Early Perceptual Effects (or lack thereof) of Conflicting Grammatical Genders: ERP Evidence from Simultaneous Bilinguals

Oleksandra Osypenko

Department of Linguistics and English Language, Lancaster University, Lancaster, UK

Silke Brandt

Department of Linguistics and English Language, Lancaster University, Lancaster, UK

Aina Casaponsa

Department of Linguistics and English Language, Lancaster University, Lancaster, UK

Corresponding Author:

Oleksandra Osypenko
Department of Linguistics and English Language,
County South, Lancaster University,
Lancaster, LA1 4YL, United Kingdom
Email: o.osypenko@lancaster.ac.uk

Early Perceptual Effects (or lack thereof) of Conflicting Grammatical Genders: ERP Evidence from Simultaneous Bilinguals

Previous studies revealed that grammatical gender may shape perception and categorisation (Sato et al., 2020), yet the extent of this influence remains unclear, specifically in simultaneous bilinguals with two gendered languages. This pre-registered ERP study investigates how two partially contrasting grammatical gender systems modulate perception in Ukrainian-Russian bilinguals. Twenty-six participants completed a non-verbal categorisation task assessing associations between primes (depicted nouns with matching or mismatching grammatical genders across L1s) and target (male/female faces). Behavioural results (response types and reaction times) showed that bilinguals were not affected by the prime-target gender congruency for matching primes or by more dominant/proficient L1 for mismatched primes. ERP analyses showed no significant modulations of predicted components (N1, P2/VPP, N300) by grammatical gender for either type of primes. This suggests that grammatical gender alone may not independently modulate categorisation processes in bilinguals with two gendered L1s, especially when explicit conceptual or semantic activation is not required.

Keywords: grammatical gender, categorisation, perception, linguistic relativity, simultaneous bilingualism, event-related potentials, language proficiency, language dominance

Introduction

Over the past few decades, linguistic relativity studies have grown exponentially, leading to more nuanced investigations of how language affects perception and cognition. Although linguistic relativity research has extensively addressed *lexical* properties and their effect on thought and perception (e.g., *colour discrimination*: Athanasopoulos et al., 2010; Drivonikou et al., 2007; Gilbert et al., 2006; Roberson et al., 2005; Thierry et al., 2009; Winawer et al., 2007; *spatio-temporal metaphors*: Athanasopoulos & Bylund, 2023; Bylund &

Athanasopoulos, 2017), Whorf (1956) himself focused on the effects that *grammatical* properties have on one's thought. One grammatical property that has received a lot of attention is grammatical gender (see Fatemi, 2024; Samuel et al., 2019 for most recent reviews). The early studies in this line of research focused on finding a yes/no answer to whether grammatical gender of one's native language affects thought, and mostly recruited monolingual speakers (Flaherty, 2001; Sera et al., 1994; 2002; Vigliocco et al., 2005). However, more recently these studies have expanded the scope of research and included bilingual speakers to investigate if these grammatical gender effects on perception remain when participants are tested in a genderless L2 (e.g., Chen & Faitaki, 2024; Pavlidou & Alvanoudi, 2013, 2019; Sato & Athanasopoulos, 2018), or whether learning a gendered L2 (while having a genderless L1) can result in a shift in perception (Athanasopoulos & Boutonnet, 2016; Kurinski & Sera, 2011), as summarised in more detail below.

Grammatical gender effects on perception: behavioural evidence across bilinguals

Research involving bilinguals with a gendered L1 and a genderless L2 predominantly shows that the influence of grammatical gender from the native language persists even when participants perform tasks in a genderless L2 (cf. Sato et al., 2013). Conversely, studies involving bilingual participants with a genderless L1 and a gendered L2 reveal that exposure to a gendered L2 begins to affect object categorisation as early as 10 weeks after learning commences (Athanasopoulos & Boutonnet, 2016; Kurinski & Sera, 2011). Additionally, Athanasopoulos and colleagues showed that individual differences (e.g., length of residence, language use, etc.) can influence bilinguals' categorical perception linked to grammatical constructs (Athanasopoulos, 2006; Athanasopoulos & Bylund, 2023). For example, looking at grammatical number, Athanasopoulos (2006) found that increased language proficiency, as measured by a standardised proficiency test, led to shifts in categorisation patterns among Japanese-English bilinguals. Specifically, participants with intermediate English proficiency displayed categorisation patterns resembling those of Japanese monolinguals, whereas participants with advanced English proficiency exhibited patterns similar to English monolinguals. Yet, in the domain of grammatical gender, we still know relatively little about how these individual-difference factors (e.g., L1/L2 proficiency, patterns of language use, age of acquisition) modulate gender-related categorisation in bilingual and multilingual speakers, especially when both L1 and L2 have contrasting grammatical gender systems (Osypenko & Athanasopoulos, 2026).

An important consideration when comparing findings across languages and experimental paradigms is how bilinguals' language profiles are assessed. Earlier studies examining the effects of grammatical gender in bilingual populations predominantly relied on self-reported proficiency (e.g., Bassetti, 2007; Boutonnet et al., 2012; Sato et al., 2020), while a number of more recent investigations (e.g., Athanasopoulos & Bylund, 2023; Chen & Faitaki, 2024; Kurinski & Sera, 2011; Osypenko et al., 2025a, 2025b) provide a more comprehensive assessment.

One way of integrating the wide range of individual differences shown to influence bilingual cognition is through the concept of *language dominance*. The construct captures the relative strength of a bilingual's languages rather than an isolated measure of language proficiency (Gertken et al., 2014; Olson, 2023; Solís-Barroso et al., 2019). Overall, language dominance reflects not only linguistic knowledge but also patterns of language use and language history, and is therefore, conceptualised as inherently gradient rather than a categorical construct. Tools such as the Bilingual Language Profile (BLP; Gertken et al., 2014) operationalise this multidimensional view by assessing language history, use, self-rated proficiency, and attitudes, thereby providing a continuous measure of relative language dominance. This continuous conceptualisation reflects the view that dominance consists of graded asymmetries across self-perceived proficiency, language use, and processing, which are relative and dynamic rather than discreet (Gertken et al., 2014; Solís-Barroso et al., 2019).

Importantly, while language proficiency and dominance often correlate, they are not interchangeable, as they index different aspects of bilingual language experience (Olson, 2023), the latter capturing more nuances of language use and exposure. Language dominance is commonly assessed through self-reported measures, whereas language proficiency is measured using standardised language tests. Combining self-reported and standardised measures therefore provides a principled and comprehensive approach, particularly in populations with complex socio-political contexts, such as Ukrainian-Russian bilinguals. Overall, by combining standardised and self-rated measures, researchers can obtain critical information on both the standardised levels of language proficiency - found to influence grammatical gender effects (Athanasopoulos, 2006; Athanasopoulos & Bylund, 2023; cf. Chen & Faitaki, 2024) - and self-reported measures, such as age of acquisition, which have also been shown to impact bilinguals' perception of gendered stimuli (Athanasopoulos et al., 2010; Boroditsky et al., 2003).

Importantly, existing behavioural research has primarily focused on *sequential bilinguals* (Boutonnet et al., 2012; Forbes et al., 2008; Sato et al., 2020; Sato & Athanasopoulos, 2018 among others), with limited evidence of grammatical gender effects in *simultaneous bilinguals* (Bassetti, 2007; Osypenko et al., 2025a, 2025b; Osypenko & Athanasopoulos, 2026). Inclusion of simultaneous bilinguals can provide further insights not only whether there are effects of grammatical gender on perception of speakers of two gendered languages (L1s) but also into how two grammatical systems interact with each other in one mind. Early behavioural research on simultaneous bilinguals (i.e., Bassetti, 2007) used stimuli exclusively with mismatching grammatical gender assignments across bilinguals' two languages, making it unclear if the lack of grammatical gender effects was due to co-activation of two opposite grammatical genders or other methodological reasons (e.g., small sample size).

To address this, Osypenko et al. (2025b) conducted two behavioural experiments examining grammatical gender effects on categorisation in Ukrainian-Russian simultaneous bilinguals, while tested in a genderless L2 (English). One of the key reasons to look at simultaneous bilinguals of Ukrainian and Russian languages in particular is their typological distinctions. Both languages are three-gendered and assign masculine, feminine, and neuter grammatical genders to nouns (Budzhak-Jones, 1997). Importantly, while some nouns have the same gender across both languages (e.g., "a pencil" is masculine in both Ukrainian ["olivets"] and Russian ["karandash"]), others differ in their grammatical gender assignment (e.g., "a notebook" is feminine ["tetrad"] in Russian and masculine ["zoshyt"] in Ukrainian). This feature allows us to test the effects of grammatical gender when they match or mismatch across two L1s.

Osypenko et al. (2025b) employed a similarity judgement paradigm, where participants are presented with object-character pairs and asked to rate how similar they are on a 9-point Likert scale. The characters paired with the objects were of either male (e.g., a prince) or female biological sex (e.g., a queen). Stimuli were chosen to be conceptually neutral, and participants' language proficiency in Ukrainian, Russian, and English was assessed using standardised tests. Experiment 1 included objects from all three grammatical gender categories (masculine, feminine, and neuter), while Experiment 2 included only masculine and feminine objects. This manipulation aimed to evaluate whether the presence of neuter grammatical gender weakened associations between grammatical gender and biological sex, potentially diminishing grammatical gender effects. Consistent with this hypothesis, Experiment 1, which included neuter nouns, revealed no significant grammatical gender effects. In contrast, Experiment 2, without neuter-gender nouns, produced significant grammatical gender effects. Specifically, bilingual participants rated object-character pairs as significantly more similar when the grammatical gender of the object and the biological sex of the character were congruent across both Ukrainian and Russian. Moreover, for objects with mismatching grammatical gender, participants' similarity judgements aligned closely with grammatical gender in their more proficient L1.

Limitations of behavioural findings

A crucial limitation of all behavioural studies, including those discussed above, is the inability to determine whether observed effects genuinely reflect perceptual changes driven by grammatical gender or whether they reflect participants' strategic use of explicit metalinguistic knowledge (Samuel et al., 2019; Sedlmeier et al., 2016). For instance, studies that use voice attribution (Bassetti, 2007) or similarity judgement (Osypenko et al., 2025b) have been characterised by Samuel et al. (2019) as having high gender/sex salience, making it possible for participants to consciously engage grammatical gender as a task-solving strategy. Additionally, such studies are particularly vulnerable to verbal interference, where participants may implicitly or explicitly rely on verbal information (see Athanasopoulos & Casaponsa, 2020). Consequently, findings from behavioural studies often reflect language activation in the moment of task completion rather than providing evidence supporting long-lasting effects of language on perceptual encoding, which stem from neural adaptations over time.

To address this issue, researchers have increasingly turned to neural measures, particularly event-related potentials (ERPs). ERPs offer high temporal resolution, enabling precise tracking of automatic and unconscious cognitive processes that are time-locked to specific stimuli. As Thierry et al. (2024) point out, most language effects on cognition occur outside of conscious awareness, but to fully assess the strength of these effects, two key research design elements are needed. First, experimental paradigms should focus on non-verbal tasks that examine perception and conceptualisation mechanisms, as these processes are expected to be influenced by language. Second, it is crucial to incorporate neural methods that can track the temporal unfolding of effects. By differentiating between effects arising at early perceptual stages and those occurring later in the processing stream - when linguistic activation may influence processing - ERPs help distinguish unconscious language effects at the pre-attentive and pre-verbal stages of processing from the overt language use (Casaponsa et al., 2024; Maier & Abdel Rahman, 2024; Thierry et al., 2009; Xue & Williams, 2024).

Grammatical gender effects on perception: ERP evidence and ERP components related to grammatical gender processing

While ERPs have been used in linguistic relativity research on domains such as colour (see review by Thierry, 2016) and motion events (Flecken et al., 2015), EEG studies examining the influence of grammatical gender on perceptual categorisation remain comparatively scarce (see Boutonnet et al., 2012; Sato et al., 2020). In the study by Boutonnet et al. (2012), Spanish-English bilinguals and English monolinguals completed a semantic categorisation task involving triplets of depicted objects in English. Participants had to determine whether the third object (target) belonged to the same semantic category as the first two objects, while ERPs were being recorded. Critically, the target's grammatical gender in Spanish was covertly manipulated, being either gender-congruent or gender-incongruent with the first two objects. Grammatical gender congruency was expected to modulate the left anterior negativity (LAN), a marker of morphosyntactic processing (Friederici et al., 1993; Friederici & Jacobsen, 1999) in the bilingual group only. The presence of LAN would indicate automatic and unconscious retrieval of grammatical gender, rather than a strategic use of such information. As predicted, in the Spanish-English bilingual group, but not the English monolingual group, LAN showed a statistically significant effect of gender congruency. It was more negative in gender-incongruent trials compared to congruent ones, providing evidence in favour of spontaneous and unconscious retrieval of grammatical gender information in bilinguals regardless of semantic category.

Similarly to Boutonnet and colleagues (2012), Sato et al. (2020) investigated the effects of grammatical gender on perception and categorisation but shifted the focus from semantic relatedness to conceptual gender. French-English bilinguals and English monolinguals were presented with object-face pairs and tested in English. Each object had both a grammatical gender in French (masculine or feminine) and a conceptual gender (stereotypically male or female). Each object was paired with either a male or female face, resulting in four conditions: (1) conceptually related and grammatically congruent (e.g., "*poupée*" [doll], grammatically feminine and conceptually female, paired with a female face); (2) conceptually related and grammatically incongruent (e.g., "*collier*" [necklace], grammatically masculine and conceptually female, paired with a female face); (3) conceptually unrelated and grammatically congruent (e.g., "*cravate*" [tie], grammatically feminine and conceptually male, paired with a female face); and (4) conceptually unrelated and grammatically incongruent (e.g., "*cigare*" [cigar], grammatically masculine and conceptually male, paired with a female face). Participants were tasked with responding "yes" or "no" to whether the object made them think of the face, while their EEG data were recorded. N1, P2/VPP, and N300 components were examined to assess whether grammatical gender influenced perceptual processes in bilingual speakers during early face processing, even though it was irrelevant to the task. The N1 component, an early negative-going ERP waveform, has two distinct subcomponents: anterior (N1a) and posterior. Sato et al. (2020) focused on the anterior N1 (N1a), which typically peaks over fronto-central electrodes around 100 milliseconds after stimulus onset. This subcomponent is associated with the initial sensory processing of visual or auditory information, as well as perceptual expectations and evaluations (Marzecová et al., 2018; Vogel & Luck, 2000). The N1 component is known to be modulated by the expectations of concurrent stimuli (e.g., Federmeier & Kutas, 2001; Lee et al., 2012). In their study, Sato et al. (2020) reported that grammatical gender congruency modulated N1 amplitude for bilingual participants but not for monolinguals. Specifically, the greater negativity was found for grammatically incongruent trials compared to congruent ones, suggesting an early, automatic attentional shift driven by grammatical gender. Interestingly, conceptual gender relatedness did not significantly influence the N1 component in either participant group, indicating that this component was solely sensitive to grammatical, rather than conceptual, gender information.

Another component examined in the study by Sato et al. (2020) is the P2 or vertex positive potential (VPP). This positive-going potential is observed around 150 milliseconds post-stimulus, particularly at anterior-central sites (Jeffreys, 1989; Joyce & Rossion, 2005). The P2/VPP is widely regarded as a key ERP marker for categorical face perception (i.e., N170 counterpart when bi-mastoids are used as reference electrodes); it is thought to reflect the structural encoding of faces and is modulated by categorical aspects such as facial expression, race, and gender (Ito & Urland, 2005; Kecskés-Kovács et al., 2013). Sato and colleagues (2020) found that bilinguals, but not monolinguals, showed greater P2/VPP positivity for grammatically congruent object-face pairs compared to incongruent pairs, suggesting facilitated early face processing when grammatical gender matched the biological sex of the face.

Lastly, Sato et al. (2020) predicted an N300 amplitude modulation in the later stages of visual processing. The N300, a late visual negative-going component, has an anterior and central-parietal region distribution, peaking approximately 300 milliseconds after the onset of a visual stimulus (Holcomb & McPherson, 1994; Kumar et al., 2021). The N300 has been linked with mechanisms of perceived object identification, particularly when semantic expectations are violated (Federmeier & Kutas, 2001) and is sensitive to the global features of visual stimuli (McPherson & Holcomb, 1999; Schendan & Kutas, 2002, 2003). This component typically shows greater negative amplitude for images that conflict with expected semantic categories compared to those that align with them. In their study, Sato et al. (2020) found N300 modulations for monolinguals with conceptually unrelated stimuli showing greater negativity than conceptually related stimuli pairs. However, bilingual speakers failed to show these modulations. Instead, the N300 component was modulated by grammatical gender congruency. Hence, grammatical gender seems to have exerted a stronger influence than conceptual gender stereotypes, suggesting that grammatical gender overrides conceptual associations.

In addition to the components examined by Sato et al. (2020), we originally pre-registered an exploratory analysis of a very early ERP modulations (approximately 45-90 ms) reported by Mouchetant-Rostaing and colleagues (Mouchetant-Rostaing et al., 2000; Mouchetant-Rostaing & Giard, 2003). Their studies showed reliable differences between male and female faces (and between face and non-face stimuli such as hands) emerging around 45-90 ms post-stimulus over occipito-parietal sites, in both explicit gender categorisation tasks and passive viewing. The found effects were interpreted as reflecting very rapid, automatic discrimination processes (Mouchetant-Rostaing et al., 2000; Mouchetant-Rostaing & Giard, 2003). These findings suggested that coarse gender-related distinctions might already be detectable at an early visual stage, motivating our pre-registered analysis of this effect. However, because these effects overlap with the C1 time range and are likely to be driven by low-level physical differences between stimulus categories, consistent with evidence that C1 is generated in primary visual cortex (V1) and is sensitive to basic visual properties rather than higher-level categories (Di Russo et al., 2002), we will now treat this analysis as exploratory and report it only in the Appendix C[1], while our main hypotheses and conclusions focus on the N1, P2/VPP, and N300 components.

Overall, previous research using ERPs suggests that grammatical gender is unconsciously activated even in tasks when it is not explicitly required. However, to our knowledge, no ERP study has explored these early categorical effects in bilingual speakers with two gendered L1s. Furthermore, a methodological question arises from previous findings. Specifically, were the effects amplified by the presence of conceptual gender (Sato et al., 2020) or semantic associations (Boutonnet et al., 2012)? If so, would grammatical gender still be accessed if the stimuli were conceptually and semantically neutral? Building on the statement by Samuel et al. (2019, p. 1773), while studies like Sato et al. (2020) and Boutonnet et al. (2012) show evidence in favour of access to grammatical information when it is not required, "this is not the same as demonstrating that objects are conceptualised as more masculine or feminine as a function of their gender assignment. Such results might be explained in terms of an effect of membership in the same grammatical category, independently of biological sex information". Therefore, designing an EEG study looking at the unconscious access to grammatical gender on its own without any additional semantic or conceptual manipulations, could provide insights to answer these questions.

[1] We thank the anonymous reviewer for this insightful suggestion.

Aims and scope of the current study

In the current pre-registered study (https://doi.org/10.17605/OSF.IO/2VR7K), we recorded ERPs in Ukrainian-Russian simultaneous bilinguals as they were completing an association judgement task involving an object and a gendered face (male or female). The task was completed fully in English (participants' genderless L2) to avoid any associations with grammatical gender of L1s. The experimental paradigm was adapted from Sato et al. (2020). However, unlike Sato et al. (2020), where participants judged a set of stereotypical gender-associated objects, our aim was to analyse whether grammatical gender effects persist in both behavioural and neurophysiological measures when using conceptually neutral items.

Two key research questions were formulated. First, we sought to determine whether the effects of matching grammatical gender across two L1s would be present in both behavioural (response types and reaction times) and ERP measures (N1, P2/VPP, N300). Specifically, we predicted faster reaction times and/or gender-biased choices in pairs where the prime's grammatical gender in the two L1s is congruent with the biological sex of the face (e.g., faster reaction times when a "pencil", masculine in both Ukrainian and Russian, is paired with a male face), compared to incongruent pairs. Additionally, we predicted to find decreased N1 and N300 amplitudes, as well as a greater P2/VPP modulation, for pairs where the grammatical gender of the object is congruent with the biological sex of the face in the two L1s, compared to incongruent pairs. Decreased N1 amplitude under congruent condition would be taken to reflect facilitated early perceptual discrimination, greater P2/VPP positivity to indicate enhanced structural encoding and categorical processing of faces, and reduced N300 negativity to reflect easier integration and categorisation of the object-face pair (Federmeier & Kutas, 2001; Ito & Urland, 2005; Jeffreys, 1989; Kecskés-Kovács et al., 2013; Kumar et al., 2021).

Second, we aimed to investigate potential L1 effects when grammatical gender of an item differs across first languages (Bassetti, 2007). Specifically, we hypothesised that participants would exhibit priming effects of the grammatical gender of their more proficient or their more dominant L1 (Ukrainian or Russian). For behavioural measures, we expected to observe faster reaction times and/or biased choices towards the pairs where the prime's grammatical gender and the target's biological sex are congruent in participants' more proficient/dominant L1. If grammatical gender in gendered languages impacted categorisation, we would expect modulation of N1, P2/VPP and N300 for the most proficient/dominant L1s.

Furthermore, for each research question we examined effects of English proficiency/dominance on participants' responses, based on earlier findings (Athanasopoulos, 2006; Athanasopoulos & Bylund, 2023; Osypenko & Athanasopoulos, 2026). We hypothesised that higher English dominance/proficiency would reduce grammatical gender effects, diluting the binary masculine-feminine gender distinctions from participants' L1s. Specifically, we predicted that participants who are more proficient in English would exhibit weaker grammatical gender effects reflected in both reaction times and response types (i.e., smaller differences in reaction times/responses between congruent and incongruent pairs). For each ERP component, amplitudes for congruent pairs were expected to be less pronounced in participants with higher English proficiency/dominance, compared to those with lower English proficiency/dominance.

In summary, the present study investigates the role of grammatical gender in shaping cognitive and perceptual processes in simultaneous Ukrainian-Russian bilinguals using both behavioural and neurophysiological measures. By examining congruent and incongruent grammatical gender conditions across two distinct grammatical systems, this study aims to disentangle the effects of grammatical gender from potential covert verbal strategies and contribute to the broader understanding of linguistic relativity in simultaneous bilingual cognition.

Methods

Participants

26 Ukrainian-Russian simultaneous bilinguals with English as a second language (L2) participated in the current experiment at Lancaster University. Proficiency levels were measured using standardised tests, such as the advanced university-entry level proficiency tests for Ukrainian and Russian proficiency (Ukrainian Center for Educational Quality Assessment, 2020) and the Cambridge English Proficiency test (Cambridge University Press, 2024). Proficiency test scores were measured on a scale from 0 to 100.

Additionally, participants completed the Bilingual Language Profile (BLP; Gertken et al., 2014) to determine their language dominance in all three languages. The BLP consists of four modules assessing language history, language use, self-rated language proficiency, and language attitudes, with responses collected separately for each language. Item responses within each module are converted into numerical values and summed to yield language-specific module scores, which are then combined into a global composite score for each language (see Table 1). In the original BLP, all four module scores are weighted to contribute equally to the global score. In the present study, however, to adapt the BLP questionnaire to the current socio-political context in Ukraine, we excluded the “language attitudes” module and based the composite scores on the history, use, and self-rated proficiency modules. The resulting composite scores were then used to calculate a continuous dominance coefficient for each language (ranging from -194 to +194).

Table 1

Structure and scoring of the Bilingual Language Profile (BLP; Gertken et al., 2014)

BLP module	Number of items	Item score range	Raw module score range	Adjustment coefficient	Contribution to global score
Language history	6	0–20	0–120	0.454	Weighted to equal module contribution
Language use	5	0–10	0–50	1.09	Weighted to equal module contribution
Self-rated proficiency	4	0–6	0–24	2.27	Weighted to equal module contribution
Language attitudes	4	0–6	0–24	2.27	Weighted to equal module contribution
Total (per language)	**19**	—	—	—	**Maximum composite score = 218**

Note. The maximum composite score (218) corresponds to the full four-module BLP. Because the 'language attitudes' module was removed in the current study, our maximum composite score was 194.

Proficiency tests and the BLP questionnaire were completed prior to the main testing. Details of the participants' language profiles are provided in Table 2. Following data collection, we calculated scores for Ukrainian and Russian separately, and then derived two coefficients to reflect relative proficiency and dominance. Specifically, the L1 Proficiency coefficient was calculated by subtracting the Russian proficiency score from the Ukrainian proficiency score, resulting in values ranging from -100 (only proficient in Russian) to +100 (only proficient in Ukrainian). Similarly, the L1 Dominance coefficient ranged from -194 (fully dominant in Russian) to +194 (fully dominant in Ukrainian), based on the difference between language dominance scores in the BLP. Importantly, the standardised proficiency tests for Ukrainian and Russian used in this study are university entry/placement exams in Ukraine (Ukrainian center for educational quality assessment, 2020) calibrated to the B2-C2 range of the CEFR, making them well suited to capture meaningful variation even among highly proficient simultaneous bilinguals. Consistent with this, the derived L1 Proficiency and L1 Dominance coefficients show considerable dispersion (see Table 4), indicating meaningful variability in relative L1 strength rather than a ceiling pattern.

All participants reported little to no knowledge of additional gendered languages (two participants reported learning German and Polish at school, but not using it at the time of testing; another participant reported using learning applications to recreationally learn Spanish, German and Italian, but chose 0% of time in the "language use" section).

As reported in the pre-registration (https://doi.org/10.17605/OSF.IO/2VR7K), we excluded data from participants who exhibited response biases, such as consistently responding "yes" or "no" or showing alternating patterns (e.g., "yes-no-yes-no") for 90% or more of their responses within a block. Data were also excluded due to poor electroencephalogram (EEG) signal quality, heavy artifact contamination (see "EEG analysis"), or fewer than 25 usable trials per condition after artifact rejection. This led to the exclusion of 7 participants (2 due to response biases and 5 due to poor EEG signal quality), resulting in a final sample of 19 Ukrainian-Russian bilinguals (18 females; $Mean_{age}$ = 39.0, *SD* = 8.70, *Range* = 22–50). All participants had normal or corrected-to-normal vision. Informed consent was obtained prior to the experiment, and participants received £10 Amazon voucher for their participation. The study protocol was approved by the Ethics Committee at Lancaster University (reference number FASSLUMS-2023-3489-AmendPaper-1).

Table 2.

Participants' language profiles

	Ukrainian			Russian			English		
	Mean	**SD**	**Range**	**Mean**	**SD**	**Range**	**Mean**	**SD**	**Range**
Language Proficiency	40.40	13.96	25–67	41.70	10.87	17–67	58.20	19.87	24–92
Language Dominance	105.45	31.08	41–148	118.05	26.49	52–164	56.90	20.02	15–87

Note. The maximum score for Language Proficiency across all languages was 100, for Language Dominance for Ukrainian and Russian – maximum was 194 (without language attitudes) and for English - 218 (with language attitudes)

The study was advertised via flyers and emails sent out to Ukrainian-Russian speakers living in Lancashire, UK. A convenience sampling strategy was employed due to the specific linguistic and demographic characteristics of the target population. Given the unique profile of this group and the logistical constraints associated with accessing this community, the sample size was limited. The predominance of female participants (n = 18) can be attributed to the timing of data collection, which occurred after the onset of the war in Ukraine. However, based on the analysis by Flaherty (2001) and Osypenko et al. (2025b), the gender imbalance is not expected to have a significant impact on the findings, as both studies found no significant differences between responses from male and female participants, suggesting that their gender did not affect the responses.

Materials

The stimuli consisted of 320 image pairs, each featuring an object alongside either a male or female face (a complete list of materials is available on the Open Science Framework: https://doi.org/10.17605/OSF.IO/2VR7K). For this study, we used the conceptually neutral stimuli from Osypenko et al. (2025b), originally normed for a behavioural study. However, given the increased trial requirements for ERP research, we conducted an additional pre-test to obtain a larger set of stimuli.

Pre-test

A pre-test following the approach of Sato and Athanasopoulos (2018) and Osypenko et al. (2025b) was carried out to select additional conceptually neutral stimuli. Fourteen Ukrainian-Russian bilinguals (10 females; $Mean_{age}$ = 24.0, SD = 4.33, $Range$ =18-33) and eleven English monolinguals (7 females; $Mean_{age}$ = 28.0, SD = 3.32, $Range$ = 23-33) were recruited online. None of these participants took part in the main experiment. Participants rated a set of 200 black-and-white images of inanimate objects and animals on a 7-point Likert scale for conceptual gender association (1 = "*Very feminine*", 7 = "*Very masculine*"). Of the 200 images, 80 depicted items with masculine grammatical gender in both Ukrainian and Russian, 80 depicted items with feminine grammatical gender in both languages, and 44 depicted items with differing grammatical gender between Ukrainian and Russian. Specifically, 28 items were feminine in Ukrainian but masculine in Russian, and 16 items were masculine in Ukrainian but feminine in Russian. Based on these ratings, 34 additional conceptually neutral images were selected (*Mean* = 4.48, *SD* = 0.98), which, combined with the 46 images from our previous study (Osypenko et al., 2025b), resulted in a total of 80 conceptually neutral stimuli.

These nouns were divided into two groups (Table 3). The first group included 40 items with matching grammatical gender across both languages, such as "strawberry", which is feminine in both Ukrainian and Russian, and "telephone", which is masculine in both languages. The second group consisted of 40 items with mismatching grammatical gender across Ukrainian and Russian, such as "moon" (masculine in Ukrainian and feminine in Russian) or "sock" (feminine in Ukrainian and masculine in Russian). Full details about stimulus presentation and pairings are provided in the next section.

Main testing

Eighty depicted black-and-white objects selected during the pre-test were each paired with an image of a female face and with a male face resulting in 160 pairs. To prevent repetition effects, an additional set of images of the selected eighty objects was included and paired with a different male and a different female face. This resulted in 320 unique pairs in total. The images depicting eighty nouns were only presented once per block (four separate experimental blocks, 80 critical trials per block). Congruency between grammatical gender and biological sex of the face was fully counterbalanced across four experimental conditions (see Table 3): (1) 80 grammatically congruent trials in both Ukrainian and Russian (CR-CU), (2) 80 grammatically incongruent in both languages (IR-IU), (3) 80 grammatically congruent in Ukrainian and incongruent in Russian (IR-CU), and (4) 80 grammatically congruent in Russian and incongruent in Ukrainian (CR-IU). To simplify condition labels throughout the text, we use the following abbreviations: C = congruent, I = incongruent, R = Russian, and U = Ukrainian. For example, CR–IU refers to trials where the object's grammatical gender is congruent with the face's sex in Russian but incongruent in Ukrainian.

Additionally, 8 conceptually gendered nouns were selected as fillers: 4 objects rated as conceptually masculine (*Mean* = 4.93, *SD* = 0.18) and 4 conceptually feminine (*Mean* = 3.07, *SD* = 0.67). These fillers resulted in a total of 32 conceptually gendered trials, which were equally distributed across the blocks (8 filler trials per block). Each object appeared only once in each block, with block order counterbalanced across participants. All trials were randomised individually for each participant. This procedure was repeated twice using different male and female faces, and a different picture of each object. In addition, 8 pairs with conceptually gendered objects (e.g., "lipstick", "briefcase") were used for practice trials. The purpose of the practice trials was to encourage participants to focus on stereotypical associations rather than on grammatical gender.

Table 3.

Examples of Stimuli based on Grammatical Gender Congruency in Ukrainian and Russian

Stimuli type		Grammatical gender in L1s	English translation	Ukrainian translation	Russian translation	Conceptual gender
1. Target stimuli	1.1. Grammatical gender matching across L1s	a) masculine in both L1s	garlic	часник	чеснок	neutral
		b) feminine in both L1s	cello	віолончель	віолончель	neutral
	1.2. Grammatical gender mismatching across L1s	a) masculine in Ukrainian, feminine in Russian	tent	намет	палатка	neutral
		b) feminine in Ukrainian, masculine in Russian	sock	шкарпетка	носок	neutral
2. Filler stimuli	2.1. Grammatical gender matching across L1s	a) masculine in both L1s	convertible	кабріолет	кабриолет	male
		b) feminine in both L1s	cooking pot	каструля	кастрюля	female
	2.2. Grammatical gender mismatching across L1s	a) masculine in Ukrainian, feminine in Russian	hat	капелюх	шляпа	female
		b) feminine in Ukrainian, masculine in Russian	axe	сокира	топор	male

Procedure

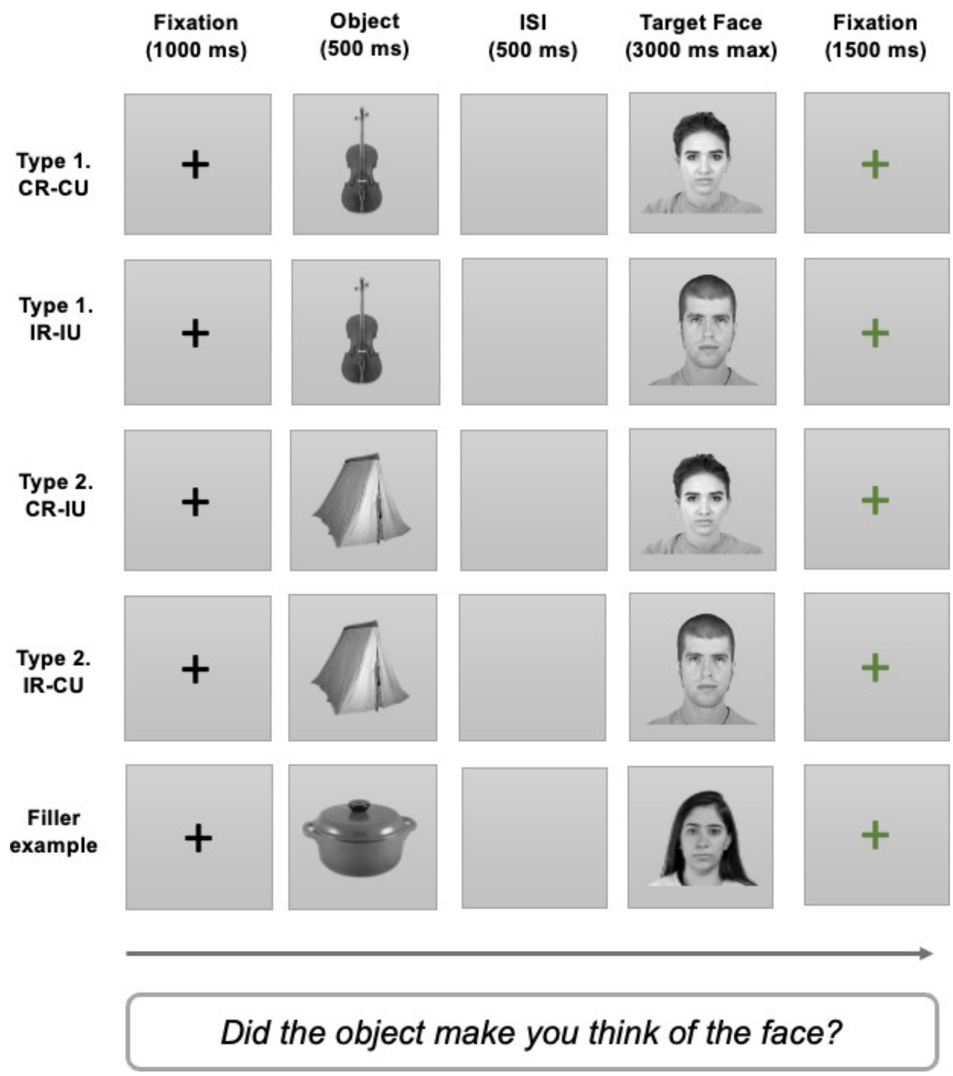
Prior to testing, we assessed participants' language proficiency and language dominance in Ukrainian, Russian, and English. On the testing day, all experimental sessions were conducted in English to minimise any cues suggesting the experiment's connection to participants' native languages. Participants were tested individually in a dimly lit booth, seated approximately 80 cm from a 17' CRT monitor. Stimuli were presented using E-Prime software (Version 2.0) foveally in a greyscale against a grey background. Responses to the target face were recorded via a button box while EEG activity was monitored.

Following the procedure of Sato et al. (2020), each trial began with a 1000 ms pre-stimulus fixation point presented at the centre of the screen, followed by the object image displayed for 500 ms (Figure 1). Then, a 500 ms interstimulus interval (ISI; blank screen) preceded the presentation of the target face image. Participants were instructed to respond whether the object made them think of the face, by pressing a “yes” or “no” button located on the outer edges of the button box. The target face remained on the screen until the participants’ responses were registered or 3000 ms had elapsed. While participants were not explicitly instructed to respond quickly, a feedback display prompted them to react faster if no response was registered within the allotted time. Each trial finished with a green prompt (+) for 1500 ms where participants were encouraged to blink if needed to minimise eye movements. A self-timed break was provided every 22 trials, and eight practice items were presented prior to the main experiment task.

Following the main experimental task, a structured debriefing session was conducted to assess whether participants recognised that grammatical gender was relevant to the study and whether such awareness led them to consciously use grammatical gender information to guide their task responses (i.e., metalinguistic strategies). The use of such strategies has been previously noted as a methodological concern in grammatical-gender judgement tasks (Chen & Faitaki, 2024; Osypenko et al., 2025b). Therefore, to control for this possibility, participants were asked open-ended questions regarding what they believed the experiment was testing and how they approached the association judgements. Participants who reported relying on grammatical gender to guide their responses would have been excluded from the analysis. However, no participants indicated such reliance, and no exclusions were necessary.

Figure 1

Sequential stimulus presentation in each trial



Note. Type 1 refers to stimuli with matching grammatical gender across Ukrainian and Russian, while Type 2 refers to stimuli with mismatching grammatical gender across both languages. The four experimental conditions included: (1) CR-CU – congruent in both Ukrainian and Russian, (2) IR-IU – incongruent in both languages, (3) CR-IU – congruent in Russian but incongruent in Ukrainian, and (4) IR-CU – incongruent in Russian but congruent in Ukrainian. A filler condition was also included. We selected filler items based on pre-test ratings, choosing only those rated as highly masculine or highly feminine.

Electroencephalography recording

EEG recordings were obtained using Neuroscan Curry7 software (Compumedics Neuroscan, NuAmps amplifier Compumedics, Charlotte, NC, USA) attached to a 38-channel elastic cap (Electro-Cap International, n.d.) positioned according to the standard 10-20 system. The electrode configuration included the following sites: FP1, FP2, Fz, F3, F4, F7, F8, FCz, FC1, FC2, FC5, FC6, Cz, C3, C4, T7, T8, CPz, CP1, CP2, CP5, CP6, Pz, P3, P4, P7, P8, Oz, O1, O2, POz, PO1, PO2, two additional electrodes to capture the horizontal eye movements, two off-line bi-mastoid reference electrodes, and the ground electrode. Two additional electrodes were used above and below the right eye to capture blinks and vertical eye movements. EEG recordings were amplified and digitised with NuAmps amplifier (Compumedics Neuroscan NuAmps amplifier Compumedics, Charlotte, NC, USA) at a sampling rate of 1 kHz. Electrodes were referenced to the average of left and right mastoids offline, and data down sampled to 250 Hz. The impedance of the mastoid and scalp electrodes was maintained below 5 kΩ and the eye electrodes below 10 kΩ throughout the recordings.

Data analysis

The behavioural and EEG analyses were divided into two parts based on our research questions. In the first part, we analysed stimuli with matching grammatical gender across Ukrainian and Russian (CR-CU and IR-IU conditions). In the second part, we examined stimuli with mismatching grammatical gender across two languages (CR-IU and IR-CU conditions).

Behavioural analysis

Across all conditions, responses below 200 ms and above 3500 ms as well as time outs were excluded from the analysis (3.56%). In addition, responses below and above 2.5 SD from the mean for each intra-participant (2.65%) and intra-item (2.49%) condition (total = 3.80%) were excluded. Importantly, in the pre-registration, we aimed to examine reaction times and accuracy of participants' responses. However, during the analysis stage, we found that participants had a response bias, and they were more likely to press "yes" in all conditions. This bias is analytically relevant, as accuracy-based measures can conflate response frequency with condition effects, potentially giving rise to spurious differences driven by the overall response tendencies rather than sensitivity to the experimental manipulation. Therefore, we adapted the analysis to model participants' response tendencies directly by analysing the proportion of yes/no responses across conditions. To do so, variable "Response Type" was dummy coded with possible values of 1 (response "yes") or 0 (response "no") across all conditions. Reaction times were analysed as reported in our pre-registration. Data analysis with accuracy can be found in Supplementary materials (Appendix B).

For both parts of the analysis, Reaction times and Response types were analysed using linear and logistic mixed effect models in R (R Core Team, 2022) with the lme4 package (Linck & Cunnings, 2015). Maximal models were fitted for all the analyses. When these failed to converge, the random-effects structure was simplified by removing those random effects that contributed the least variance. All categorical predictors were dummy coded (-0.5 and 0.5) in all models.

In the first part, we examined the effect of condition (CR-CU and IR-IU) on reaction times and response types. In the second part of the analysis (conditions CR-IU and IR-CU), to quantify participants' more proficient/dominant L1 (Ukrainian or Russian), we used L1 Proficiency and Dominance coefficients (see "Participants" section). The coefficients exhibited a wide range of variability (Figure 2), with no ceiling effects observed (Table 4). Because of a high correlation between the two variables, we analysed the coefficients in separate models.

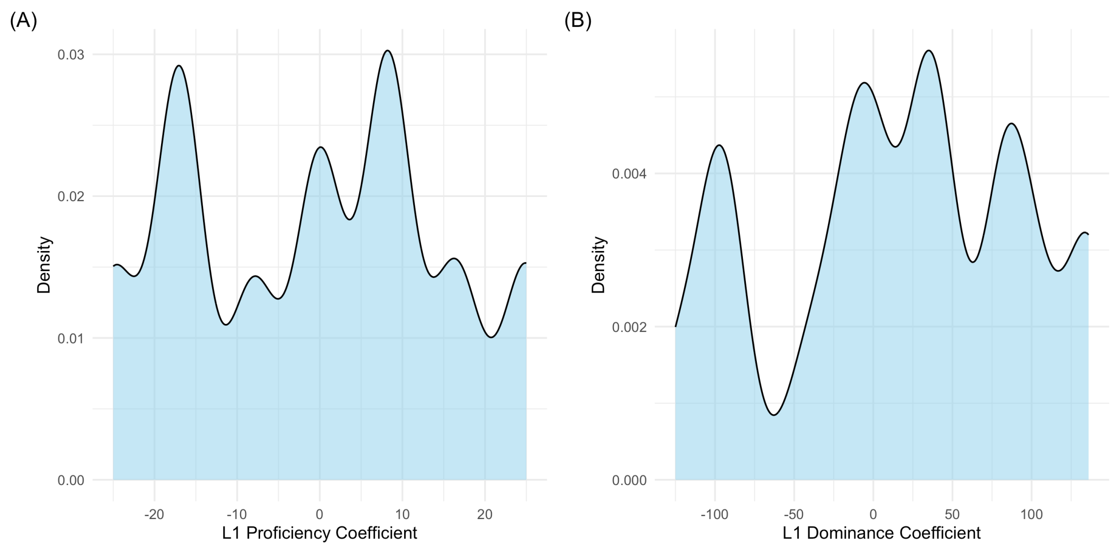
Table 4

L1 Proficiency and L1 Dominance Coefficients Across Participants

	Mean	SD	Range
L1 Proficiency Coefficient	-0.6	15.6	-25 to 25
L1 Dominance Coefficient	16.84	77.97	-125 to 136

Figure 2

Density plots for (a) L1 Proficiency and (b) L1 Dominance coefficients



Note. Each plot illustrates the distribution of participants' L1 coefficient scores, reflecting a wide range of variability without ceiling effects.

Additionally, in both analyses, we explored whether English proficiency and English dominance affected the grammatical gender influence of Ukrainian and Russian. We assessed the correlation between the two variables to determine whether they could be included in the same model. A Pearson correlation analysis revealed a low correlation between English proficiency and dominance scores (*r* = 0.082). Consequently, both variables were included in the same model. Additionally, to facilitate model convergence, both variables were scaled using the base "scale" function in R. The maximal model included random intercepts and slopes for grammatical gender congruency across both participants and items. For CR-CU and IR-IU conditions, a generalised linear mixed-effects model (glmer) was then built to analyse Response Type as a function of grammatical gender congruency, scaled English proficiency, and scaled English dominance. For CR-IU and IR-CU conditions, to investigate English proficiency effects, we analysed a three-way interaction between English proficiency, L1 Proficiency coefficient, and Condition. Similarly, for English dominance, we analysed a three-way interaction between English dominance, L1 Dominance coefficient, and Condition. To understand the direction and nature of the effects, we fitted separate glmer models using each condition as a baseline level.

EEG analysis

For the analysis of EEG data, we followed the approach of Sato et al. (2020), focusing on the N1, P2/VPP, and N300 components. The data were analysed using EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014). All EEG data pre-processing steps were scripted and run in MATLAB (v. R2022b). Obtained EEG data were down-sampled offline at 250 Hz and the average of two mastoids was used for re-referencing. Ocular artefacts were corrected using independent component analysis (ICA; Makeig et al., 1996), specifically with the "runica" function, excluding ocular electrodes from the analysis (see Casaponsa et al., 2024; Sato et al., 2020). Components associated with horizontal and vertical eye blinks were identified and manually removed using ICLabel (Pion-Tonachini et al., 2019). A high-pass filter of 0.5 Hz was applied before ICA, and a low-pass filter of 30 Hz was applied after ICA. While the pre-registration specified a high-pass filter of 0.01 Hz, the adjustment to 0.5 Hz was made based on recommendations from Delorme (2023) to optimise data quality. We computed mean ERPs time-locked to target onset off-line from trials free of ocular artefacts; epochs with activity exceeding ±75µV at any cap electrode site were automatically discarded. Baseline correction was applied using the averaged EEG activity in the 100 ms preceding the onset of the stimuli.

Similar to the behavioural analysis, the EEG analysis was divided into two parts. However, parsimonious EEG models only included random structures for participants. In line with our pre-registration protocol, we examined EEG activity in regions where previous studies reported relevant effects. Specifically, for the N1, P2/VPP, and N300 components (based on findings reported by Sato et al., 2020), we focused on the anterior region (Fz, F3, F4, FC1, FCz, FC2 electrodes). Temporal windows for each component were selected based on the latency peak. For the N1 and P2/VPP, a 60 ms time window (30 ms before and after the peak) was selected for amplitude analyses. For the N300, a 100 ms time window (50 ms before and after the peak) was selected for amplitude analyses. For each component and region of interest, mean ERP amplitudes were computed for each participant, condition, and electrode over the relevant time window. These condition-averaged mean amplitudes served as the dependent variables in the statistical analyses.

For all selected components, we fitted linear mixed-effects models (lmer) models. In the first part, we examined the effect of condition (CR-CU and IR-IU, contrast coded as 0.5 and -0.5) on the modulations of selected components. In the second part of the analysis, we also investigated whether the interaction between condition (CR-IU and IR-CU, contrast coded as -0.5 and 0.5) and proficiency/dominance coefficient influenced the modulations of selected components. For all models, random intercepts and random slopes for condition were included per participant. Additionally, an exploratory analysis was conducted to examine the interaction between condition and participants' English proficiency/dominance on the modulation of the components of interest. These models also included random intercepts and random slopes for condition per participant.

Results

Analysis for Research Question 1 (stimuli with matching grammatical gender; conditions CR-CU and IR-IU)

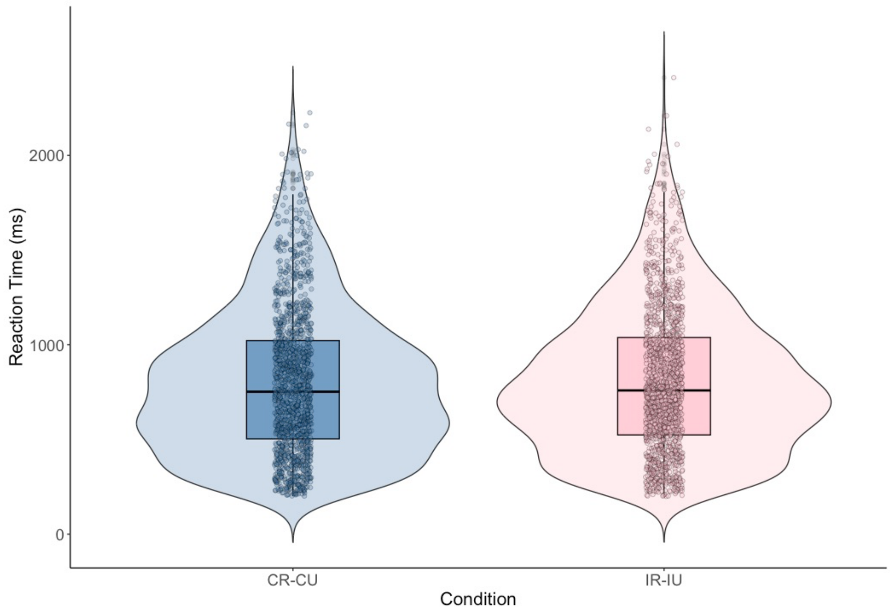
For CR-CU and IR-IU conditions, we predicted faster reaction times (RTs) and/or biased choices that align with congruent grammatical gender across participants' two L1s. For ERP measures, we hypothesised that unconscious access to grammatical gender will manifest in decreased N1 and N300 amplitudes, as well as a greater P2/VPP modulation, for pairs in the anterior region from the CR-CU condition, compared to IR-IU.

Behavioural analysis

Reaction times. The mean reaction time for the CR-CU condition was 797 ms (*SD* = 267.0), while for the IR-IU condition, it was 810 ms (*SD* = 246.7) (see Figure 3). The linear mixed effects model was built to determine whether the congruency between primes' grammatical gender and targets' biological sex (condition CR-CU) yielded faster reaction times compared to the incongruent condition (IR-IU). The parsimonious model included random intercepts for subjects and items, as well as by-subject random slopes for Condition. The model did not reveal a significant effect of grammatical gender congruency on reaction times (*Estimate* = -12.90, *SE* = 16.12, *t* = -0.80, *p* = .434, 95% CI[-46.90, 21.10]), suggesting that participants' response times were comparable across the two conditions. These findings did not support our hypothesis that grammatical gender congruency facilitates faster responses.

Figure 3

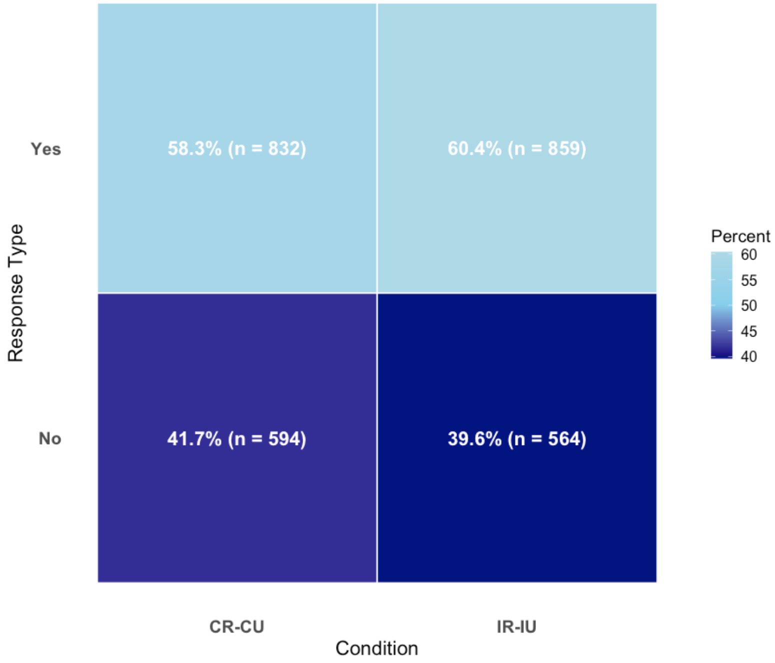
Distribution of reaction times by Condition (CR-CU vs IR-IU) and Response Type (Yes vs No)



Response Types. Similarly to the reaction times, analysis of the response types suggested no sensitivity to grammatical gender congruency. The parsimonious generalised linear mixed effects (glmer) model included fixed effects of condition, random intercepts for subjects and face images, as well as by-subject and by-item random slopes for condition. The model revealed no significant effect of Condition (*Estimate* = -0.128, *SE* = 0.257, *z* = -0.496, *p* = .620, 95% CI[-0.631, 0.376]), indicating no significant difference in the CR-CU condition compared to the IR-IU condition. Furthermore, as shown in Figure 4, participants displayed a bias in their responses and tended to press “yes” more frequently than “no” when judging associations between objects and faces regardless of the congruency of the object’s grammatical gender and the face’s biological sex.

Figure 4

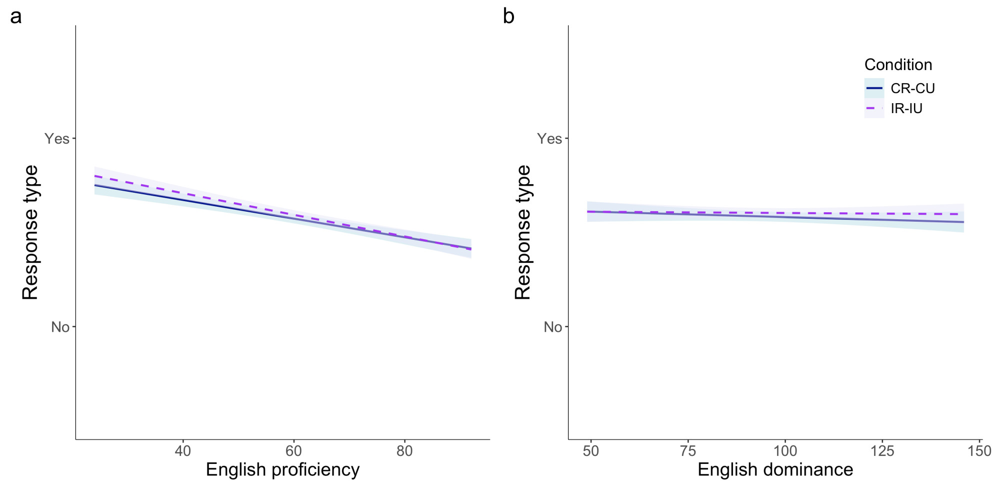
Proportion and Count of Yes and No responses by Condition (CR-CU vs IR-IU)



English proficiency and dominance. To examine the effects of English proficiency and dominance, we built a glmer model that included the interaction between Condition, English Proficiency, and English Dominance, as well as random intercepts and by-condition slopes for both subjects and items. Contrary to our predictions, we found no significant interactions. Neither the interaction between Condition and English Proficiency (*Estimate* = -0.040, *SE* = 0.257, *z* = -0.157, *p* = .875), nor the interaction between Condition and English Dominance (*Estimate* = -0.025, *SE* = 0.230, *z* = -0.110, *p* = .912) was significant (see Figure 5). Therefore, English proficiency and dominance did not reliably modulate the effect of grammatical gender for these conditions. Full model output is available on OSF: https://doi.org/10.17605/OSF.IO/2VR7K.

Figure 5

Effects of (a) English Proficiency and (b) English Dominance on Response Types (Yes vs No) across Conditions (CR-CU vs IR-IU)



ERP analysis

For each ERP component, we built a linear mixed-effects model with Condition as a fixed effect and random intercepts and by-condition slopes for each participant. The model structure was consistent across all components.

N1 time window (60-120 ms). No significant effect of Condition was found for N1 amplitudes (*Estimate* = 0.303, *SE* = 0.209, *df* = 18, *t* = 1.452, *p* = .164, 95% CI[-0.136, 0.742]). This indicates that, contrary to our predictions, N1 amplitude was not significantly greater[2] for congruent pairs compared to incongruent ones (see Figure 6).

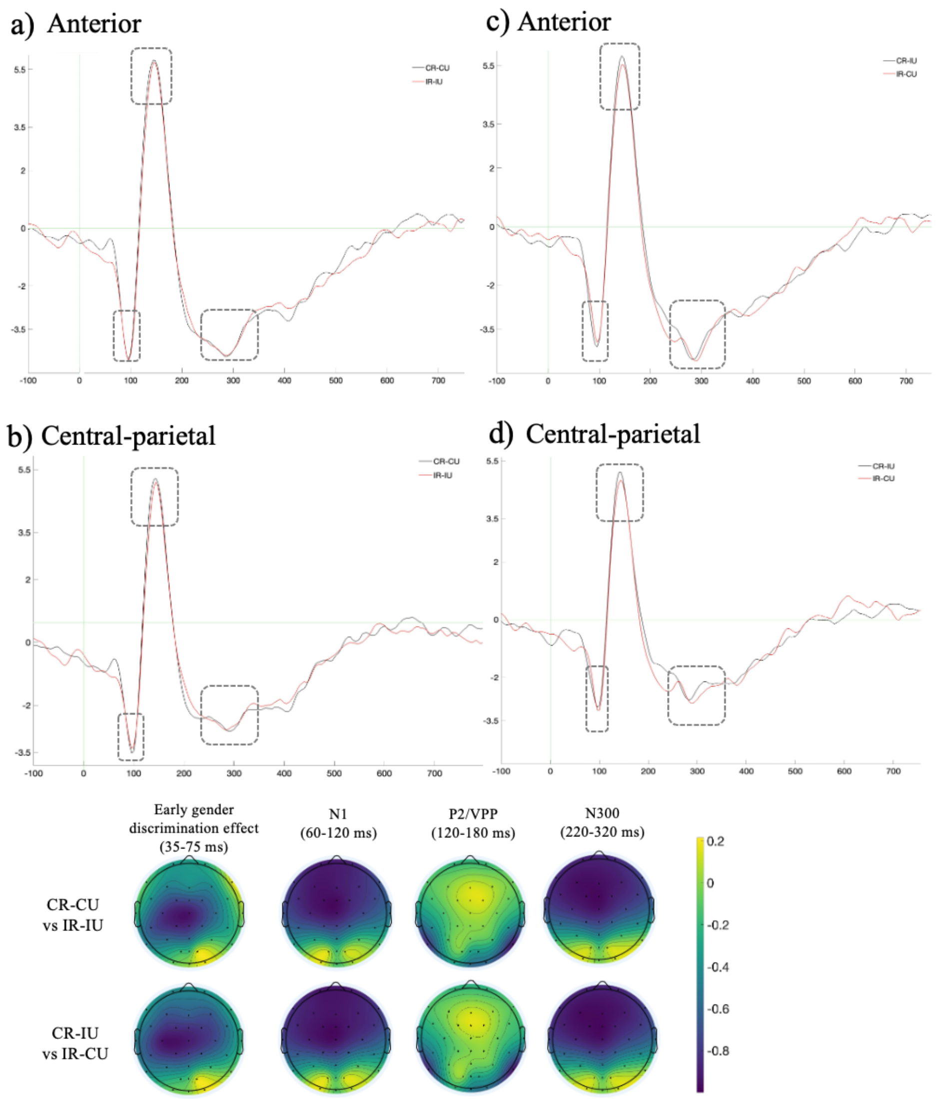
Similarly, including English Proficiency or Dominance as factors revealed no significant effects in the anterior region. The interaction between English Proficiency and Condition was non-significant (*Estimate* = -0.003, *SE* = 0.010, *df* = 17, *t* = -0.249, *p* = .807),

[2] Results were replicated using a repeated-measures ANOVA, as suggested by an anonymous reviewer. The analysis revealed no significant effect of Condition on N1 amplitudes, $F(1, 18) = 2.11$, $p = .164$, $\eta^2_p = .10$ (CR-CU: $M = -1.90$ µV; IR-IU: $M = -2.21$ µV).

as was the interaction between English Dominance and Condition (*Estimate* = -0.003, *SE* = 0.008, *df* = 17, *t* = -0.380, *p* = .709). These results suggest that increasing English proficiency and/or dominance had no measurable effect on N1 amplitude in this task.

Figure 6

Grand average ERP waveforms of the N1, P2/VPP, and N300 components (circled with a grey line) showing the effects of grammatical gender congruency for Ukrainian-Russian bilinguals across conditions



Note. A and B represent conditions CR-CU and IR-IU, indicated by black and red lines, respectively. C and D represent conditions CR-IU and IR-CU, also indicated by black and red lines, respectively. The analysis includes two regions of interest: the anterior region (A and C), encompassing electrodes Fz, F3, F4, FC1, FCz, and FC2; and the central-parietal region

(B and D), encompassing electrodes C3, Cz, C4, CP1, CP2, and CPz. Average ERPs were calculated from pooled electrodes, with time zero marking the onset of the target face image. Scalp topographies represent the effects of the Grammatical gender congruency as differences obtained between the CR-CU and IR-IU condition (E) and CR-IU and IR-CU conditions (F)

P2/VPP time window (120-180 ms). No significant effect of Condition was observed for P2/VPP amplitudes (*Estimate* = 0.295, *SE* = 0.257, *df* = 18, *t* = 1.147, *p* = .266, 95% CI[-0.245, 0.835]). We had predicted that when the biological sex of the face and grammatical gender of the object were congruent, P2/VPP amplitude would be significantly greater compared to incongruent pairs. However, our findings did not confirm this prediction[3].

Including interactions with English Proficiency/Dominance and Condition did not reveal any significant effects. Neither the interaction between Condition and English Proficiency (*Estimate* = -0.013, *SE* = 0.013, *df* = 17, *t* = -1.066, *p* = .301) nor the Condition - English Dominance interaction (*Estimate* = -0.004, *SE* = 0.010, *df* = 17, *t* = -0.417, *p* = .682) reached significance. These findings suggest that higher English proficiency or dominance did not reduce the P2/VPP amplitude for congruent pairs, contrary to our predictions.

N300 time window (220-320 ms). No significant main effect of Condition was found for the N300 amplitude in the anterior region (*Estimate* = -0.026, *SE* = 0.130, *df* = 208, *t* = -0.201, *p* = .841, 95% CI[- 0.283, 0.23]). These results indicate that N300 amplitudes did not differ between congruent and incongruent pairs, contrary to our predictions[4]. When English Proficiency was added as a factor, the interaction between Condition and English Proficiency

---

[3] Furthermore, a repeated-measures ANOVA revealed no significant effect of Condition on P2/VPP amplitudes, $F(1, 18) = 2.11$, $p = .164$, $\eta^2_p = .10$ (CR-CU: *M* = 3.83 µV; IR-IU: *M* = 3.54 µV).

[4] A repeated-measures ANOVA also showed no effect of Condition on N300 amplitudes, $F(1, 18) = 0.01$, $p = .918$, $\eta^2_p = .001$ (CR-CU: *M* = -3.89 µV; IR-IU: *M* = -3.86 µV).

was also non-significant (*Estimate* = 0.011, *SE* = 0.006, *df* = 17, *t* = -1.738, *p* = .084). Finally, adding English Dominance revealed no significant effects for Condition (*Estimate* = 0.219, *SE* = 0.484, *df* = 17, *t* = 0.452, *p* = .652) or a Condition-Dominance interaction (*Estimate* = -0.002, *SE* = 0.005, *df* = 17, *t* = -0.413, *p* = .680).

Results for Research Question 2 (stimuli with mismatching grammatical gender; conditions CR-IU and IR-CU)

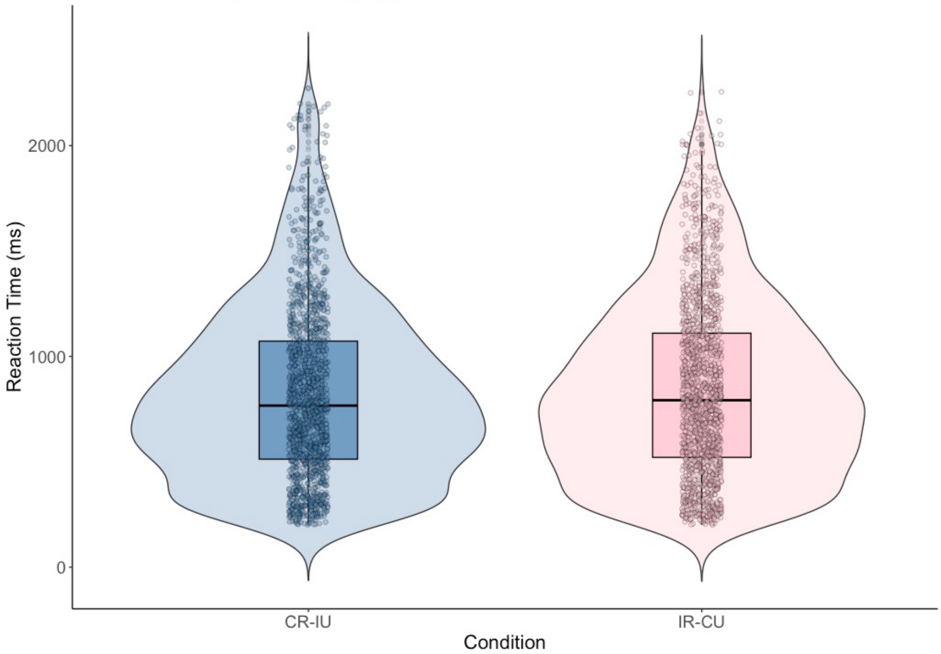
For the CR-IU and IR-CU conditions, we predicted that participants would show priming effects for the grammatical gender of their more proficient/dominant L1 (as determined by proficiency tests and the language dominance questionnaire). For behavioural measures, we predicted faster reaction times and/or biased choices towards the congruent grammatical/conceptual gender associations in participants' more proficient/dominant L1. For ERP measures, we predicted decreased N1 and N300 amplitudes and greater P2/VPP modulation for pairs where the biological sex of the face aligns with the grammatical gender of the object in the participant's more proficient/dominant L1.

Behavioural analysis

Reaction times. The mean reaction time for the CR-IU condition was 824.8 ms (*SD* = 266.4), while for the IR-CU condition, it was 846.4 ms (*SD* = 262.9) (Figure 7). Parsimonious linear mixed-effects models were fitted and included the interaction between Condition and L1 Proficiency/Dominance, as well as random intercepts and by-condition slopes for both subjects and items. Consistent with previous analyses, we found no significant interaction between the L1 Proficiency coefficient and Condition (*Estimate* = -0.598, *SE* = 0.845, *t* = -0.707, *p* = .489) or between the L1 Dominance coefficient and Condition (*Estimate* = -5.91, *SE* = 13.29, *t* = -0.45, *p* = .662) on reaction times. This suggests that neither greater proficiency nor dominance in a particular L1 influenced participants' response times when presented with pairs in which the objects' grammatical gender was congruent with the biological sex of the face in that L1 (see full model specifications on OSF: https://doi.org/10.17605/OSF.IO/2VR7K).

Figure 7

Distribution of reaction times by Condition (CR-IU vs IR-CU) and Response Type (Yes vs No)



Response Types. For the selected conditions, participants also displayed a bias towards "yes" responses (see Figure 8), especially in the IR-CU condition. To examine the role of L1 Dominance and Proficiency on Response types, we built glmer models including the interaction between Condition and L1 Dominance or L1 Proficiency, and random intercepts and by-condition slopes for both subjects and items. As in Reaction times analyses, there was no significant interaction between L1 Dominance coefficient and Condition (*Estimate* = -0.037, *SE* = 0.101, *z* = -0.363, *p* = .716) or between L1 Proficiency and Condition (*Estimate* =

-0.004, *SE* = 0.006, *z* = -0.571, *p* = .568). However, we observed a significant main effect of Condition with participants' bias to respond "yes" more often in both conditions (*Estimate* = 0.441, *SE* = 0.141, *z* = 3.137, *p* = .002). In addition, there was a significant main effect of L1 Dominance (*Estimate* = 0.435, *SE* = 0.201, *z* = 2.170, *p* = .030, see Figure 9). It suggests that participants with higher L1 Dominance scores were more likely to give "yes" responses, regardless of condition.

Figure 8

Proportion and Count of Yes and No responses by Condition (CR-IU vs IR-CU)

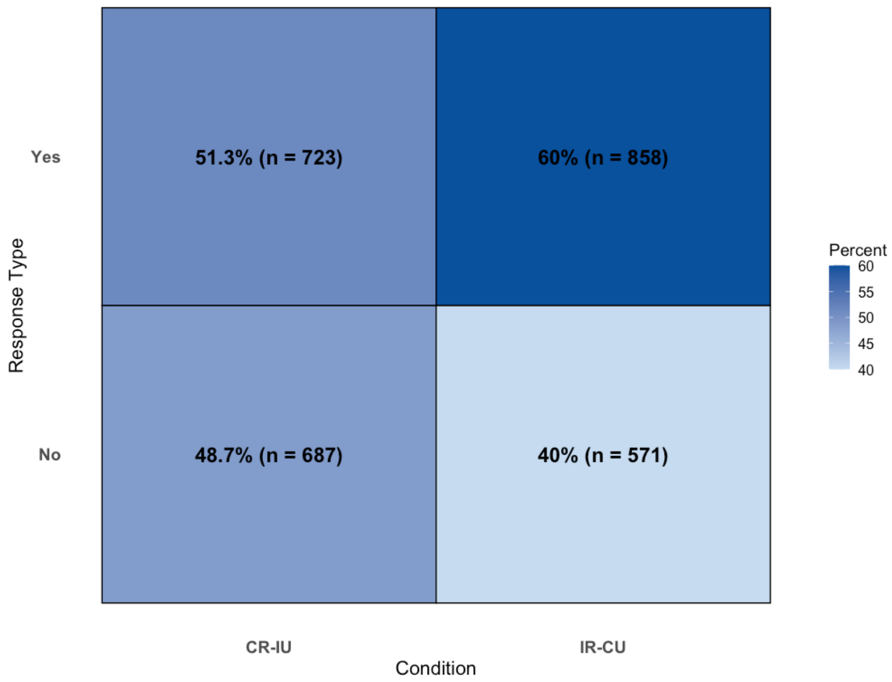
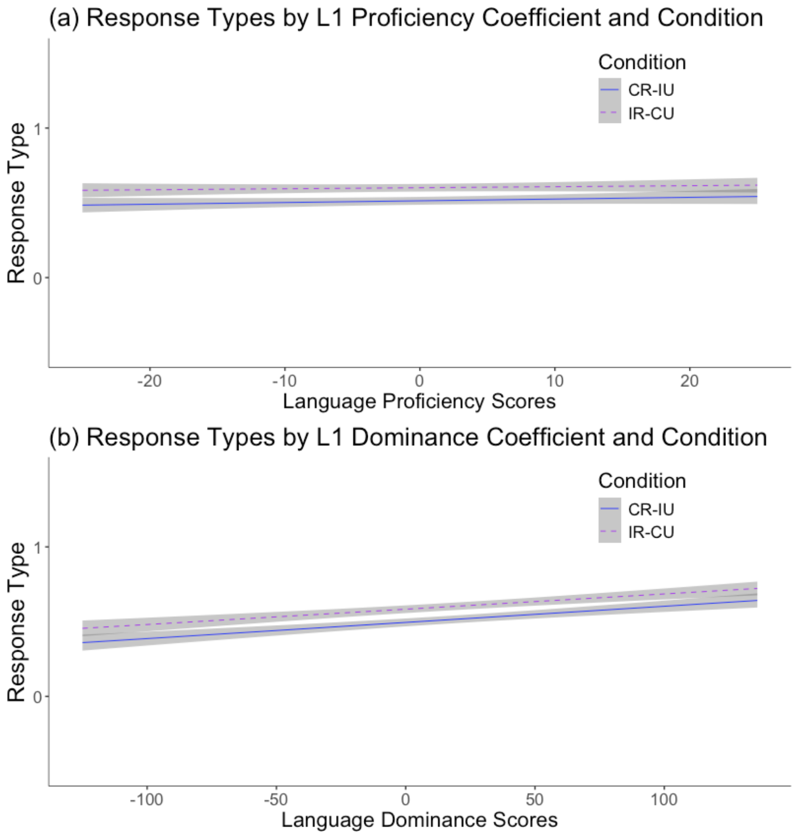


Figure 9

Response types (Yes vs No) as a function of (a) L1 Proficiency coefficient and (b) L1 Dominance coefficient across CR-IU and IR-CU conditions.



English proficiency and dominance. To examine how English proficiency and dominance influenced responses across conditions, we built a series of glmer models. These models included interactions between Condition, English Proficiency/Dominance, and the corresponding L1 measures (Proficiency or Dominance), along with random intercepts and by-condition slopes for both subjects and items.

When analysing the effects of English language proficiency on these stimuli, we did not find a significant interaction between English Proficiency, Condition, and the L1 Proficiency coefficient (IR-CU used as the reference level: *Estimate* = -0.049, *SE* = 0.078, *z* =

-0.623, *p* = .533). Similarly, for English dominance, the interaction between English Dominance, Condition and L1 Dominance coefficient was also not significant (IR-CU used as the reference level: *Estimate* = -0.061, *SE* = 0.108, *z* = -0.566, *p* = .571). Full model specifications and results are available on OSF: https://doi.org/10.17605/OSF.IO/2VR7K.

ERP analysis

For each ERP component, we fitted linear mixed-effects (lmer) models including the interaction between Condition and either the L1 Proficiency or the L1 Dominance coefficient, with random intercepts and by-condition slopes for each ERP set.

N1 time window (60-120 ms). No significant interactions were observed between Condition and the L1 Proficiency coefficient (*Estimate* < 0.001, *SE* = 0.02, *df* = 17, *t* = -0.004, *p* = .997), contradicting our predictions, showing no reduction in N1 amplitude for pairs where the biological sex of the face and the grammatical gender of the object were congruent in participants' more proficient L1.

The results for the L1 Dominance coefficient mirrored those for L1 Proficiency. No significant interaction was found in the anterior region (*Estimate* = 0.003, *SE* = 0.01, *df* = 17, *t* = 0.56, *p* = .581). This further indicates that participants' more dominant L1 had no measurable influence on N1 amplitude, regardless of congruency. For example, participants dominant in Ukrainian did not exhibit reduced N1 amplitudes for pairs congruent in Ukrainian but incongruent in Russian.

P2/VPP time window (120-180 ms). No significant interactions were observed between Condition and the L1 Proficiency coefficient (*Estimate* = 0.02, *SE* = 0.02, *df* = 17, *t* = 0.71, *p* = .487) or Condition and the L1 Dominance coefficient (*Estimate* = 0.01, *SE* = 0.01, *df* = 17, *t* = 1.16, *p* = .261). A main effect of L1 Proficiency coefficient was approaching significance

(*Estimate* = 0.09, *SE* = 0.05, *df* = 17, *t* = 1.92, *p* = .072). Yet, contrary to our predictions, these findings indicate no significant increase in P2/VPP amplitude for pairs congruent in participants' more proficient or dominant L1 compared to those congruent in their less proficient or dominant L1.

N300 time window (220-320 ms). For the anterior N300, no significant interactions were observed between Condition and the L1 Proficiency coefficient (*Estimate* = 0.031, *SE* = 0.022, *df* = 17, *t* = 1.389, *p* = .183) or Condition and the L1 Dominance coefficient (*Estimate* = 0.008, *SE* = 0.008, *df* = 17*, t* = 0.953, *p* = .354). These findings failed to support our prediction that N300 amplitude would decrease for pairs congruent in participants' more proficient or dominant L1 compared to their less proficient or dominant L1.

Discussion

In the current study, we aimed to investigate the role of two grammatical gender systems in modulating categorisation and perception in simultaneous Ukrainian-Russian bilinguals. To achieve this, we analysed behavioural (response types and reaction times) and neural (ERPs) data elicited when participants were presented with gendered (female and male) faces paired with two types of primes (objects): those with matching grammatical gender across participants' two L1s (conditions CR-CU and IR-IU) and those with mismatching grammatical gender (conditions IR-CU and CR-IU).

We first examined the overall effects of overlapping grammatical gender systems on categorical perception by looking at behavioural and ERP outcomes in trials where the grammatical gender of the prime and the biological sex of the target were either congruent or incongruent in both L1s. Behaviourally, we found no significant effects of congruency on reaction times and response types, consistent with previous studies on sequential bilinguals with one grammatical gender system (e.g., French-English: Sato et al., 2020; Spanish-English: Boutonnet et al., 2012). In terms of ERP results, contrary to our predictions, no significant effects of grammatical gender congruency were observed in the N1, P2/VPP, or N300 components. The N1 component, typically associated with early sensory discrimination, has been shown to reflect participants' ability to differentiate stimuli at a perceptual level (Marzecová et al., 2018). The absence of significant N1 modulations suggests that congruency between prime's grammatical gender and target's biological sex did not influence early perceptual discrimination in this task. The P2/VPP component, linked to attention allocation and stimulus evaluation (Sato et al., 2020; Yu et al., 2017), similarly showed no significant effects, indicating that prime-target congruency did not lead to enhanced attention or evaluation for congruent pairs. Finally, the N300, often associated with semantic integration and categorisation (Federmeier & Kutas, 2001), showed no significant modulations. This lack of differentiation implies that grammatical gender did not affect higher-order cognitive processing during the categorisation process. The significant overlap in ERP waveforms across congruent and incongruent conditions supports the conclusion that grammatical gender congruency alone (i.e., conceptually neutral stimuli) was not sufficient to modulate neural processing in our task.

We then examined whether participants would exhibit a grammatical gender bias consistent with their more proficient or dominant L1, by looking at primes with mismatching grammatical gender across Ukrainian and Russian. Behaviourally, reaction times and response types revealed a similar pattern of results, independently of participant's relative L1 proficiency or dominance. Similarly, ERP results for the N1, P2/VPP, and N300 components showed no significant modulations based on L1 dominance or proficiency coefficients. The absence of significant effects in these components suggests that the presented pairs that were congruent in grammatical gender with the more proficient/dominant L1 and incongruent in the less proficient/dominant L1 did not influence early sensory discrimination (N1), attention allocation or stimulus evaluation (P2/VPP), or higher-order semantic integration and categorisation processes (N300). Regardless of participants' more proficient or dominant L1, there was a consistent decrease in amplitude for pairs congruent in Ukrainian but incongruent in Russian. One explanation for this result relates to participants' language use on the day of testing. During debriefing, all participants included in the final analysis reported speaking Ukrainian exclusively prior to arriving at the lab, which may have biased their perception and led to pairs congruent in Ukrainian being perceived as more similar overall.

The consistent null effects observed across the ERP and behavioural data contrast sharply with findings from our previous behavioural studies (Osypenko et al., 2025b, 2025a; see also Sato et al., 2020). One possible explanation for these discrepancies lies in differences between experimental paradigms and task demands. In our earlier behavioural experiments (Osypenko et al., 2025b, 2025a), which were conducted online, there may have been unmonitored linguistic input (e.g., participants could have had oral or written communication with someone in either L1 when completing the task) that influenced results and that we could not account for. Additionally, while participants in Osypenko et al. (2025b) were instructed to respond based on their first impressions, there was no strict time limit, which may have allowed for more deliberate processing. In contrast, the current study's design emphasised rapid categorisation, potentially reducing the opportunity for conscious grammatical gender effects to emerge. Furthermore, those paradigms likely engaged grammatical gender processing more explicitly than the current study, due to the high gender saliency in the task (see Samuel et al., 2019 for a detailed discussion). However, as this study constitutes the first replication of Sato et al.'s (2020) early ERP effects, our null findings suggest that these modulations should be interpreted cautiously and would benefit from further replication across different populations and task contexts.

In this context, it is also important to acknowledge a potential limitation related to the nature of the association judgment task used in the present study, as the question posed to participants (i.e., "Does the object make you think of the face?") may not have been fully self-explanatory in the absence of explicit semantic relations[5]. However, this task was selected for two main reasons. First, it allowed grammatical gender to remain fully task-irrelevant, thereby minimising explicit gender engagement and reducing the likelihood of strategic or metalinguistic processing, which is known to affect tasks with high gender or sex salience, such as similarity judgements, voice attribution, or explicit gender classification (Chen & Faitaki, 2024; Samuel et al., 2019; Sedlmeier et al., 2016). Second, given the scarcity of ERP studies on grammatical gender and the aim to directly assess whether previously reported effects extend to simultaneous bilinguals with two partially conflicting gendered L1s, methodological comparability was prioritised. Importantly, participants were not left without guidance: prior to the main task, they completed practice trials using conceptually gendered items (e.g., a lipstick paired with a female face). The latter was done to establish the intended associative framework and ensure that participants understood how to approach the task. Following this rationale, stereotypically gendered trials were also added as filler trials during the main task.

Nevertheless, a key consideration is how our results differ from those reported by Boutonnet et al. (2012) and Sato et al. (2020), that found significant effects of grammatical gender on perception and categorisation in bilingual participants, even when grammatical gender activation was not explicitly required by the task. Unlike the present research, both studies introduced overt manipulations: Sato et al. (2020) used conceptual gender, while Boutonnet et al. (2012) employed semantic associations. Rather than directly testing whether

[5] We thank an anonymous reviewer for this suggestion.

grammatical gender alone modulates categorisation, these studies provided evidence for the unconscious activation of grammatical gender when participants engaged with semantic or conceptual associations. By contrast, the present study examined whether grammatical gender effects emerge independently when participants are exposed to conceptually neutral stimuli. Our findings, which revealed no significant effects, raise the question of whether grammatical gender effects are task-dependent, requiring co-activation with other processes, such as conceptual representations or semantic associations, to emerge.

Besides, language typology may also explain the discrepancy between the findings. Both French and Spanish, studied by Sato et al. (2020) and Boutonnet et al. (2020) respectively, are two-gender systems (with masculine and feminine genders) that maintain a more transparent relationship between grammatical and natural gender. By contrast, Ukrainian and Russian are three-gender systems (masculine, feminine, and neuter). Although our study excluded neuter stimuli, the mere presence of neuter gender in the language may have diluted the effects of grammatical gender. This aligns with the longstanding debate by Sera et al. (2002) and Vigliocco et al. (2005), who argued that grammatical gender effects are more evident in speakers of two-gendered languages than in speakers of three-gendered languages due to the more transparent masculine-feminine distinction in two-gendered languages. While comparing our results with those of Boutonnet et al. (2012) and Sato et al. (2020) provides some confirmation of this point, further research is needed to strengthen this claim. Specifically, it would be beneficial to recruit speakers of two-gendered languages to complete the task employed in the present study, without additional manipulations, and compare their findings with those of speakers of three-gendered languages in this paper. This approach would help assess the robustness of grammatical gender effects across typologically different language systems.

In addition, French and Spanish use articles as prominent gender markers, allowing faster access to grammatical gender information. Ukrainian and Russian, on the other hand, lack articles, and gender information is encoded at the semantic and morphological levels, such as in adjectives, pronouns, and determiners. Finally, a critical typological difference is the grammatical gender of the word “face.” In French, “*le visage*” is masculine, which may have amplified incongruency effects in Sato et al.’s study (2020). In Ukrainian and Russian, however, the word “face” (“*обличчя*” [oblychchya] and “*лицо*” [litso], respectively) is neuter, which may have further weakened the salience of gender congruency effects.

Another factor that requires consideration is the role of cognates in Ukrainian and Russian, which were also present in our stimuli list. A post-hoc analysis (see Supplementary Materials) revealed that cognates were more prevalent in matching-gender primes, with mismatching-gender primes showing a much lower proportion of cognates. However, ERP analyses showed no significant differences in the processing of the two types of primes, irrespective of the cognate ratio. This suggests that presence of cognates was unlikely to be the reason for the null effects observed in our experiment.

We also investigated whether participants’ L2 proficiency or dominance in English, a genderless language, influenced the findings. Both behavioural and ERP analyses revealed limited effects. While a marginally significant interaction emerged for N300 amplitudes, with higher English proficiency reducing the difference between conditions, no other significant effects were observed, contrary to previous findings (Athanasopoulos, 2006; Athanasopoulos & Bylund, 2023; Osypenko & Athanasopoulos, 2026). It is possible that English proficiency does influence participants’ response types, as predicted; however, it is challenging to determine whether it diminishes the effects of Ukrainian and Russian grammatical genders, given that these effects were not themselves significant. To further explore the role of L2 proficiency in studies using ERP, additional studies are required. For example, longitudinal studies could examine whether increasing L2 proficiency and length of exposure to English modulate these effects over time, as observed in earlier research (Athanasopoulos et al., 2010). Additionally, alternative paradigms that more robustly elicit the effects of grammatical gender on modulation of ERP components in participants' L1s could provide greater insight into how L2 proficiency contributes to cognitive restructuring.

Furthermore, characteristics of our participant sample may have influenced the findings. Our participants were older and less familiar with experimental testing compared to typical student samples, which could have impacted their performance. However, the complete overlap of ERP waveforms across conditions suggests that increasing the sample size would likely yield similar results. This consistency suggests that the null effects are more likely due to the absence of grammatical gender effects on categorical perception in this task or may be attributable to other factors discussed earlier in this section.

Our null findings, especially in contrast to previous studies that reported significant grammatical gender effects, prompt us to reflect on a point raised by Athanasopoulos and Casaponsa (2020). They argue that rather than framing the question of language effects on cognitive processes as a binary "yes/no", researchers should focus on investigating when and how these effects emerge - or, as in the case of our findings, disappear. Following this approach, we are prompted to consider: Could having two three-gendered grammatical systems, which are also partially contrasting, reduce the salience of grammatical gender effects? By contrast, do two-gendered languages such as Spanish or French provide clearer connections to biological sex and a more transparent grammatical system (i.e., presence of articles), thereby amplifying these effects? Alternatively, could the absence of effects in our findings be due to the experimental paradigm used here, which did not involve any conceptual or semantic processing that might unconsciously activate grammatical gender effects?

To address the questions arising from this study, future research could include a monolingual Russian control group to determine whether the presence of two three-gender systems negates grammatical gender effects or whether the three-gender typology itself contributes to the observed pattern. Additionally, examining other simultaneous bilingual groups with greater typological differences between their languages could provide additional insights into the interplay between grammatical systems and cognitive processing. Finally, employing tasks that require Ukrainian-Russian bilinguals to engage in conceptual or semantic processing, rather than isolated categorisation, may provide a deeper understanding of the influence of grammatical gender on perceptual processes and our findings presented in the current study.

To sum up, our findings suggest that grammatical gender does not exert a significant influence on categorisation and perception when assessed in isolation. The null effects observed in this study may be attributable to a combination of factors, including the experimental paradigm, the structural properties of Ukrainian and Russian, and the linguistic profiles of the participants. These findings highlight the need for further research to elucidate the complex interplay between grammatical gender, task demands, and cognitive processing in bilinguals.

Acknowledgements

We sincerely thank Funds for Women Graduates (FfWG) for their support. FfWG provide grants, bursaries and fellowships to women graduates to help with living expenses.

Declaration of interest statement

The authors report there are no competing interests to declare.

References

Athanasopoulos, P. (2006). Effects of the grammatical representation of number on cognition in bilinguals. *Bilingualism: Language and Cognition*, *9*(1), 89–96. https://doi.org/10.1017/S1366728905002397

Athanasopoulos, P., & Boutonnet, B. (2016). Learning grammatical gender in a second language changes categorization of inanimate objects: Replications and new evidence from English learners of L2 French. In R. Alonso Alonso (Ed.), *Crosslinguistic influence in second language acquisition* (pp. 173–192). Multilingual Matters. https://doi.org/10.21832/9781783094837-011

Athanasopoulos, P., & Bylund, E. (2023). Cognitive restructuring: Psychophysical measurement of time perception in bilinguals. *Bilingualism: Language and Cognition*, *26*(4), 809–818. https://doi.org/10.1017/S1366728922000876

Athanasopoulos, P., Dering, B., Wiggett, A., Kuipers, J.-R., & Thierry, G. (2010). Perceptual shift in bilingualism: Brain potentials reveal plasticity in pre-attentive colour perception. *Cognition*, *116*(3), 437–443. https://doi.org/10.1016/j.cognition.2010.05.016

Bassetti, B. (2007). Bilingualism and thought: Grammatical gender and concepts of objects in Italian-German bilingual children. *The International Journal of Bilingualism : Cross-Disciplinary, Cross-*

Linguistic Studies of Language Behavior, *11*(3), 251–273. https://doi.org/10.1177/13670069070110030101

Boroditsky, L., Scmidt, L. A., & Phillips, W. (2003). Sex, syntax and semantics. In *Language in mind: Advances in the study of language and thought* (pp. 61–79). Boston Review. https://doi.org/10.7551/mitpress/4117.001.0001

Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Research*, *1479*, 72–79. https://doi.org/10.1016/j.brainres.2012.08.044

Budzhak-Jones, S. (1997). Quantitative analysis of gender assignment in mono/bilingual discourse. *Journal of Quantitative Linguistics*, *4*(1–3), 67–91. https://doi.org/10.1080/09296179708590080

Bylund, E., & Athanasopoulos, P. (2017). The Whorfian time warp: Representing duration through the language hourglass. *Journal of Experimental Psychology: General*, *146*(7), 911–916. https://doi.org/10.1037/xge0000314

Cambridge University Press. (2024). *Test your English | Cambridge English*. https://www.cambridgeenglish.org/test-your-english/

Casaponsa, A., García-Guerrero, M. A., Martínez, A., Ojeda, N., Thierry, G., & Athanasopoulos, P. (2024). Electrophysiological Evidence for a Whorfian Double Dissociation of Categorical Perception Across Two Languages. *Language Learning*, *74*(S1), 136–156. https://doi.org/10.1111/lang.12648

Chen, Z., & Faitaki, F. (2024). Effects of the French grammatical gender system on bilingual adults' perception of objects. *Bilingualism: Language and Cognition*, 1–14. https://doi.org/10.1017/S1366728924000464

Delorme, A. (2023). EEG is better left alone. *Scientific Reports*, *13*(1), 2372. https://doi.org/10.1038/s41598-023-27528-0

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Di Russo, F., Martínez, A., Sereno, M. I., Pitzalis, S., & Hillyard, S. A. (2002). Cortical sources of the early components of the visual evoked potential. *Human Brain Mapping*, *15*(2), 95–111. https://doi.org/10.1002/hbm.10010

Drivonikou, G. V., Kay, P., Regier, T., Ivry, R. B., Gilbert, A. L., Franklin, A., & Davies, I. R. L. (2007). Further evidence that Whorfian effects are stronger in the right visual field than the left. *Proceedings of the National Academy of Sciences*, *104*(3), 1097–1102. https://doi.org/10.1073/pnas.0610132104

Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning,* *Memory, and Cognition*, *27*(1), 202–224. https://doi.org/10.1037/0278-7393.27.1.202

Flaherty, M. (2001). How a Language Gender System Creeps into Perception. *Journal of Cross-Cultural Psychology*, *32*(1), 18–31. https://doi.org/10.1177/0022022101032001005

Forbes, J., Poulin-Dubois, D., Rivero, M. R., & Sera, M. D. (2008). Grammatical Gender Affects Bilinguals' Conceptual Gender: Implications for Linguistic Relativity and Decision Making. *The Open Applied Linguistics Journal*, *1*, 68–76. https://doi.org/10.2174/1874913500801010068

Friederici, A. D., & Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *Journal of Psycholinguistic Research*, *28*(5), 467–484. https://doi.org/10.1023/A:1023264209610

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*(3), 183–192. https://doi.org/10.1016/0926-6410(93)90026-2

Gertken, L. M., Amengual, M., & Birdsong, D. (2014). Assessing language dominance with the Bilingual language profile. In *Measuring L2 Proficiency: Perspectives from SLA* (pp. 208–225).

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the*

National Academy of Sciences, *103*(2), 489–494. https://doi.org/10.1073/pnas.0509868103

Holcomb, P. J., & McPherson, W. B. (1994). Event-related brain potentials reflect semantic priming in an object decision task. *Brain and Cognition*, *24*(2), 259–276. https://doi.org/10.1006/brcg.1994.1014

Ito, T. A., & Urland, G. R. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(1), 21–36. https://doi.org/10.3758/CABN.5.1.21

Jeffreys, D. A. (1989). A face-responsive potential recorded from the human scalp. *Experimental Brain Research*, *78*(1), 193–202. https://doi.org/10.1007/BF00230699

Joyce, C., & Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: The effect of reference electrode site. *Clinical Neurophysiology*, *116*(11), 2613–2631. https://doi.org/10.1016/j.clinph.2005.07.005

Kecskés-Kovács, K., Sulykos, I., & Czigler, I. (2013). Is it a face of a woman or a man? Visual mismatch negativity is sensitive to gender category. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00532

Kumar, M., Federmeier, K. D., & Beck, D. M. (2021). The N300: An Index for Predictive Coding of Complex Visual Objects and Scenes. *Cerebral*

Cortex Communications, *2*(2), tgab030. https://doi.org/10.1093/texcom/tgab030

Kurinski, E., & Sera, M. (2011). Does learning Spanish grammatical gender change English-speaking adults' categorization of inanimate objects? *Bilingualism: Language and Cognition*, *14*, 203–220. https://doi.org/10.1017/S1366728910000179

Lee, C.-Y., Liu, Y.-N., & Tsai, J.-L. (2012). The Time Course of Contextual Effects on Visual Word Recognition. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00285

Linck, J. A., & Cunnings, I. (2015). The Utility and Application of Mixed-Effects Models in Second Language Research. *Language Learning*, *65*(S1), 185–207. https://doi.org/10.1111/lang.12117

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213. https://doi.org/10.3389/fnhum.2014.00213

Maier, M., & Abdel Rahman, R. (2024). Transient and long-term linguistic influences on visual perception: Shifting brain dynamics with memory consolidation. *Language Learning*, *74*(Suppl 1), 157–184. https://doi.org/10.1111/lang.12631

Makeig, S., Bell, A., Jung, T.-P., & Sejnowski, T. (1996). *Independent Component Analysis of Electroencephalographic Data*. *8*.

Marzecová, A., Schettino, A., Widmann, A., SanMiguel, I., Kotz, S. A., & Schröger, E. (2018). Attentional gain is modulated by probabilistic feature expectations in a spatial cueing task: ERP evidence. *Scientific Reports*, *8*(1), 54. https://doi.org/10.1038/s41598-017-18347-1

McPherson, W. B., & Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology*, *36*(1), 53–65. https://doi.org/10.1017/S0048577299971196

Mouchetant-Rostaing, Y., & Giard, M. H. (2003). Electrophysiological correlates of age and gender perception on human faces. *Journal of Cognitive Neuroscience*, *15*(6), 900–910. https://doi.org/10.1162/089892903322370816

Mouchetant-Rostaing, Y., Giard, M. H., Bentin, S., Aguera, P. E., & Pernier, J. (2000). Neurophysiological correlates of face gender processing in humans. *The European Journal of Neuroscience*, *12*(1), 303–310. https://doi.org/10.1046/j.1460-9568.2000.00888.x

Olson, D. J. (2023). Measuring bilingual language dominance: An examination of the reliability of the Bilingual Language Profile. *Language Testing*, *40*(3), 521–547. https://doi.org/10.1177/02655322221139162

Osypenko, O., & Athanasopoulos, P. (2026). Asymmetric cognitive restructuring of grammatical gender effects: How genderless English dilutes effects of gendered L1s in categorisation and memory. *Journal of*

Multilingual and Multicultural Development, *0*(0), 1–24. https://doi.org/10.1080/01434632.2026.2621116

Osypenko, O., Brandt, S., & Athanasopoulos, P. (2025a). Between Two Grammatical Gender Systems: Exploring the Impact of Grammatical Gender on Memory Recall in Ukrainian−Russian Simultaneous Bilinguals. *Cognitive Science*, *49*(10), e70117. https://doi.org/10.1111/cogs.70117

Osypenko, O., Brandt, S., & Athanasopoulos, P. (2025b). The influence of three-gendered grammatical systems on simultaneous bilingual cognition: The case of Ukrainian-Russian bilinguals. *Language and Cognition*, *17*, e25. https://doi.org/10.1017/langcog.2024.73

Pavlidou, T.-S., & Alvanoudi, A. (2013). Grammatical Gender and Cognition. *Versita/de Gruyter, Major Trends in Theoretical and Applied Linguistics*, *2*(2), 109–124.

Pavlidou, T.-S., & Alvanoudi, A. (2019). Conceptualizing the world as 'female' or 'male': Further remarks on grammatical gender and speakers' cognition. *Selected Papers on Theoretical and Applied Linguistics*, *23*(0), Article 0. https://doi.org/10.26262/istal.v23i0.7351

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

R Core Team. (2022). *R: The R Project for Statistical Computing*. The R Project for Statistical Computing. https://www.r-project.org/

Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology*, *50*(4), 378–411. https://doi.org/10.1016/j.cogpsych.2004.10.001

Samuel, S., Cole, G., & Eacott, M. J. (2019). Grammatical gender and linguistic relativity: A systematic review. *Psychonomic Bulletin & Review*, *26*(6), 1767–1786. https://doi.org/10.3758/s13423-019-01652-3

Sato, S., & Athanasopoulos, P. (2018). Grammatical gender affects gender perception: Evidence for the structural-feedback hypothesis. *Cognition*, *176*, 220–231. https://doi.org/10.1016/j.cognition.2018.03.014

Sato, S., Casaponsa, A., & Athanasopoulos, P. (2020). Flexing Gender Perception: Brain Potentials Reveal the Cognitive Permeability of Grammatical Information. *Cognitive Science*, *44*(9), e12884. https://doi.org/10.1111/cogs.12884

Schendan, H. E., & Kutas, M. (2002). Neurophysiological evidence for two processing times for visual object identification. *Neuropsychologia*, *40*(7), 931–945. https://doi.org/10.1016/S0028-3932(01)00176-2

Schendan, H. E., & Kutas, M. (2003). Time Course of Processes and Representations Supporting Visual Object Identification and Memory.

Journal of Cognitive Neuroscience, *15*(1), 111–135. https://doi.org/10.1162/089892903321107864

Sedlmeier, P., Tipandjan, A., & Jänchen, A. (2016). How persistent are grammatical gender effects? The case of German and Tamil. *Journal of Psycholinguistic Research*, *45*(2), 317–336. https://doi.org/10.1007/s10936-015-9350-x

Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodríguez, W., & Dubois, D. P. (2002). When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General*, *131*(3), 377–397. https://doi.org/10.1037/0096-3445.131.3.377

Solís-Barroso, C., Stefanich, S., Solís-Barroso, C., & Stefanich, S. (2019). Measuring Language Dominance in Early Spanish/English Bilinguals. *Languages*, *4*(3). https://doi.org/10.3390/languages4030062

Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J.-R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences*, *106*(11), 4567–4570. https://doi.org/10.1073/pnas.0811155106

Thierry, G., Rahman, R. A., & Athanasopoulos, P. (2024). An Introduction to the Cognitive Neuroscience of Language Embodiment and Relativity Special Issue of the Language Learning Cognitive Neuroscience Series. *Language Learning*, *74*(S1), 5–19. https://doi.org/10.1111/lang.12643

Ukrainian center for educational quality assessment. (2020). *Ukrainian center for educational quality assessment*. https://testportal.gov.ua/en/

Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical Gender Effects on Cognition: Implications for Language Learning and Language Use. *Journal of Experimental Psychology: General*, *134*(4), 501–520. https://doi.org/10.1037/0096-3445.134.4.501

Vogel, E. K., & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology*, *37*(2), 190–203.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, *104*(19), 7780–7785. https://doi.org/10.1073/pnas.0701644104

Xue, Y., & Williams, J. (2024). Inducing Shifts in Attentional and Preattentive Visual Processing Through Brief Training on Novel Grammatical Morphemes: An Event-Related Potential Study. *Language Learning*, *74*(S1), 185–223. https://doi.org/10.1111/lang.12642

Yu, M., Li, Y., Mo, C., & Mo, L. (2017). Newly learned categories induce pre-attentive categorical perception of faces. *Scientific Reports*, *7*(1), 14006. https://doi.org/10.1038/s41598-017-14104-6