope for the interaction effect of *gender congruency* and *cognate status*. Voltage amplitudes were more positive for *congruent cognate* nouns compared to *incongruent non-cognates* with $\beta = -0.684$, $SE = 0.342$, $t = -2.002$, $p = 0.045$. The difference in amplitude between the remaining conditions was not significant.[3]

4. Discussion

The aim of this study was twofold: first, we examined CLI of the gender systems and phonological systems to obtain a better characterization of the time course of non-native NP production. Secondly, we explored which production stage is associated with the selection of the target language in a multilingual language production configuration.

We studied the gender congruency effect to highlight CLI of the gender systems during *lexical retrieval*. We predicted higher naming accuracy, shorter naming latencies, less positive P300 amplitudes and less negative N400 amplitudes for congruent compared to incongruent nouns. Critically, this would indicate that the target language was not selected prior to *lexical retrieval*. We also explored the cognate facilitation effect to illustrate CLI during *phonological encoding* and expected higher naming accuracy, shorter naming latencies, less positive P300 amplitudes and less negative N400 voltage amplitudes for cognates compared to non-cognates. The presence of a cognate facilitation effect would imply that the lexical entries from both the target and non-target language actively competed during *phonological encoding*, placing the locus of target language selection beyond *lexical retrieval* (Christoffels et al., 2007; Colomé, 2001; Hoshino and Kroll, 2008; Hoshino and

Thierry, 2011; Peterson and Savoy, 1998; Pulvermüller et al., 2009; Rodriguez-Fornells et al., 2005).

In line with our predictions, we found that participants were significantly more accurate and faster at naming congruent nouns compared to incongruent nouns. These behavioural findings are important for two reasons: First, the presence of the gender congruency effect suggests CLI during gender processing. Second, the gender congruency effect also implies that the target language was not selected before *lexical retrieval*. Yet, results from the gender congruency effect alone cannot clarify whether CLI continued beyond gender processing. Therefore, we complemented these findings with results from the cognate facilitation effect. Despite a clear descriptive trend, we found no evidence for a cognate facilitation effect at the behavioural level – in contrast to previous research on the cognate status in non-native production (Acheson et al., 2012; Christoffels et al., 2007; Peeters et al., 2013). There are two possible interpretations of this outcome: first, our late language learners did not face CLI during *phonological encoding*. As a result, there were no detectable processing differences between cognates and non-cognates. Critically, this would imply that CLI may be resolved prior to *phonological encoding* and that only the lexical entry from the target language is phonologically encoded. The second interpretation is that the behavioural measures lacked the power to pick up on a fine-grained modulation based on cognate status. Our EEG data are able to discriminate between these two possible interpretations.

Despite clear evidence for a P300 effect, we did not find evidence for an N400 effect in the time window of interest. This is somewhat surprising given that previous research linked the N400 to language co-activation, and to the neural correlates of the gender congruency effect and the cognate facilitation effect (Chen et al., 2017; Paolieri et al., 2020). However, studies also showed a reduced or delayed N400 in speakers with lower proficiency levels (Heidlmayr et al., 2021; Midgley et al., 2009; Weber-Fox and Neville, 1996). Therefore, the N400 effect may have been absent, or delayed and masked by articulatory artefacts.

Regarding the P300, its topographic characteristics are in line with a P300 component in the time window between 350 ms and 600 ms, more specifically a P3b component (Barry et al., 2020; Hruby and Marsalek, 2003; Polich, 2007; Squires et al., 1975). The P300, in particular the P3b, has been linked to classical inhibitory tasks as well as inhibitory tasks combined with a linguistic task (Bosma and Pablos, 2020; Eriksen and Eriksen, 1974; Jiao et al., 2020; Soares Pereira et al., 2019). Critically, it was proposed to reflect general cognitive mechanisms such as inhibition, conflict resolution and cognitive interference, and more recently the recruitment and allocation of attentional resources and working memory load (Barker and Bialystok, 2019; González Alonso et al., 2020; Neuhaus et al., 2010a,b; Polich, 2007, 2012; Wu and Thierry, 2013). In order to successfully produce the correct NP in the target language, speakers not only had to go through the multi-stage process of language production, but had to simultaneously mitigate CLI effects between the target and non-target language. Here, we argue that the P300 directly taps into this latter notion and that it provides an index for this ongoing conflict between the target and the non-target language. Our EEG data revealed a small, but robust by-condition modulation of P300 voltage amplitudes, reflected in the interaction effect of gender congruency and cognate status. As predicted, P300 amplitudes were significantly different for trials with high processing costs and a larger involvement of the inhibitory control system, i.e., incongruent non-cognate trials compared to congruent cognate trials. Therefore, our results suggest quantitatively different neural patterns for producing NPs subject to differential processing costs and inhibitory demands. More importantly, this modulation of P300 amplitudes appeared to last until 600 ms post-stimulus onset (and possibly beyond). This notion has direct implications for the time course of non-native production because it provides a clear time frame for the cognitive mechanisms underlying the mitigation of CLI.

Our EEG results were indicative of the following: first, our speakers faced CLI both during the processing of gender and cognate status.
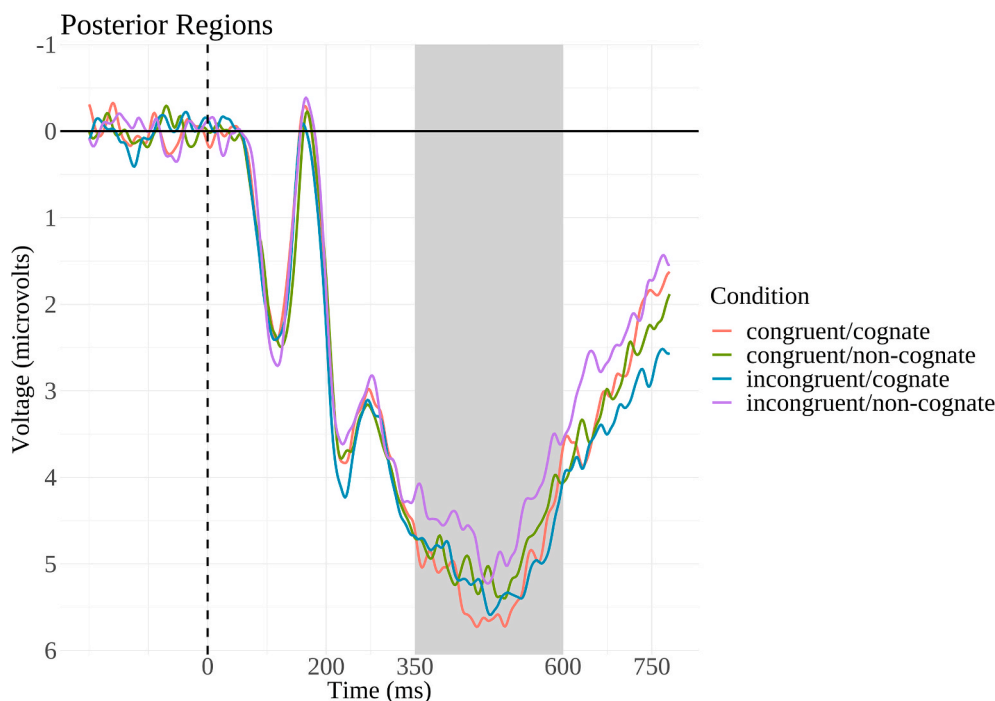
---

[3] As per suggestion of a reviewer, we also explored left anterior negativity (LAN) effects as a function of condition. Based on previous literature (Barber and Carreiras, 2005; Friederici et al., 1999; Hahne and Friederici, 1999; Steinhauer et al., 2009; Valente et al., 2016; Weber-Fox and Neville, 1996), we determined channels Fp1, F3, F7, FC3 and FT7 in left anterior regions as our ROI, and the time-window of interest between 300 ms and 500 ms post stimulus onset. The data show a negative-going wave peaking at around 450 ms post-stimulus onset, consistent with the topography of a delayed LAN. However, there seemed to be little difference between the conditions in terms of LAN voltage amplitudes. This was confirmed in our statistical analysis: The model that contained condition as fixed effect was not significantly better than the model that did not contain condition (χ^2 (3, 30) = 0.072, $p = 0.995$). We therefore found no evidence for a by-condition modulation of the LAN in this particular study.

Fig. 4. *Voltage amplitudes by condition over time for channels TP8, P8, O2, Oz, O1, P7, TP7, CP4, P4, Pz, P3, CP3 and CPz for the picture naming task (n = 30). The time window of interest from* 350 ms *to* 600 ms *is highlighted in grey. Negativity is plotted up.*

Secondly, in line with the findings by Hoshino and Thierry (2011) on the locus of target language selection, CLI appears to continue beyond gender processing until at least *phonological encoding* in late language learners. This finding favours the interpretation that target language selection takes place after gender processing. To the best of our knowledge, this is the first study to report a P300 effect during overt non-native NP production in a paradigm that was not explicitly about inhibitory control, but instead included an implicit inhibitory control component. Similar tentative EEG results are reported by González-Alonso et al. (2020) within the framework of third language acquisition of artificial mini-grammars.

An interesting feature of the P300 effect was the elicitation of more positive amplitudes for congruent and cognate nouns compared to incongruent and non-cognate nouns. This is in contrast to our original hypothesis, where we predicted less positive amplitudes for congruent and cognate nouns compared to incongruent and non-cognate. Notably, this particular pattern of behavioural and EEG results has been previously reported in the literature in connection to the cognate facilitation effect (Acheson et al., 2012; Christoffels et al., 2007). For example, Acheson et al. (2012) used an overt picture naming task with unbalanced German-Dutch speakers to study conflict monitoring during bilingual language production with respect to the Error-Related Negativity (ERN). They found faster naming latencies for cognates compared to non-cognates. However, this was linked to more negative amplitudes for cognates from about 150 ms post-stimulus onset at the FCz electrode. Furthermore, Jiao et al. (2020) measured EEG during a picture naming task combined with a flanker task in unbalanced Chinese-English bilinguals. Their ERP results showed more positive P300 amplitudes for congruent compared to incongruent flankers in centro-parietal regions, while response times for congruent flanker trials were shorter compared to incongruent flanker trials (see also Bosma and Pablos, 2020). These results mirror those from our study. On the other hand, studies have also supported the more traditional notion of faster response times or shorter naming latencies in combination with smaller ERP amplitudes for cognates compared to non-cognates (Comesaña et al., 2012; Peeters et al., 2013; Strijkers et al., 2010; Xiong et al., 2020) and smaller P300 amplitudes for congruent trials in the flanker task (Wu and Thierry, 2013).

Therefore, our behavioural and EEG patterns are by no means unusual. Instead, they suggest a clear involvement of inhibitory control and acutely reflect the critical processes linked to successful non-native NP production. We propose that this study highlights the significance of the P300 as an index for the cognitive processes underlying the mitigation of CLI and the selection of the target language. Nevertheless, given the scarcity in terms of research, the directionality of the P300 effect elicited during non-native NP production warrants closer inspection in the future.

Taken together, we found traceable effects of CLI both at the behavioural and at the neural level, establishing CLI as a significant modulator of the time course of non-native NP production. This has implications for the time course of the production processes in non-native NP production, as reflected in naming accuracy, naming latencies and ERP patterns with respect to gender processing and phonological processing. CLI acts both as a facilitator and a hindrance during the production process: on one hand, there is a processing advantage for congruent and cognate nouns. On the other hand, this appears to be less the case for incongruent and non-cognate nouns. Moreover, our findings suggest that late language learners not only face CLI during early production stages and *lexical retrieval*, but possibly also during later phonological processing stages of *phonological encoding*.

In turn, this implies that lexical entries from both the target and non-target language are selected for phonological processing, thereby shifting the locus of target language selection until *phonological encoding* or after it. Arguably, this highlights the complexity of non-native production processes compared to the production process in native-like speakers. Given the design of our study, we cannot exclude the possibility that our speakers resolve CLI between target and non-target language at even later production stages, e.g., the *phonetic encoding* stage. Yet, our findings have important implications for characterising the theoretical and neural underpinnings of the time course of non-native production processes, in particular for speakers with intermediate proficiency levels. Further, our findings add novel evidence to the debate about the locus of target language selection in late language learners.
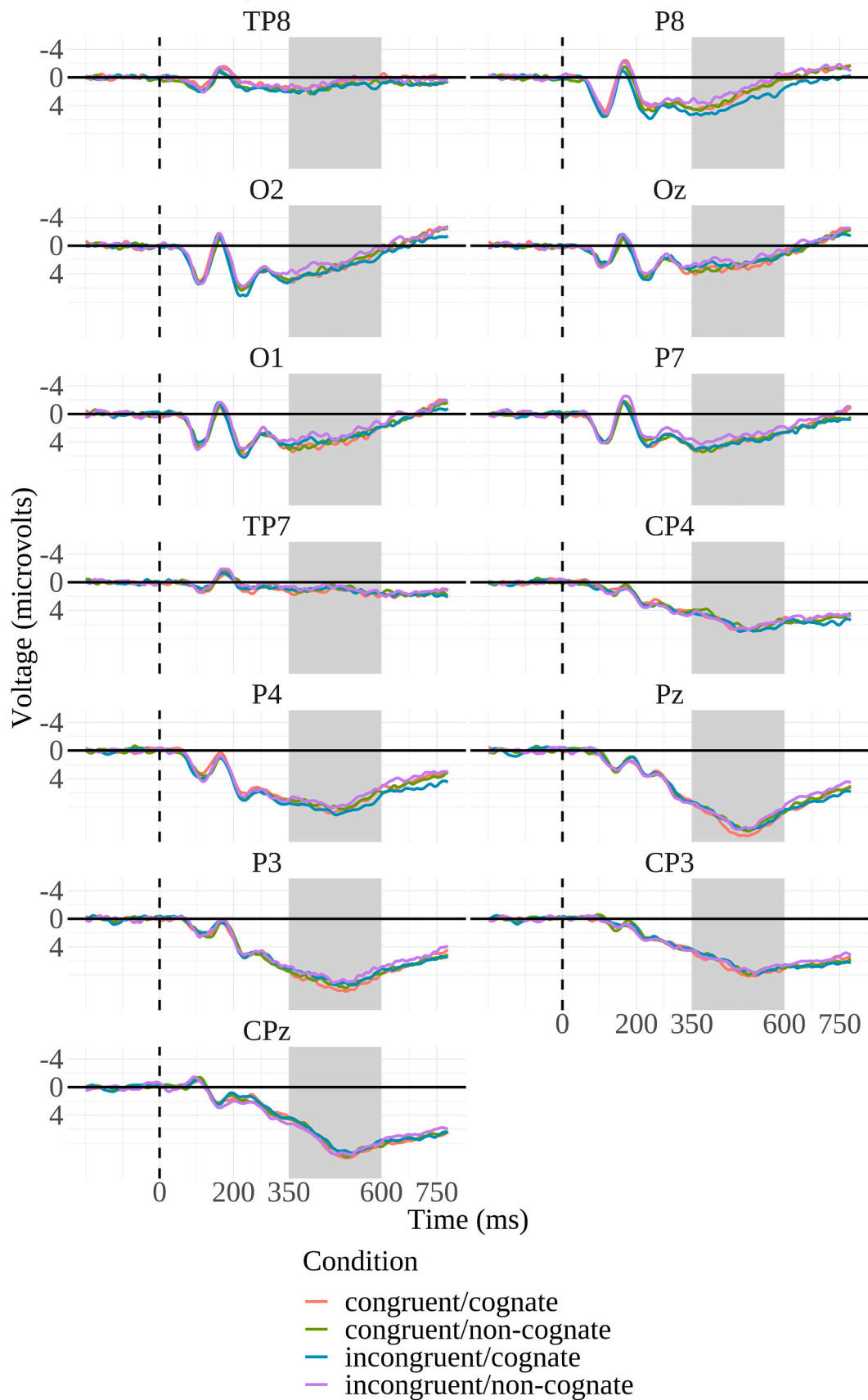
Fig. 5. Voltage amplitudes by condition over time for each individual electrode included in the analysis of the picture naming task (n = 30). The time window of interest from 350 ms to 600 ms is highlighted in grey. Negativity is plotted up.

5. Conclusion

In this study, we found traceable CLI effects at the behavioural and the neural level. More specifically, speakers faced CLI during gender processing and during phonological processing, which in turn impacted the time course of non-native production. In terms of the locus of target language selection, our findings suggested that the target and non-target language remained active at least until *phonological encoding.* Our findings have important theoretical implications for the conceptualisation of non-native production mechanisms, and warrant further exploration with regard to subsequent production stages and the exact involvement of inhibitory control in non-native NP production. Finally, we argue that there should be an increased focus on both the P300 component as an index of CLI and lower proficiency levels in studies on non-native NP production.

6. Citation diversity statement

To shed light on the systematic underrepresentation of work by female scientists and scientists identifying as members of a minority compared to the papers published in the field, we included a Citation diversity statement (Dworkin et al., 2020; Rust and Mehrpour, 2020; Torres et al., 2020; Zurn et al., 2020). For this, we classified the first and last author in each paper from our reference list based on their preferred gender (wherever information was available).

Our reference list consisted of 26% woman/woman authors, 38% man/man, 21% woman/man and finally, 13% man/woman authors. This compares to 6.7% for woman/woman, 58.4% for man/man, 25.5% woman/man, and lastly, 9.4% for man/woman authored references for the field of neuroscience (Dworkin et al., 2020). A clear limitation of this classification is the rather broad binary woman/man distinction. However, we have full confidence that the classification system will vastly improve in the future with the routine addition of the preferred gender to personal and academic websites.

Credit author contribution statement

Sarah von Grebmer zu Wolfsthurn: Conceptualisation, Methodology, Programming, Data Acquisition, Formal Analysis, Data Curation, Writing-Original draft preparation, Writing-Review and Editing. **Leticia Pablos**: Conceptualisation, Methodology, Data Curation, Writing-Review and Editing, Supervision. **Niels O. Schiller**: Conceptualisation, Writing-Review and Editing, Supervision, Funding Acquisition
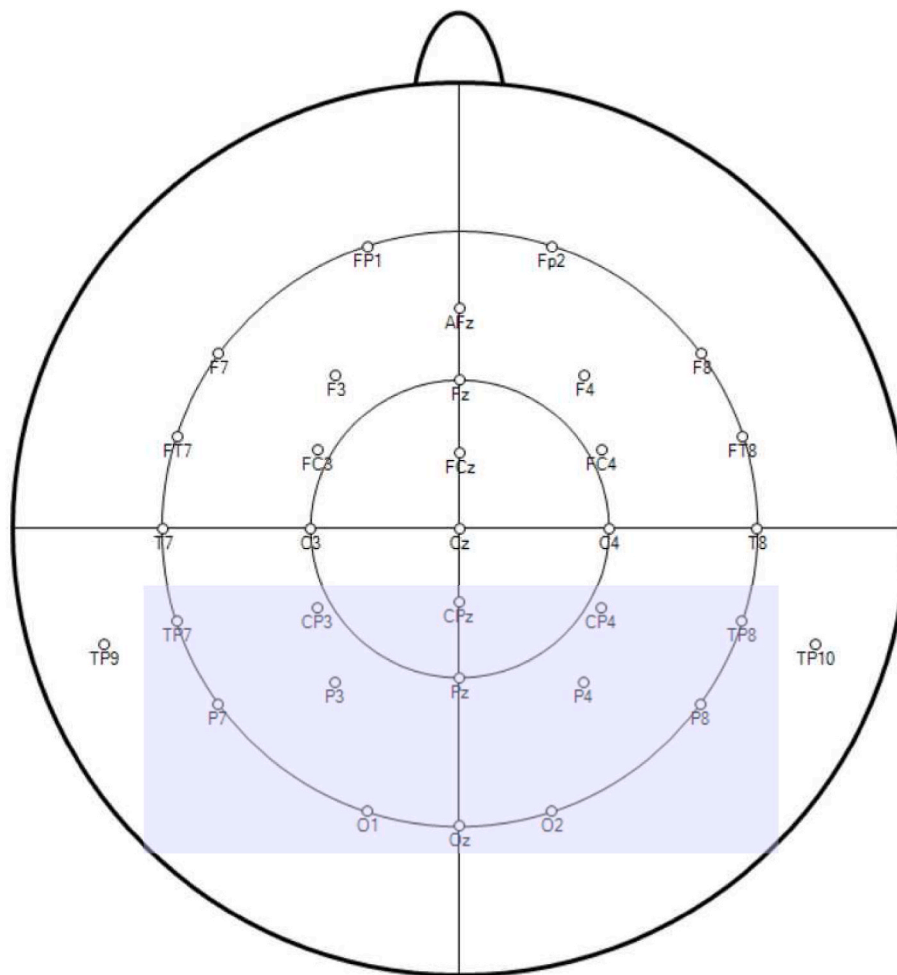
Appendix

Appendix A

Overview of the native and non-native languages acquired by the participants of the current study (N = 33) according to the LEAP-Q (Marian et al., 2007).

Language	Native language	First foreign language	Second foreign language	Third foreign language	Fourth foreign language	Total
German	n = 33					**33**
Spanish			n = 16	n = 15	n = 2	**33**
English		n = 31	n = 2			**33**
French		n = 2	n = 11	n = 5		**18**
Latin			n = 3	n = 1	n = 1	**5**
Russian			n = 1		n = 1	**2**
Swedish				n = 1		**1**
Portuguese					n = 2	**2**
Arabic					n = 1	**1**
Catalan					n = 1	**1**
Italian					n = 1	**1**
Mandarin					n = 1	**1**
Total	**33**	**33**	**33**	**22**	**10**	

Appendix B

Electrode positions following a 10/20 montage. Electrodes included in the analysis are highlighted in purple.

Appendix C

Model of best fit for Naming Accuracy, including estimated means, confidence intervals errors and z-values.

Formula: naming accuracy ~ gender congruency (congruent vs. incongruent) + cognate status (cognate vs. non-cognate) + Lextale-Esp score + familiarisation phase performance (none correct vs. one correct vs. two correct vs. three correct) + (1|subject) + (1|item)

Fixed effects	Estimate	95% CI [low – high]	z-value	Pr (>|z|)
(Intercept)	−0.698	−1.38–−0.010	−1.99	0.047
Gender congruency incongruent	−0.452	−0.908–0.003	−1.95	**0.052***
Cognate status Non-cognate	-0–279	−0.736–0.179	−1.19	0.233
Lextale-Esp score	0.022	0.006–0.038	2.72	**0.007****
Familiarisation phase performance: one correct	1.98	1.52–2.44	8.44	**< 0.001*****
Familiarisation phase performance: two correct	3.14	2.94–3.90	13.91	**< 0.001*****
Familiarisation phase performance: three correct	4.26	3.75–4.78	16.14	**< 0.001*****
Random effects				
σ^2	3.29			
$\tau_{00 \text{ Item}}$	0.66			
$\tau_{00 \text{ Subject}}$	0.53			
ICC	0.27			
N_{Subject}	30			
N_{Item}	96			
Observations	2880			
Marginal R^2/Conditional R^2	0.341/0.516			

Appendix D

Model of best fit for Naming Latencies, including estimated means, confidence intervals and z-values.

Formula: naming latency ~ gender congruency (congruent vs. incongruent) + familiarisation phase performance (none correct vs. one correct vs. two correct vs. three correct) + (gender congruency|subject) + (1|item)

Fixed effects	Estimate	95% CI [low – high]	z-value	Pr (>|z|)
(Intercept)	1.43	1.32–1.55	25.04	<0.001
Gender congruency incongruent	0.059	0.002–0.116	2.04	**0.041***
Familiarisation phase performance: one correct	−0.180	−0.267––0.092	−4.02	**< 0.001***
Familiarisation phase performance: two correct	−0.419	−0.501––0.337	−9.98	**< 0.001***
Familiarisation phase performance: three correct	−0.495	−0.577––0.412	−11.76	**< 0.001***
Random effects				
σ^2	0.05			
$\tau_{00\ Item}$	0.00			
$\tau_{00\ Subject}$	0.00			
$\tau_{11\ Subject.Congruency\_Typeincongruent}$	0.00			
$\rho_{01\ Subject}$	−0.14			
ICC	0.16			
$N_{Subject}$	30			
N_{Item}	96			
Observations	2459			
Marginal R^2/Conditional R^2	0.184/0.312			

Appendix E

Rejection rates for each condition for the EEG data of the picture naming task for n = 30.

Condition	Rejection rate (%)	Rejected trials	Total valid trials
congruent/cognate	3.47	23	663
congruent/non-cognate	4.02	25	622
incongruent/cognate	3.34	21	628
incongruent/non-cognate	4.95	27	546
Average per condition	**3.94**	**24**	**614.75**
Total across conditions	**3.90**	**96**	**2459**

Appendix F

Model of best fit for Voltage Amplitudes, including estimated means, confidence intervals and t-values.

Formula: voltage amplitudes ~ gender congruency (congruent vs. incongruent) * cognate status (cognate vs. non-cognate) + hemisphere (left vs. midline vs. right) + familiarisation phase performance (none correct vs. one correct vs. two correct vs. three correct) + (gender congruency * cognate status|subject) + (1|item)

Fixed effects	Estimate	95% CI [low – high]	t-value	Pr (>|t|)
(Intercept)	5.27	4.29–6.25	10.84	<0.001
incongruent/cognate	−0.075	−0.810–0.660	−0.203	0.839
congruent/non-cognate	−0.254	−0.923–0.415	−0.751	0.453
incongruent/non-cognate	−0.684	−1.36––0.008	−2.002	**0.045***
Hemisphere midline	1.98	1.96–2.01	170.06	**< 0.001***
Hemisphere right	−0.710	−0.730––0.691	−70.30	**< 0.001***
Familiarisation phase performance: one correct	−0.106	−0.180––0.033	−2.85	**< 0.001***
Familiarisation phase performance: two correct	−0.026	−0.043–0.095	−0.742	0.458
Familiarisation phase performance: three correct	−0.333	−0.402––0.263	−9.35	**< 0.001***
Random effects				
σ^2	76.06			
$\tau_{00\ Item}$	1.01			
$\tau_{00\ Subject}$	5.80			
$\tau_{11\ Subject.Conditioncogn\_incongr}$	1.60			
$\tau_{11\ Subject.Conditionnoncogn\_incongr}$	0.90			
$\tau_{11\ Subject.Conditionnoncogn\_congr}$	0.97			
$\rho_{01\ Subject.Conditionnoncogn\_congr}$	−0.33			
$\rho_{01\ Subject.Conditioncogn\_incongr}$	−0.39			
$\rho_{01\ Subject.Conditionnoncogn\_incongr}$	−0.42			
ICC	0.08			
$N_{Subject}$	30			

(*continued*)

Formula: voltage amplitudes ~ gender congruency (congruent vs. incongruent) * cognate status (cognate vs. non-cognate) + hemisphere (left vs. midline vs. right) + familiarisation phase performance (none correct vs. one correct vs. two correct vs. three correct) + (gender congruency * cognate status|subject) + (1|item)				
Fixed effects	Estimate	95% CI [low – high]	t-value	Pr (>|t|)
N $_{Item}$	96			
Observations	3,873,870			
Marginal R^2/Conditional R^2	0.014/0.089			

References

Abdel Rahman, R., Sommer, W., 2008. Seeing what we know and understand: how knowledge shapes perception. Psychonomic Bull. Rev. 15 (6), 1055–1063. https://doi.org/10.3758/PBR.15.6.1055.

Abutalebi, J., Green, D., 2007. Bilingual language production: the neurocognition of language representation and control. J. Neurolinguistics 20 (3), 242–275. https://doi.org/10.1016/j.jneuroling.2006.10.003.

Acheson, D.J., Ganushchak, L.Y., Christoffels, I.K., Hagoort, P., 2012. Conflict monitoring in speech production: physiological evidence from bilingual picture naming. Brain Lang. 123 (2), 131–136. https://doi.org/10.1016/j.bandl.2012.08.008.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Control 19 (6), 716–723.

Alario, F.-X., Caramazza, A., 2002. The production of determiners: evidence from French. Cognition 82 (3), 179–223. https://doi.org/10.1016/S0010-0277(01)00158-5.

Alday, P.M., Schlesewsky, M., Bornkessel-Schlesewsky, I., 2017. Electrophysiology reveals the neural dynamics of naturalistic auditory language processing: event-related potentials reflect continuous model updates. ENeuro 4 (6), 1–19. https://doi.org/10.1523/ENEURO.0311-16.2017.

Aristei, S., Melinger, A., Abdel Rahman, R., 2011. Electrophysiological chronometry of semantic context effects in language production. J. Cognit. Neurosci. 23 (7), 1567–1586. https://doi.org/10.1162/jocn.2010.21474.

Baayen, H.R., Davidson, D.J., Bates, D.M., 2008. Mixed-effects modeling with crossed random effects for subjects and items. J. Mem. Lang. 59 (4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005.

Badecker, W., Miozzo, M., Zanuttini, R., 1995. The two-stage model of lexical retrieval: evidence from a case of anomia with selective preservation of grammatical gender. Cognition 57 (2), 193–216. https://doi.org/10.1016/0010-0277(95)00663-J.

Barber, H.A., Carreiras, M., 2005. Grammatical gender and number agreement in Spanish: an ERP comparison. J. Cognit. Neurosci. 17 (1), 137–153. https://doi.org/10.1162/0898929052880101.

Barker, R.M., Bialystok, E., 2019. Processing differences between monolingual and bilingual young adults on an emotion n-back task. Brain Cognit. 134, 29–43. https://doi.org/10.1016/j.bandc.2019.05.004.

Barr, D.J., 2013. Random effects structure for testing interactions in linear mixed-effects models. Front. Psychol. 4, 1–2. https://doi.org/10.3389/fpsyg.2013.00328.

Barry, R.J., Steiner, G.Z., Blasio, F.M.D., Fogarty, J.S., Karamacoska, D., MacDonald, B., 2020. Components in the P300: don't forget the novelty P3! Psychophysiology 57 (7), e13371. https://doi.org/10.1111/psyp.13371.

Bates, D.M., Kliegl, R., Vasishth, S., Baayen, H.R., 2018. Parsimonious Mixed Models. ArXiv:1506.04967 [Stat]. http://arxiv.org/abs/1506.04967.

Bates, D.M., Maechler, M., Bolker, B., Walker, S., Christensen, R.H.B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., Fox, J., Bauer, A., Krivitsky, P.N., 2020. Package 'lme4'.

Berthele, R., 2021. Introduction: what's special about multilingualism? Lang. Learn. 71 (S1), 5–11. https://doi.org/10.1111/lang.12436.

Bordag, D., Pechmann, T., 2007. Factors influencing L2 gender processing. Biling. Lang. Cognit. 10 (3), 299–314. https://doi.org/10.1017/S1366728907003082.

Bosch, J.E., Unsworth, S., 2020. Cross-linguistic influence in word order: effects of age, dominance and surface overlap. Linguistic Approaches to Bilingualism 1–34. https://doi.org/10.1075/lab.18103.bos.

Bosma, E., Blom, E., Hoekstra, E., Versloot, A., 2019. A longitudinal study on the gradual cognate facilitation effect in bilingual children's Frisian receptive vocabulary. Int. J. Biling. Educ. BiLing. 22 (4), 371–385. https://doi.org/10.1080/13670050.2016.1254152.

Bosma, E., Pablos, L., 2020. Switching direction modulates the engagement of cognitive control in bilingual reading comprehension: an ERP study. J. Neurolinguistics 55, 100894. https://doi.org/10.1016/j.jneuroling.2020.100894.

Broersma, P., Weenink, D., 2019. Praat: Doing Phonetics by Computer (6.0.49) [Computer software]. https://www.fon.hum.uva.nl/praat/.

Bull, W., 1965. Spanish for Teachers: Applied Linguistics. Ronald Press.

Bultena, S., Dijkstra, T., van Hell, J.G., 2014. Cognate effects in sentence context depend on word class, L2 proficiency, and task. Q. J. Exp. Psychol. 67 (6), 1214–1241. https://doi.org/10.1080/17470218.2013.853090.

Bürki, A., Laganaro, M., 2014. Tracking the time course of multi-word noun phrase production with ERPs or on when (and why) cat is faster than the big cat. Front. Psychol. 5, 1–13. https://doi.org/10.3389/fpsyg.2014.00586.

Bürki, A., Sadat, J., Dubarry, A.-S., Alario, F.-X., 2016. Sequential processing during noun phrase production. Cognition 146, 90–99. https://doi.org/10.1016/j.cognition.2015.09.002.

Camen, C., Morand, S., Laganaro, M., 2010. Re-evaluating the time course of gender and phonological encoding during silent monitoring tasks estimated by ERP: serial or parallel processing? J. Psycholinguist. Res. 39 (1), 35–49. https://doi.org/10.1007/s10936-009-9124-4.

Cárdenas-Hagan, E., Carlson, C.D., Pollard-Durodola, S.D., 2007. The cross-linguistic transfer of early literacy skills: the role of initial L1 and L2 skills and language of instruction. Lang. Speech Hear. Serv. Sch. 38 (3), 249–259. https://doi.org/10.1044/0161-1461(2007/026.

Casaponsa, A., Antón, E., Pérez, A., Duñabeitia, J.A., 2015. Foreign language comprehension achievement: insights from the cognate facilitation effect. Front. Psychol. 6, 1–15. https://doi.org/10.3389/fpsyg.2015.00588.

Chen, P., Bobb, S.C., Hoshino, N., Marian, V., 2017. Neural signatures of language co-activation and control in bilingual spoken word comprehension. Brain Res. 1665, 50–64. https://doi.org/10.1016/j.brainres.2017.03.023.

Cheng, X., Schafer, G., Akyürek, E.G., 2010. Name agreement in picture naming: an ERP study. Int. J. Psychophysiol. 76 (3), 130–141. https://doi.org/10.1016/j.ijpsycho.2010.03.003.

Christoffels, I.K., Firk, C., Schiller, N.O., 2007. Bilingual language control: an event-related brain potential study. Brain Res. 1147, 192–208. https://doi.org/10.1016/j.brainres.2007.01.137.

Clegg, J.H., 2011. A frequency-based analysis of the norms for Spanish noun gender. Hispania 94 (2), 303–319.

Colomé, À., 2001. Lexical activation in bilinguals' speech production: language-specific or language-independent? J. Mem. Lang. 45 (4), 721–736. https://doi.org/10.1006/jmla.2001.2793.

Comesaña, M., Sánchez-Casas, R., Soares, A.P., Pinheiro, A.P., Rauber, A., Frade, S., Fraga, I., 2012. The interplay of phonology and orthography in visual cognate word recognition: an ERP study. Neurosci. Lett. 529 (1), 75–79. https://doi.org/10.1016/j.neulet.2012.09.010.

Corbett, G.G., 1991. Gender. Cambridge University Press. https://doi.org/10.1017/CBO9781139166119.

Costa, A., Caramazza, A., Sebastian-Galles, N., 2000. The cognate facilitation effect: implications for models of lexical access. J. Exp. Psychol. Learn. Mem. Cognit. 26 (5), 1283–1296. https://doi.org/10.1037/0278-7393.26.5.1283.

Costa, A., Heij, W.L., Navarrete, E., 2006. The dynamics of bilingual lexical access. Biling. Lang. Cognit. 9 (2), 137–151. https://doi.org/10.1017/S1366728906002495.

Costa, A., Kovacic, D., Franck, J., Caramazza, A., 2003. On the autonomy of the grammatical gender systems of the two languages of a bilingual. Biling. Lang. Cognit. 6 (3), 181–200. https://doi.org/10.1017/S1366728903001123.

Costa, A., Santesteban, M., 2004. Lexical access in bilingual speech production: evidence from language switching in highly proficient bilinguals and L2 learners. J. Mem. Lang. 50 (4), 491–511. https://doi.org/10.1016/j.jml.2004.02.002.

Costa, A., Santesteban, M., Caño, A., 2005. On the facilitatory effects of cognate words in bilingual speech production. Brain Lang. 94 (1), 94–103. https://doi.org/10.1016/j.bandl.2004.12.002.

Costa, A., Strijkers, K., Martin, C., Thierry, G., 2009. The time course of word retrieval revealed by event-related brain potentials during overt speech. Proc. Natl. Acad. Sci. Unit. States Am. 106 (50), 21442–21446. https://doi.org/10.1073/pnas.0908921106.

Council of Europe, 2001. Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Cambridge University Press.

Duñabeitia, J.A., Crepaldi, D., Meyer, A.S., New, B., Pliatsikas, C., Smolka, E., Brysbaert, M., 2018. MultiPic: a standardized set of 750 drawings with norms for six European languages. Q. J. Exp. Psychol. 71 (4), 808–816. https://doi.org/10.1080/17470218.2017.1310261.

Dworkin, J.D., Linn, K.A., Teich, E.G., Zurn, P., Shinohara, R.T., Bassett, D.S., 2020. The Extent and Drivers of Gender Imbalance in Neuroscience Reference Lists. BioRxiv. https://doi.org/10.1101/2020.01.03.894378, 2020.01.03.894378.

Eddington, D., 2002. Spanish gender assignment in an analogical framework. J. Quant. Ling. 9 (1), 49–75. https://doi.org/10.1076/jqul.9.1.49.8482.

Eriksen, B.A., Eriksen, C.W., 1974. Effects of noise letters upon the identification of a target letter in a nonsearch task. Percept. Psychophys. 16 (1), 143–149. https://doi.org/10.3758/BF03203267.

Eulitz, C., Hauk, O., Cohen, R., 2000. Electroencephalographic activity over temporal brain areas during phonological encoding in picture naming. Clin. Neurophysiol. 111 (11), 2088–2097. https://doi.org/10.1016/S1388-2457(00)00441-7.

Francis, W.S., Gallard, S.L.K., 2005. Concept mediation in trilingual translation: evidence from response time and repetition priming patterns. Psychonomic Bull. Rev. 12 (6), 1082–1088. https://doi.org/10.3758/BF03206447.

Friederici, A.D., Steinhauer, K., Frisch, S., 1999. Lexical integration: sequential effects of syntactic and semantic information. Mem. Cognit. 27 (3), 438–453. https://doi.org/10.3758/BF03211539.

Fröber, K., Stürmer, B., Frömer, R., Dreisbach, G., 2017. The role of affective evaluation in conflict adaptation: an LRP study. Brain Cognit. 116, 9–16. https://doi.org/10.1016/j.bandc.2017.05.003.

Frömer, R., Maier, M., Abdel Rahman, R., 2018. Group-level EEG-processing pipeline for flexible single trial-based analyses including linear mixed models. Front. Neurosci. 12, 1–15. https://doi.org/10.3389/fnins.2018.00048.

Ganushchak, L.Y., Christoffels, I.K., Schiller, N.O., 2011a. The use of electroencephalography in language production research: a review. Front. Psychol. 2, 1–6. https://doi.org/10.3389/fpsyg.2011.00208.

Ganushchak, L.Y., Verdonschot, R.G., Schiller, N.O., 2011b. When leaf becomes neuter: event-related potential evidence for grammatical gender transfer in bilingualism. Neuroreport 22 (3), 106–110. https://doi.org/10.1097/WNR.0b013e3283427359.

Gollan, T.H., Montoya, R.I., Fennema-Notestine, C., Morris, S.K., 2005. Bilingualism affects picture naming but not picture classification. Mem. Cognit. 33 (7), 1220–1234. https://doi.org/10.3758/BF03193224.

González Alonso, J., Alemán Bañón, J., DeLuca, V., Miller, D., Pereira Soares, S.M., Puig-Mayenco, E., Slaats, S., Rothman, J., 2020. Event related potentials at initial exposure in third language acquisition: implications from an artificial mini-grammar study. J. Neurolinguistics 56, 100939. https://doi.org/10.1016/j.jneuroling.2020.100939.

Green, D.W., 1998. Mental control of the bilingual lexico-semantic system. Biling. Lang. Cognit. 1 (2), 67–81. https://doi.org/10.1017/S1366728998000133.

Grözinger, B., Kornhuber, H.H., Kriebel, J., 1975. Methodological problems in the investigation of cerebral potentials preceding speech: determining the onset and suppressing artefacts caused by speech. Neuropsychologia 13 (3), 263–270. https://doi.org/10.1016/0028-3932(75)90002-0.

Guo, T., Peng, D., 2006. Event-related potential evidence for parallel activation of two languages in bilingual speech production. Neuroreport 17 (17), 1757–1760. https://doi.org/10.1097/01.wnr.0000246327.89308.a5.

Habets, B., Jansma, B.M., Münte, T.F., 2008. Neurophysiological correlates of linearization in language production. BMC Neurosci. 9 (1), 1–8. https://doi.org/10.1186/1471-2202-9-77.

Hahne, A., Friederici, A.D., 1999. Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. J. Cognit. Neurosci. 11 (2), 194–205. https://doi.org/10.1162/089892999563328.

Hanulová, J., Davidson, D.J., Indefrey, P., 2011. Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. Lang. Cognit. Process. 26 (7), 902–934. https://doi.org/10.1080/01690965.2010.509946.

Heidlmayr, K., Ferragne, E., Isel, F., 2021. Neuroplasticity in the phonological system: the PMN and the N400 as markers for the perception of non-native phonemic contrasts by late second language learners. Neuropsychologia 156, 107831. https://doi.org/10.1016/j.neuropsychologia.2021.107831.

Heim, S., Friederici, A.D., Schiller, N.O., Rüschemeyer, S.-A., Amunts, K., 2009. The determiner congruency effect in language production investigated with functional MRI. Hum. Brain Mapp. 30 (3), 928–940. https://doi.org/10.1002/hbm.20556.

Hermans, D., Bongaerts, T., De Bot, K., Schreuder, R., 1998. Producing words in a foreign language: can speakers prevent interference from their first language? Biling. Lang. Cognit. 1 (3), 213–229. https://doi.org/10.1017/S1366728998000364.

Hoshino, N., Kroll, J.F., 2008. Cognate effects in picture naming: does cross-language activation survive a change of script? Cognition 106 (1), 501–511. https://doi.org/10.1016/j.cognition.2007.02.001.

Hoshino, N., Thierry, G., 2011. Language selection in bilingual word production: electrophysiological evidence for cross-language competition. Brain Res. 1371, 100–109. https://doi.org/10.1016/j.brainres.2010.11.053.

Hruby, T., Marsalek, P., 2003. Event-related potentials—the P3 wave. Acta Neurobiol. Exp. 63, 55–63.

Indefrey, P., 2011. The spatial and temporal signatures of word production components: a critical update. Front. Psychol. 2, 1–16. https://doi.org/10.3389/fpsyg.2011.00255.

Indefrey, P., Levelt, W.J.M., 2004. The spatial and temporal signatures of word production components. Cognition 92 (1–2), 101–144. https://doi.org/10.1016/j.cognition.2002.06.001.

Ivanova, I., Costa, A., 2008. Does bilingualism hamper lexical access in speech production? Acta Psychol. 127 (2), 277–288. https://doi.org/10.1016/j.actpsy.2007.06.003.

Izura, C., Cuetos, F., Brysbaert, M., 2014. Lextale-Esp: a test to rapidly and efficiently assess the Spanish vocabulary size. Psicologica 35 (1), 49–67.

Jiao, L., Grundy, J.G., Liu, C., Chen, B., 2020. Language context modulates executive control in bilinguals: evidence from language production. Neuropsychologia 142, 107441. https://doi.org/10.1016/j.neuropsychologia.2020.107441.

Koester, D., Schiller, N.O., 2008. Morphological priming in overt language production: electrophysiological evidence from Dutch. Neuroimage 42 (4), 1622–1630. https://doi.org/10.1016/j.neuroimage.2008.06.043.

Kroll, J.F., Bobb, S.C., Misra, M., Guo, T., 2008. Language selection in bilingual speech: evidence for inhibitory processes. Acta Psychol. 128 (3), 416–430. https://doi.org/10.1016/j.actpsy.2008.02.001.

Kroll, J.F., Bobb, S.C., Wodniecka, Z., 2006. Language selectivity is the exception, not the rule: arguments against a fixed locus of language selection in bilingual speech. Biling. Lang. Cognit. 9 (2), 119–135. https://doi.org/10.1017/S1366728906002483.

Kroll, J.F., Stewart, E., 1994. Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. J. Mem. Lang. 33, 149–174.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62 (1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123.

Laganaro, M., Morand, S., Schwitter, V., Zimmermann, C., Camen, C., Schnider, A., 2009. Electrophysiological correlates of different anomic patterns in comparison with normal word production. Cortex 45 (6), 697–707. https://doi.org/10.1016/j.cortex.2008.09.007.

Lago, S., Mosca, M., Garcia, A.S., 2021. The role of crosslinguistic influence in multilingual processing: lexicon versus syntax. Lang. Learn. 71 (S1), 163–192. https://doi.org/10.1111/lang.12412.

Lange, V.M., Perret, C., Laganaro, M., 2015. Comparison of single-word and adjective-noun phrase production using event-related brain potentials. Cortex 67, 15–29. https://doi.org/10.1016/j.cortex.2015.02.017.

Lau, E.F., Phillips, C., Poeppel, D., 2008. A cortical network for semantics: (De) constructing the N400. Nat. Rev. Neurosci. 9 (12), 920–933. https://doi.org/10.1038/nrn2532.

Leckey, M., Federmeier, K.D., 2019. Electrophysiological methods in the study of language processing. In: de Zubicaray, G.I., Schiller, N.O. (Eds.), The Oxford Handbook of Neurolinguistics. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190672027.013.3.

Lee, M.-W., Williams, J.N., 2001. Lexical access in spoken word production by bilinguals: evidence from the semantic competitor priming paradigm. Biling. Lang. Cognit. 4 (3), 233–248.

Lemhöfer, K., Broersma, M., 2012. Introducing LexTALE: a quick and valid lexical test for advanced learners of English. Behav. Res. Methods 44 (2), 325–343. https://doi.org/10.3758/s13428-011-0146-0.

Lemhöfer, K., Dijkstra, T., Michel, M., 2004. Three languages, one ECHO: cognate effects in trilingual word recognition. Lang. Cognit. Process. 19 (5), 585–611. https://doi.org/10.1080/01690960444000007.

Lemhöfer, K., Schriefers, H., Hanique, I., 2010. Native language effects in learning second-language grammatical gender: a training study. Acta Psychol. 135 (2), 150–158. https://doi.org/10.1016/j.actpsy.2010.06.001.

Lemhöfer, K., Spalek, K., Schriefers, H., 2008. Cross-language effects of grammatical gender in bilingual word recognition and production. J. Mem. Lang. 59 (3), 312–330. https://doi.org/10.1016/j.jml.2008.06.005.

Levelt, W.J.M., Roelofs, A., Meyer, A.S., 1999. A theory of lexical access in speech production. Behav. Brain Sci. 22, 1–38. https://doi.org/10.1017/S0140525X99001776.

Li, C., Gollan, T.H., 2018. Cognates interfere with language selection but enhance monitoring in connected speech. Mem. Cognit. 46 (6), 923–939. https://doi.org/10.3758/s13421-018-0812-x.

Luck, S.J., 1998. Sources of dual-task interference: evidence from human electrophysiology. Psychol. Sci. 9 (3), 223–227. https://doi.org/10.1111/1467-9280.00043.

Marian, V., Blumenfeld, H.K., Kaushanskaya, M., 2007. the Language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. J. Speech Lang. Hear. Res. 50 (4), 940–967. https://doi.org/10.1044/1092-4388(2007/067).

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H.R., Bates, D.M., 2017. Balancing Type I error and power in linear mixed models. J. Mem. Lang. 94, 305–315. https://doi.org/10.1016/j.jml.2017.01.001.

Midgley, K.J., Holcomb, P.J., Grainger, J., 2009. Language effects in second language learners and proficient bilinguals investigated with event-related potentials. J. Neurolinguistics 22 (3), 281–300. https://doi.org/10.1016/j.jneuroling.2008.08.001.

Midgley, K.J., Holcomb, P.J., Grainger, J., 2011. Effects of cognate status on word comprehension in second language learners: an ERP investigation. J. Cognit. Neurosci. 23 (7), 1634–1647. https://doi.org/10.1162/jocn.2010.21463.

Miozzo, M., Caramazza, A., 1999. The selection of determiners in noun phrase production. J. Exp. Psychol. Learn. Mem. Cognit. 25 (4), 907–922. https://doi.org/10.1037/0278-7393.25.4.907.

Misra, M., Guo, T., Bobb, J.F., Kroll, J.F., 2012. When bilinguals choose a single word to speak: electrophysiological evidence for inhibition of the native language. J. Mem. Lang. 67 (1), 224–237. https://doi.org/10.1016/j.jml.2012.05.001.

Morales, L., Paolieri, D., Dussias, P.E., Valdés Kroff, J.R., Gerfen, C., Bajo, M.T., 2016. The gender congruency effect during bilingual spoken-word recognition. Biling. Lang. Cognit. 19 (2), 294–310. https://doi.org/10.1017/S1366728915000176.

Müller, N., Hulk, A., 2001. Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. Biling. Lang. Cognit. 4 (1), 1–21. https://doi.org/10.1017/S1366728901000116.

Neath, A.A., Cavanaugh, J.E., 2012. the bayesian information criterion: background, derivation, and applications. WIREs Comput. Stat. 4 (2), 199–203. https://doi.org/10.1002/wics.199.

Neuhaus, A.H., Trempler, N.R., Hahn, E., Luborzewski, A., Karl, C., Hahn, C., Opgen-Rhein, C., Urbanek, C., Schaub, R., Dettling, M., 2010a. Evidence of specificity of a visual P3 amplitude modulation deficit in schizophrenia. Schizophr. Res. 124 (1–3), 119–126. https://doi.org/10.1016/j.schres.2010.08.014.

Neuhaus, A.H., Urbanek, C., Opgen-Rhein, C., Hahn, E., Ta, T.M.T., Koehler, S., Gross, M., Dettling, M., 2010b. Event-related potentials associated with attention network test. Int. J. Psychophysiol. 76 (2), 72–79. https://doi.org/10.1016/j.ijpsycho.2010.02.005.

Newman, A.J., Tremblay, A., Nichols, E.S., Neville, H.J., Ullman, M.T., 2012. The influence of language proficiency on lexical semantic processing in native and late

learners of English. J. Cognit. Neurosci. 24 (5), 1205–1223. https://doi.org/10.1162/jocn_a_00143.

Paolieri, D., Demestre, J., Guasch, M., Bajo, T., Ferré, P., 2020. The gender congruency effect in Catalan–Spanish bilinguals: behavioral and electrophysiological evidence. Biling. Lang. Cognit. 23 (5), 1–11. https://doi.org/10.1017/S1366728920000073.

Paolieri, D., Padilla, F., Koreneva, O., Morales, L., Macizo, P., 2019. Gender congruency effects in Russian–Spanish and Italian–Spanish bilinguals: the role of language proximity and concreteness of words. Biling. Lang. Cognit. 22 (1), 112–129. https://doi.org/10.1017/S1366728917000591.

Parker-Jones, 'Ōiwi, Green, D.W., Grogan, A., Pliatsikas, C., Filippopolitis, K., Ali, N., Lee, H.L., Ramsden, S., Gazarian, K., Prejawa, S., Seghier, M.L., Price, C.J., 2012. Where, when and why brain activation differs for bilinguals and monolinguals during picture naming and reading aloud. Cerebr. Cortex 22 (4), 892–902. https://doi.org/10.1093/cercor/bhr161.

Peeters, D., Dijkstra, T., Grainger, J., 2013. The representation and processing of identical cognates by late bilinguals: RT and ERP effects. J. Mem. Lang. 68 (4), 315–332. https://doi.org/10.1016/j.jml.2012.12.003.

Peterson, R.R., Savoy, P., 1998. Lexical selection and phonological encoding during language production: evidence for cascaded processing. J. Exp. Psychol. Learn. Mem. Cognit. 24 (3), 539–557. https://doi.org/10.1037/0278-7393.24.3.539.

Pivneva, I., Palmer, C., Titone, D., 2012. Inhibitory control and L2 proficiency modulate bilingual language production: evidence from spontaneous monologue and dialogue speech. Front. Psychol. 3 https://doi.org/10.3389/fpsyg.2012.00057.

Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. 118 (10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019.

Polich, J., 2012. Neuropsychology of P300. In: Kappenman, E.S., Luck, S.J. (Eds.), The Oxford Handbook of Event-Related Potential Components. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0089.

Porcaro, C., Medaglia, M.T., Krott, A., 2015. Removing speech artifacts from electroencephalographic recordings during overt picture naming. Neuroimage 105, 171–180. https://doi.org/10.1016/j.neuroimage.2014.10.049.

Pulvermüller, F., 2007. Word processing in the brain as revealed by neurophysiological imaging. In: Gaskell, M.G. (Ed.), The Oxford Handbook of Psycholinguistics. Oxford University Press, pp. 118–140. https://doi.org/10.1093/oxfordhb/9780198568568711.013.0008.

Pulvermüller, F., Shtyrov, Y., Hauk, O., 2009. Understanding in an instant: neurophysiological evidence for mechanistic language circuits in the brain. Brain Lang. 110 (2), 81–94. https://doi.org/10.1016/j.bandl.2008.12.001.

R Core Team, 2019. RStudio: Integrated Development for R (1.4.1106) [Computer software]. https://rstudio.com/.

Ritter, W., Vaughan, H.G., 1969. Averaged evoked responses in vigilance and discrimination: a reassessment. Science 164 (3877), 326–328. https://doi.org/10.1126/science.164.3877.326.

Rodriguez-Fornells, A., Lugt, A. van der, Rotte, M., Britti, B., Heinze, H.-J., Münte, T.F., 2005. Second language interferes with word production in fluent bilinguals: brain potential and functional imaging evidence. J. Cognit. Neurosci. 17 (3), 422–433. https://doi.org/10.1162/0898929053279559.

Rodriguez-Fornells, A., Schmitt, B.M., Kutas, M., Münte, T.F., 2002. Electrophysiological estimates of the time course of semantic and phonological encoding during listening and naming. Neuropsychologia 40 (7), 778–787. https://doi.org/10.1016/S0028-3932(01)00188-9.

Roelofs, A., Piai, V., Garrido Rodriguez, G., Chwilla, D.J., 2016. Electrophysiology of cross-language interference and facilitation in picture naming. Cortex 76, 1–16. https://doi.org/10.1016/j.cortex.2015.12.003.

Rosenman, R., Tennekoon, V., Hill, L.G., 2011. Measuring bias in self-reported data. Int. J. Behav. Healthc. Res. 2 (4), 320–332. https://doi.org/10.1504/IJBHR.2011.043414.

Rossi, E., Newman, S., Kroll, J.F., Diaz, M.T., 2018. Neural signatures of inhibitory control in bilingual spoken production. Cortex 108, 50–66. https://doi.org/10.1016/j.cortex.2018.07.009.

Rossi, S., Gugler, M.F., Friederici, A.D., Hahne, A., 2006. The impact of proficiency on syntactic second-language processing of German and Italian: evidence from event-related potentials. J. Cognit. Neurosci. 18 (12), 2030–2048. https://doi.org/10.1162/jocn.2006.18.12.2030.

Rothman, J., 2015. Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: timing of acquisition and proficiency considered*. Biling. Lang. Cognit. 18 (2), 179–190. https://doi.org/10.1017/S136672891300059X.

Runnqvist, E., Strijkers, K., Sadat, J., Costa, A., 2011. On the temporal and functional origin of L2 disadvantages in speech production: a critical review. Front. Psychol. 2, 1–8. https://doi.org/10.3389/fpsyg.2011.00379.

Rust, N.C., Mehrpour, V., 2020. Understanding image memorability. Trends Cognit. Sci. 24 (7), 557–568. https://doi.org/10.1016/j.tics.2020.04.001.

Sá-Leite, A.R., Fraga, I., Comesaña, M., 2019. Grammatical gender processing in bilinguals: an analytic review. Psychonomic Bull. Rev. 26 (4), 1148–1173. https://doi.org/10.3758/s13423-019-01596-8.

Sá-Leite, A.R., Luna, A., Fraga, I., Comesaña, M., 2020. The gender congruency effect across languages in bilinguals: a meta-analysis. Psychonomic Bull. Rev. 27, 677–693. https://doi.org/10.3758/s13423-019-01702-w.

Schendan, H.E., Kutas, M., 2003. Time course of processes and representations supporting visual object identification and memory. J. Cognit. Neurosci. 15 (1), 111–135. https://doi.org/10.1162/089892903321107864.

Schiller, N.O., 2006. Lexical stress encoding in single word production estimated by event-related brain potentials. Brain Res. 1112 (1), 201–212. https://doi.org/10.1016/j.brainres.2006.07.027.

Schiller, N.O., Bles, M., Jansma, B.M., 2003. Tracking the time course of phonological encoding in speech production: an event-related brain potential study. Cognit. Brain Res. 17 (3), 819–831. https://doi.org/10.1016/S0926-6410(03)00204-0.

Schiller, N.O., Caramazza, A., 2003. Grammatical feature selection in noun phrase production: evidence from German and Dutch. J. Mem. Lang. 48 (1), 169–194. https://doi.org/10.1016/S0749-596X(02)00508-9.

Schmitt, B.M., Schiltz, K., Zaake, W., Kutas, M., Münte, T.F., 2001. An electrophysiological analysis of the time course of conceptual and syntactic encoding during tacit picture naming. J. Cognit. Neurosci. 13 (4), 510–522. https://doi.org/10.1162/08989290152001925.

Schneider, W., Eschman, A., Zuccolotto, A., 2002. E-prime User's Guide. https://www.researchgate.net/publication/260296789_E-prime_User's_Guide.

Schriefers, H., 1992. Lexical access in the production of noun phrases. Cognition 45 (1), 33–54. https://doi.org/10.1016/0010-0277(92)90022-A.

Schriefers, H., 1993. Syntactic processes in the production of noun phrases. J. Exp. Psychol. Learn. Mem. Cognit. 19 (4), 841–850. https://doi.org/10.1037/0278-7393.19.4.841.

Schriefers, H., de Ruiter, J.P., Steigerwald, M., 1999. Parallelism in the production of noun phrases: experiments and reaction time models. J. Exp. Psychol. Learn. Mem. Cognit. 25 (3), 702–720. https://doi.org/10.1037/0278-7393.25.3.702.

Sholl, A., Sankaranarayanan, A., Kroll, J.F., 1995. Transfer between picture naming and translation: a test of asymmetries in bilingual memory. Psychol. Sci. 6 (1), 45–49. https://doi.org/10.1111/j.1467-9280.1995.tb00303.x.

Soares Pereira, S.M., Ong, G., Abutalebi, J., Del Maschio, N., Sewell, D., Weekes, B., 2019. A diffusion model approach to analyzing performance on the Flanker task: the role of the DLPFC. Biling. Lang. Cognit. 22 (5), 1194–1208. https://doi.org/10.1017/S1366728918000044.

Squires, N.K., Squires, K.C., Hillyard, S.A., 1975. Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. Electroencephalogr. Clin. Neurophysiol. 38 (4), 387–401. https://doi.org/10.1016/0013-4694(75)90263-1.

Steinhauer, K., White, E.J., Drury, J.E., 2009. Temporal dynamics of late second language acquisition: evidence from event-related brain potentials. Sec. Lang. Res. 25 (1), 13–41. https://doi.org/10.1177/0267658308098995.

Strijkers, K., Costa, A., Thierry, G., 2010. Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. Cerebr. Cortex 20 (4), 912–928. https://doi.org/10.1093/cercor/bhp153.

Sunderman, G., Kroll, J.F., 2006. First language activation during second language lexical processing: an investigation of lexical form, meaning, and grammatical class. Stud. Sec. Lang. Acquis. 28 (3), 387–422.

Thierry, G., Wu, Y.J., 2007. Brain potentials reveal unconscious translation during foreign-language comprehension. Proc. Natl. Acad. Sci. Unit. States Am. 104 (30), 12530–12535. https://doi.org/10.1073/pnas.0609927104.

Torres, L., Blevins, A.S., Bassett, D.S., Eliassi-Rad, T., 2020. The Why, How, and when of Representations for Complex Systems. ArXiv:2006.02870 [Cs, q-Bio]. http://arxiv.org/abs/2006.02870.

Valente, A., Bürki, A., Laganaro, M., 2014. ERP correlates of word production predictors in picture naming: a trial by trial multiple regression analysis from stimulus onset to response. Front. Neurosci. 8, 1–13. https://doi.org/10.3389/fnins.2014.00390.

Valente, A., Pinet, S., Alario, F.-X., Laganaro, M., 2016. "When" does picture naming take longer than word reading? Front. Psychol. https://doi.org/10.3389/fpsyg.2016.00031, 0.

Van der Meij, M., Cuetos, F., Carreiras, M., Barber, H.A., 2011. Electrophysiological correlates of language switching in second language learners. Psychophysiology 48 (1), 44–54. https://doi.org/10.1111/j.1469-8986.2010.01039.x.

Van Turennout, M., Hagoort, P., Brown, C.M., 1995. Electrophysiological evidence on the time course of semantic and phonological processes in speech production. J. Exp. Psychol. Learn. Mem. Cognit. 23 (4), 787–806.

Verdonschot, R.G., Middelburg, R., Lensink, S.E., Schiller, N.O., 2012. Morphological priming survives a language switch. Cognition 124 (3), 343–349. https://doi.org/10.1016/j.cognition.2012.05.019.

Voeten, C.C., 2019. *Permutes: Permutation tests for Time Series Data* (1.0) [Computer software]. https://CRAN.R-project.org/package=permutes.

von Grebmer zu Wolfsthurn, S., Pablos Robles, L., Schiller, N.O., 2021. Cross-linguistic interference in late language learners: an ERP study. Brain Lang. 221, 104993. https://doi.org/10.1016/j.bandl.2021.104993.

Weber-Fox, C.M., Neville, H.J., 1996. Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. J. Cognit. Neurosci. 8 (3), 231–256. https://doi.org/10.1162/jocn.1996.8.3.231.

White, E.J., Titone, D., Genesee, F., Steinhauer, K., 2017. Phonological processing in late second language learners: the effects of proficiency and task. Biling. Lang. Cognit. 20 (1), 162–183. https://doi.org/10.1017/S1366728915000620.

Wicha, N.Y.Y., Moreno, E.M., Kutas, M., 2003. Expecting gender: an event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. Cortex 39 (3), 483–508. https://doi.org/10.1016/S0010-9452(08)70260-0.

Wood Bowden, H., Steinhauer, K., Sanz, C., Ullman, M.T., 2013. Native-like brain processing of syntax can be attained by university foreign language learners. Neuropsychologia 51 (13), 2492–2511. https://doi.org/10.1016/j.neuropsychologia.2013.09.004.

Wu, Y.J., Thierry, G., 2013. Fast modulation of executive function by language context in bilinguals. J. Neurosci. 33 (33), 13533–13537. https://doi.org/10.1523/JNEUROSCI.4760-12.2013.

Xiong, K., Verdonschot, R.G., Tamaoka, K., 2020. The time course of brain activity in reading identical cognates: an ERP study of Chinese - Japanese bilinguals. J. Neurolinguistics 55, 1–14. https://doi.org/10.1016/j.jneuroling.2020.100911.

Yip, V., Matthews, S., 2007. Relative clauses in Cantonese-English bilingual children: typological challenges and processing motivations. Stud. Sec. Lang. Acquis. 29 (2), 277–300.

Zhang, Q., Damian, Markus F., 2009. The time course of segment and tone encoding in Chinese spoken production: an event-related potential study. Neuroscience 163 (1), 252–265. https://doi.org/10.1016/j.neuroscience.2009.06.015.

Zurn, P., Bassett, D.S., Rust, N.C., 2020. The Citation Diversity Statement: a practice of transparency, a way of life. Trends Cognit. Sci. 24 (9), 669–672. https://doi.org/10.1016/j.tics.2020.06.009.