



Neural mechanisms of bilingual speech perception: the role of the executive control network in managing competing phonological representations

Research Article

Cite this article: García-Sierra, A. and Ramírez-Esparza, N. (2025). Neural mechanisms of bilingual speech perception: the role of the executive control network in managing competing phonological representations. *Bilingualism: Language and Cognition* 1–18. <https://doi.org/10.1017/S1366728925000148>

Received: 07 March 2023
Revised: 29 October 2024
Accepted: 18 January 2025

Keywords:
bilinguals; executive control network; speech perception; sLORETA

Corresponding author:
Adrián García-Sierra;
Email: adrian.garcia-sierra@uconn.edu

This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

Adrián García-Sierra¹ and Nairán Ramírez-Esparza²

¹Department of Speech, Language & Hearing Science, CT Institute for Brain and Cognitive Science, University of Connecticut, Storrs, CT 06269-1085, USA and ²Department of Psychological Sciences, University of Connecticut, Storrs, CT 06269-1085, USA

Abstract

This study investigated the neural mechanisms underlying bilingual speech perception of competing phonological representations. A total of 57 participants were recruited, consisting of 30 English monolinguals and 27 Spanish-English bilinguals. Participants passively listened to stop consonants while watching movies in English and Spanish. Event-Related Potentials and sLORETA were used to measure and localize brain activity. Comparisons within bilinguals across language contexts examined whether language control mechanisms were activated, while comparisons between groups assessed differences in brain activation. The results showed that bilinguals exhibited stronger activation in the left frontal areas during the English context, indicating greater engagement of executive control mechanisms. Distinct activation patterns were found between bilinguals and monolinguals, suggesting that the Executive Control Network provides the flexibility to manage overlapping phonological representations. These findings offer insights into the cognitive and neural basis of bilingual language control and expand current models of second language acquisition.

Highlights

- Bilinguals and monolinguals heard two speech sounds in two language contexts.
- Speech sounds represented two phonemic categories in English and one in Spanish.
- Bilinguals exhibited interference from the non-target language.
- The Executive Control Network (ECN) was active in both language contexts.
- The ECN regulates parallel activation and adjusts processing pathways flexibly.

1. Introduction

The concept of language activation and selection is central to bilingualism research, as it involves the ability to actively use one language while inhibiting another. The present investigation examines bilinguals' brain mechanisms in language selection during the perception of speech sounds that have no lexical meaning but compete for phonological representations across languages.

Speech perception centers on the mechanisms of activation, processing, and representation in the brain. In the auditory domain, speech perception typically follows a sequential pathway that is driven purely by the speech signal (bottom-up), but prior knowledge or expectations can refine the perception of individual sounds (top-down) (McClelland & Elman, 1986; see Norris et al., 2000 for an opposite idea). Using a purely speech-driven approach, activation is understood as the process by which certain auditory dimensions or cues are prioritized or activated during the categorization of sounds (Holt & Lotto, 2006; Poeppel et al., 2008). The concept of processing refers to how the auditory system interprets and categorizes input, involving the neural and computational mechanisms that operate on representations (acoustic, phonetic, phonological, and lexical) to facilitate speech perception and production (Li et al., 2010; Obleser et al., 2005; Poeppel et al., 2008). Finally, the concept of representation involves the mental encoding and storage of phonetic, acoustic, phonological, and lexical information within speech processing. It describes how different aspects of speech signals are encoded in the brain, playing a crucial role in transforming acoustic input into meaningful linguistic information (Hickok & Poeppel, 2007). These representations, stored and organized in the brain, provide the foundation for recognizing and producing language and are shaped by linguistic experience.

Central to bilingualism research is the concept of language activation and selection, which entails the ability to use one language while inhibiting the other, introducing the notion of

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.



language control. This implies that, alongside speech perception processes, language selection mechanisms operate concurrently in the bilingual brain and must be regulated by a language control mechanism. The “Adaptive Control Hypothesis” proposed by Green and Abutalebi (2013) describes activation, processing, and representation in the context of language use in bilinguals. Activation refers to the engagement of specific neural areas during language selection, with both languages being simultaneously activated in the bilingual mind, even when only one is used. This concurrent activation creates competition between languages and requires control processes (e.g., the Executive Control Network [ECN]) to manage the selection of the appropriate language for a given context. Processing is defined as the mechanism by which bilinguals select and control representations in working memory, ensuring alignment with communicative goals. This involves various cognitive control processes, including conflict monitoring, executive control, interference suppression, and task switching, all adapting based on the interactional context. Finally, representation in bilinguals pertains to the verbal and nonverbal representations maintained in working memory to achieve communicative goals, encompassing the full range of both languages’ elements, from words and syntax to concepts.

The present investigation explores speech sounds with different mental representations across languages (e.g., the same sound representing the phoneme [k] in Spanish and [g] in English). Since the representations of these sounds compete for phonemic membership, we will present them in two language contexts (Spanish and English) to determine whether control mechanisms are involved in their perception, even when only one language is being used. We will rely on Event-Related Potentials (ERPs) to assess brain electrical activity associated with perceptual and cognitive processes (Luck, 2014) and rely on standardized Low Resolution Brain Electromagnetic Tomography (sLORETA) (Pascual-Marqui, 2002; Pascual-Marqui et al., 1994) to pinpoint the brain areas associated with those processes. This approach allows us to infer how phonemic sounds are represented in both languages. Specifically, sLORETA will help determine whether the discrimination of speech sounds with competing phonemic representations is managed by control mechanisms, such as the ECN.

Even though there is a well-established body of research showing that bilinguals rely on a control mechanism when accessing competing lexical representations across languages (Abutalebi et al., 2007; Abutalebi et al., 2011; Garbin et al., 2011; Green, 1998; Green & Abutalebi, 2013; Liu et al., 2021; Marian et al., 2014; Marian et al., 2017; Perani et al., 2003; Rodríguez-Pujadas et al., 2013; Shen et al., 2020; Sulpizio et al., 2020), most speech perception models have only recently begun to incorporate this aspect into their frameworks. For instance, the L2LP model (Escudero, 2005; Escudero et al., 2009; Escudero & Boersma, 2004; Van Leussen & Escudero, 2015) and the Speech Learning Model (SLM) (Flege, 1995; Flege et al., 2003), along with its revised version (SLM-r; Bohn & Flege, 2021), provide theoretical frameworks for understanding how second-language (L2) learners perceive and acquire new sounds. While these models emphasize the influence of a learner’s first language (L1) on L2 perception, they differ in their explanations of how L2 sound categories are formed and stabilized. Only recently has the L2LP model explicitly integrated language control mechanisms into its framework, as proposed in Escudero and Yazawa (2024) through the Language Mode Activation Hypothesis, which suggests that in the final stage of acquisition, both L1 and L2 perceptual systems remain active, with their influence varying depending on proficiency, language exposure, and cognitive control mechanisms. The

present investigation aims to empirically test this hypothesis and determine whether bilingual language selection operates under this newly proposed framework.

The L2LP model (Escudero, 2005) and the SLM (Flege, 1995; Flege et al., 2003) offer theoretical frameworks to explain how learners acquire phonetic categories in an L2, with a particular focus on the influence of the first language. Both models recognize that L2 learners initially rely heavily on their existing L1 phonetic system when encountering L2 sounds, mapping these new sounds onto the closest L1 equivalents. This reliance often leads to perceptual difficulties, especially when the L2 contains sounds absent in the L1, as unfamiliar L2 sounds are assimilated into existing L1 categories (e.g., English [g] and [k] are not contrastive in Spanish). Consequently, the accuracy of both perception and production in the L2 is affected since the nuanced distinctions of the L2 sounds may not be accurately captured.

Despite these foundational similarities, the L2LP and SLM diverge significantly in their explanations of the mechanisms underlying L2 phonetic acquisition and in their predictions about the ultimate outcomes for learners. The L2LP model employs the Gradual Learning Algorithm (GLA) (Boersma, 1998; Escudero, 2005) to simulate L2 development, mirroring the sequential processes observed in L1 acquisition. This approach involves perceptual learning, where learners adjust to new phonetic inputs, and representational learning, which entails forming new phonological categories. The L2LP model is optimistic about adult learners’ potential to achieve native-like perception through appropriate input and learning mechanisms, emphasizing that rich and sufficient L2 input can compensate for any reduced neural plasticity compared to children.

In contrast, the SLM proposes two specific mechanisms of interaction between L1 and L2 phonetic systems: category assimilation and category dissimilation. Category assimilation occurs when L2 sounds are not sufficiently distinct from L1 sounds, leading to merged phonetic categories (Davidian & Flege, 1984; Flege, 1991; Flege, 1992; Flege, 1992; Flege et al., 2003; Williams, 1977). In this scenario, L2 sounds are consistently mapped onto existing L1 categories, hindering the formation of new, distinct L2 categories. For example, native speakers of French or Spanish may produce the English /t/ with an intermediate voice onset time (VOT) (Lisker & Abramson, 1964; Abramson & Lisker, 1967), value blending characteristics of both languages (Flege et al., 2003). Category dissimilation happens when new L2 categories are established and diverge from L1 categories in common phonetic space to maintain perceptual contrast (Lindblom, 1990). This mechanism reflects the learner’s effort to enhance the distinction between L2 sounds and their L1 counterparts by minimizing overlap within the shared phonetic space (Flege & Eefting, 1987a).

Relevant to the present investigation, the two models differ in their conceptualization of language activation and the role of contextual influence. The SLM emphasizes integration and interaction within a unified phonetic space (Lindblom, 1990) but does not explicitly address how linguistic context modulates the activation levels of each language. While the SLM considers the possibility of mutual influence in a common phonetic space, where L1 categories affect L2 perception and production, and vice versa, it does not explore the contextual activation of separate linguistic systems. In contrast, the L2LP model introduces the Language Mode Activation Hypothesis, an extension of Grosjean’s Language Mode Framework (1998, 2001, 2008), which posits that bilinguals modulate the activation levels of their languages along a continuum. This continuum ranges from a monolingual mode, where one language dominates, to a bilingual mode, where both languages are

integrated based on contextual cues. The latest account of L2LP incorporates cognitive control as a key component of this hypothesis, proposing that L2 learners can activate both L1 and L2 perceptual systems to varying degrees depending on linguistic demands. This framework allows for parallel or selective activation of two grammars, enabling bilinguals to dynamically manage perceptual resources in real time. Additionally, the Language Mode Activation Hypothesis offers an alternative explanation for the merged phonetic categories proposed by SLM. Rather than viewing assimilation as a fixed process, it suggests that simultaneous activation of both languages can lead to intermediate perceptual representations, reflecting a dynamic interplay between the two languages within the learner's cognitive framework.

There is extensive behavioral research demonstrating that a specific language context can establish a language mode and alter the phonemic categorization of speech sounds with competing representations. For instance, García-Sierra et al. (2009) presented a speech continuum ranging from /ga/ to /ka/ by manipulating VOT. The VOT continuum was presented to bilingual Spanish-English speakers and monolingual English speakers in separate Spanish and English contexts. The results showed that bilinguals, but not monolinguals, shifted their phonemic boundaries based on the active language (i.e., perceiving more /ga/ sounds in the English context than in the Spanish context). This led to the concept of a “double phonemic boundary” or “double phonemic representation” (García-Sierra et al., 2012), demonstrating bilinguals' ability to shift phonemic boundaries based on the active language. The concept has been examined across various language pairs using VOT continua, including English and Spanish (Casillas & Simonet, 2018; Elman et al., 1977; García-Sierra et al., 2021; Gonzales & Lotto, 2013; Lozano-Argüelles et al., 2021; Wig & García-Sierra, 2020; Williams, 1977), English and French (Caramazza et al., 1973; Gonzales et al., 2019; Hazan & Boulakia, 1993), English and Dutch (Flege & Eefting, 1987b), and English and Greek (Antoniou et al., 2010; Antoniou et al., 2012). It has also been explored in studies on vowel perception (Yazawa et al., 2020).

Although the “double phonemic representation” research suggests that bilinguals can modify the phonemic representation of pre-lexical sounds based on linguistic contexts, other models, such as the L2LP model (Van Leussen & Escudero, 2015), propose that bilinguals' ability to distinguish between sounds or group them into categories can shift without altering the underlying mental representations. In other words, listeners may refine or adjust their perception of the physical characteristics of speech sounds (such as vowels and consonants) before these sounds are associated with any lexical meaning or phonological representations. Spivey and Marian (1999) have similarly suggested that phonetic categories may be more flexible at the processing level rather than at the representational level. This implies that when individuals process speech sounds, their brains actively interpret and integrate acoustic information, potentially adjusting or reorganizing phonetic categories as needed for efficient communication. This adjustment could involve merging previously distinct sounds or differentiating sounds that were once perceived as similar, depending on contextual or language-specific cues.

In the present investigation, by relying on sLORETA and comparing brain activation patterns between groups, we aim to gain insights into how these processing adjustments occur and whether they are associated with specific neural control mechanisms, such as those within the ECN. This approach will help us understand if bilinguals' brain activity reflects dynamic processing pathways that

accommodate linguistic context without necessarily altering the core phonemic representations.

As previously mentioned, speech perception models (McClelland & Elman, 1986) suggest that prior knowledge or expectations shape the way we perceive speech, refining the perception of individual sounds through top-down processes. The studies discussed above further suggest the presence of a language control mechanism that is sensitive to linguistic contexts, indicating that bilinguals' ability to manage competing phonological representations is shaped by both their prior linguistic knowledge and the contextual demands, such as activating only one language in monolingual settings or managing both languages in bilingual or code-switching contexts. Unfortunately, only a few studies have explored the cognitive processes underlying phonemic representations across language contexts. These studies use ERPs because, unlike behavioral paradigms, ERPs enable continuous monitoring of the processes between stimuli and responses, allowing for the identification of processing stages that are affected by experimental manipulations (Luck, 2014; Shtyrov & Pulvermüller, 2007). Because ERPs provide millisecond-level temporal precision, it is possible to investigate the speed at which processes associated with auditory discrimination are affected during the perception of speech sounds. These studies used the ERP mismatch negativity (MMN) response, which is commonly used to examine speech discrimination. The MMN is typically elicited by presenting repetitive sounds (standard) and randomly introducing a different (deviant) sound that varies in amplitude, intensity, or phonetic category (Näätänen, 1992; Shtyrov & Pulvermüller, 2007). The MMN is observed approximately 200 ms after the onset of deviant sounds, and its amplitude increases as signal discrimination improves (Tiitinen et al., 1994). Importantly, the MMN is elicited without requiring participants' active attention or explicit responses to stimuli, thereby offering an advantage in assessing the effects of language contexts.

García-Sierra et al. (2012) found that the amplitude of the MMN in bilingual Spanish-English speakers was modulated by language contexts. When “ga” and “ka” were presented, the English context elicited a greater MMN amplitude, indicating a more contrastive phonemic perception, whereas the Spanish context produced a smaller MMN amplitude, suggesting the phonemes were perceived as allophones. These findings imply that the auditory discrimination processes differentially “weight” VOT information depending on the linguistic context. In a separate study, Wig and García-Sierra (2020) relied on predictive coding – a theory suggesting that the brain continuously generates predictions about incoming sensory information based on prior knowledge (Garrido et al., 2009) – to interpret their MMN findings. Predictive coding posits that when there is a mismatch between these predictions and the actual sensory input, a prediction error occurs (MMN), prompting the brain to update its predictions to minimize future errors. In the Wig and García-Sierra study, Spanish-English bilinguals and English monolinguals were presented with a series of ten stop consonants with Spanish phonetic characteristics, ranging from prevoiced sounds to short lag sounds (i.e., /da/ to /ta/; respectively). Prevoiced sounds always served as standards, while short-lags functioned as deviant sounds. Participants were required to press a button upon hearing /ta/. Perceptual “errors” were generated by presenting speech sounds with Spanish characteristics in a mismatched English language context (short videos in English). These responses were compared to a matching condition, where the same sounds were presented in a Spanish language context (short videos in Spanish). The results revealed that bilinguals, but not monolinguals, adjusted their conceptual expectations during the early stages

of phonetic discrimination (larger MMN during the English language context, i.e., expecting English sounds but perceiving Spanish sounds). This finding indicates that linguistic contexts can establish expectations about the type of phonetic information to anticipate, and when the incoming speech sound deviates from these expectations, a mismatch response is triggered. This interpretation aligns with behavioral studies that show that bilinguals, unlike monolinguals, can be conceptually cued to an imagined language context. That is, bilinguals can “activate” a language without direct exposure to it (García-Sierra et al., 2021; Gonzales et al., 2019; Lozano-Argüelles et al., 2021). For instance, in Gonzales et al.’s study, bilinguals were told that the speech sounds to be identified (stop consonants) were produced by either an English or French speaker. The results demonstrated clear shifts in the perception of stop consonants based on the imagined language contexts.

Despite substantial evidence that bilinguals can use linguistic context to adjust their phonetic categories, no study has yet examined the language control mechanisms involved in perceiving speech sounds without lexical information but with overlapping phonological representations between languages. While the above ERP studies offer a valuable method for studying the cognitive aspects of auditory discrimination, the observed amplitude shifts only suggest changes in processing without identifying the specific brain regions involved. The present investigation aims to localize the brain areas associated with differences in auditory discrimination across language contexts, which will allow us to determine whether regions related to the ECN are involved. If the ECN is involved in the processing of speech sounds with overlapping phonological representations, it would suggest that bilinguals actively recruit cognitive control mechanisms to manage competing phonemic categories across languages. This would provide crucial insight into how the brain dynamically adjusts to language context and could offer a neural basis for the ability to resolve phonetic competition. Identifying the brain regions involved would enhance our understanding of the interaction between language control and auditory discrimination, offering a more comprehensive view of bilingual language processing.

2. Goals and overview

The primary goal of this study is to investigate the brain mechanisms involved in bilingual auditory discrimination, focusing on how cognitive control processes modulate speech sound discrimination across different language contexts. Increased brain activity in prefrontal areas, such as the dorsolateral prefrontal cortex (DLPFC), anterior cingulate cortex, and Inferior Frontal Gyrus (IFG), has been observed in bilingual individuals, reflecting their enhanced cognitive control during language processing (Abutalebi et al., 2007; Green & Abutalebi, 2013; Hernandez et al., 2001). In the present study, we aim to localize brain regions associated with these control mechanisms by examining the MMN response, and we will use sLORETA to localize the brain regions associated with these processes. The MMN has been linked to two mechanisms: a sensory memory system with temporal generators (Alho et al., 1998; Rinne et al., 2000; Scherg et al., 1989; Tervaniemi et al., 2000) and a comparator-based mechanism tied to prefrontal areas (Giard et al., 1990; Gomes et al., 2000; Maess et al., 2007; Opitz et al., 1999; Roland, 1981, 1982). Previous research suggests that MMN frontal generators are lateralized, with the right hemisphere associated with tone paradigms (Levänen et al., 1996) and the left hemisphere involved in language paradigms (Näätänen et al., 1997;

Tervaniemi et al., 2000). These frontal regions are believed to engage in processes that modulate the deviance detection system in the temporal cortex, supporting auditory change detection (Doeller et al., 2003; Garrido et al., 2009).

The present study employs both within-group and between-group designs, in which bilingual and monolingual participants will passively listen to stop consonants while watching movies in both Spanish and English without making any phonetic judgments. The within-group comparison for bilinguals aims to explore whether a language control mechanism is activated by examining their sLORETA responses across different language contexts. In contrast, the between-group comparison with monolinguals will evaluate the differences in brain activation between the two groups, focusing specifically on how bilinguals process speech sounds with competing phonological representations – an issue that monolinguals do not encounter. By comparing both groups and language contexts, we aim to determine whether (1) bilinguals can adjust how they perceive and categorize speech sounds (such as vowels and consonants) without altering their core mental representations, and this flexibility may be supported by the ECN, or (2) bilinguals can refine their perception of speech sounds based on context or experience, even before these sounds are associated with word meanings or stored in the phonological system, with the ECN playing a role in regulating these pre-lexical adjustments. This approach will offer valuable insights into the neural mechanisms that underlie language control in bilingual individuals.

3. Methods

3.1. Participants

Participants were recruited at the University of Connecticut for a large-scale study of bilinguals’ social language interactions and both cortical and subcortical auditory processing. This study involved 74 recruits, yet only 64 participated in the ERP segment. Additionally, 5 participants did not show clear ERP responses, and the final sample comprised 57 normal-hearing students aged 18–23 years. All participants passed a hearing screening at or below 20 dB HL for octave frequencies ranging from 500 to 8000 Hz. Participants completed a language-screening questionnaire to determine whether they were monolingual English speakers ($n = 30$, including 7 males) or Spanish-English bilinguals ($n = 27$, including 7 men). Monetary incentives were offered to all the participants.

Bilingual participants reported that their caregivers originated from various areas of Latin America. The language background questionnaire evaluated exposure to and use of Spanish and English from childhood to adulthood. This report includes 26 bilinguals for linguistic background questions¹ (one bilingual declined to answer the questionnaire). Questions about exposure were presented on a Likert scale ranging from 1 to 5 (1 = 100% Spanish; 2 = 75% Spanish, 25% English; 3 = 50% Spanish, 50% English; 4 = 25% Spanish, 75% English; and 5 = 100% English). Figure 1 displays violin plots illustrating bilinguals’ language exposure and use from infancy to the time of the experiment. The data shows a distinct transition from predominantly Spanish exposure and usage to predominantly English exposure and use. English monolingual participants

¹In some instances, participants did not complete one of the three measures of bilingualism (questionnaires, LENA, and PPTV). However, all participants had at least two measures, and therefore, the two implemented measures of bilingualism provided sufficient information to classify them as bilingual participants.

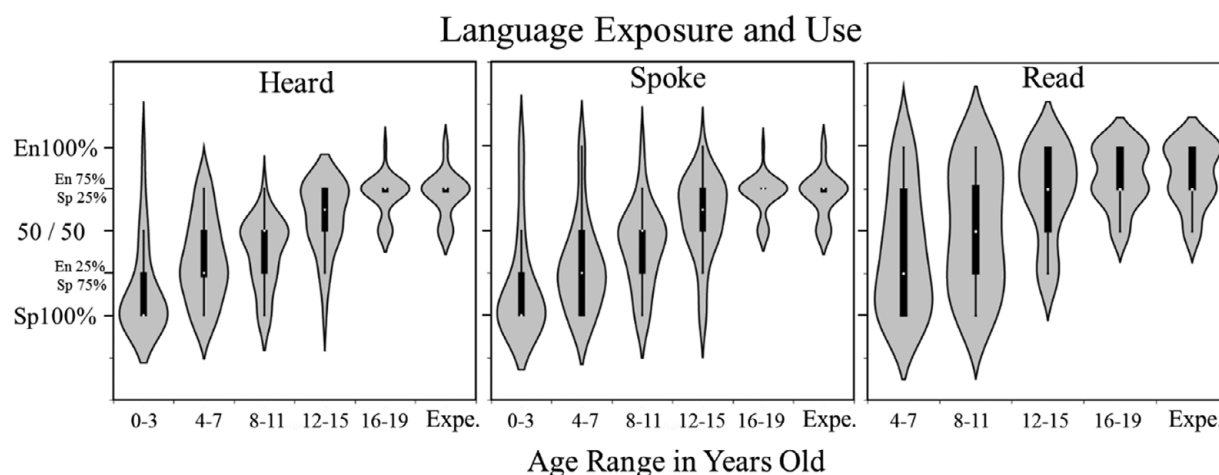


Figure 1. Violin plots for bilingual participants' language exposure and use from birth to the date of the experiment. White dots represent the median.

reported being raised entirely in English-speaking families, with only incidental exposure to Spanish.

Another series of questions measured bilinguals' present language confidence. These questions were presented separately for English and Spanish on a 1-to-5 Likert scale (1 = I cannot speak the language, have only a few words or phrases, and I cannot create sentences; 5 = I have native-like proficiency with few grammatical errors and strong vocabulary). Bilinguals' confidence in speaking English averaged 4.92 (SD = 0.276) and in Spanish 4.58 (SD = 0.634). Bilinguals' confidence in understanding English averaged 5.0 (SD = 0.00) and in Spanish 4.81 (SD = 0.402). Please see [Supplementary Tables \(S1–S4\)](#) for further descriptive questions assessing confidence in hearing, speaking, reading in relation to age and language use with family members and friends.

Bilinguals and monolinguals wore digital recorders for two days in this large-scale study. Another study describes both groups' everyday activities and language use (e.g., Ramírez-Esparza et al., 2024). We used two-day language recordings of the digital recorders to verify bilinguals' responses to the language questionnaire. A total of 27 bilinguals and 30 monolinguals were analyzed. Coders examined pre-selected and randomized speech-active parts from the digital recorders' audio files (Ramírez-Esparza et al., 2024). Bilinguals spoke 57% (SD = 16) of the time, and monolinguals spoke 62% (SD = 21) of the time. A t-test showed no significant differences in the amount of time that either group spoke $t(54) = 1.031$, $p > .05$, 95% CI [−5, 15]. Bilinguals spoke 50.2% in English, 3.7% in Spanish, and 3% code-switched. Despite speaking Spanish, bilinguals' dominant language was English at the time of the experiment, as evidenced by the digital recordings and linguistic background questionnaire.

The English-Spanish NIH Peabody picture vocabulary test (PPVT) assessed vocabulary (Gershon et al., 2014). This computer-adaptive receptive vocabulary test customizes questions based on prior responses. Participants are presented with an audio recording of a word and four images on a computer screen, and they are asked to select the image that best represents the word. The PPVT evaluates vocabulary abilities that are more dependent upon past learning experiences and are consistent across the lifespan. Age-adjusted scores over 100 indicate normal vocabulary (Dunn, 1997). Twenty-five bilinguals were assessed using the PPVT in

both languages. The age-adjusted score for English was 107.76 (SD = 13.5), and for Spanish, it was 107.32 (SD = 18.0).

Overall, the bilinguals recruited in this study exhibit a typical language development pattern that is commonly observed in numerous young bilingual individuals residing in the United States. That is, they are exposed to languages other than English throughout infancy and early childhood, but as they grow and attend school, English becomes their dominant language (Kohnert et al., 1999). This group of bilinguals is referred to as heritage bilinguals (Valdés, 2005).

3.2. Stimuli

We employed a shorter (190 ms) version of the stimuli ("ga" and "ka") used by García-Sierra et al. (2009, 2012) to capture brainstem potentials and ERPs simultaneously. Here, we report only ERPs, while ABRs are reported. ASL software from the Computerized Speech Lab (CSL) system was used to create stimuli using the cascade method (Klatt, 1980). "KA" (standard stimulus) had a + 50 ms VOT, while "GA" (deviant stimulus) had a + 15 ms VOT. Formant transitions were linearly interpolated from velar stop consonant values (180, 1725, 1725, 3200, and 3500 Hz for F1-F5) to vowel /a/ values (750, 1200, 2450, 3200, and 3500 Hz for F1-F5). Simulating consonant release required a 10-ms turbulent noise source at 60 dB amplitude. Aspiration was simulated using an aspiration source (AH) at 62 dB after consonant release and before vowel onset. The initial 100 ms formant transition is interpolated between 45 and 65 dB from F0. Insert earphones (Etymotic ER3C) delivered stimuli at approximately 67–68 LAeq dB.

3.3. Language contexts

Data collection involved presenting movies in the target language, with a Spanish-language film shown during the Spanish-language context and an English-language film during the English-language context. This approach was applied to both the bilingual and monolingual participants. To mitigate potential order effects, a single language context was established for each experimental session, with sessions separated by a minimum of three days. Additionally, the language context sessions were counterbalanced. To control the influence of lip movements on speech perception, cartoons were used in the experimental setup (Yoshida et al., 2010).

The experimental design encompassed both the movie and stimulation blocks. Each session began with a movie block in which participants watched a film in the target language, set at a comfortable audio level. During the stimulation blocks, the audio of the movie was entirely muted and subsequently reinstated to a comfortable level. A critical aspect was the continuous display of captions throughout both the movie and stimulation blocks. The movie blocks had a duration of approximately 90 s and the stimulation blocks lasted 60 s. The protocol involved 12 alternating cycles of the two block types. Each stimulation block consisted of 80 standard and 20 deviant sounds, in addition to a set of 10 stimuli introduced for familiarization before each stimulation block, which were excluded from the final data average. The entire duration of each language context session lasted approximately 30 min. A detailed visual representation of this setup is shown in [Figure S1](#) of the [supplementary materials](#).

3.4. Electroencephalogram (EEG)

An elastic cap with 64 electrodes positioned according to the worldwide 10/10 system was placed on the participants' scalps, and a camera (Cap Track Brain Vision) recorded the electrodes' x, y, and z coordinates. The anatomical markers for electrode digitization were the nasion and tragi of both ears. The digitized electrode locations were used to create sLORETA files. The EEG was referenced to FCz in DC mode and re-referenced offline to an average reference for analysis. ActiChamp amplifiers (24-bit A/D converter) recorded the electroencephalogram, and StimTrack (Brain Vision) delivered clicks (1 ms) for each stimulus. Offline filters at 0.10 Hz (6 dB/oct forward) and 40 Hz (12 dB/oct zero phase) were implemented. BESA Research 7.1 procedures (BESA GmbH, Gräfelfing, Germany) corrected eyeblinks measured by Fp1 and Fp2 electrodes. EEG segments with electrical activity exceeding ± 100 mV were eliminated from the final average, and electrode impedances were maintained below 5 k Ω . ERPs were averaged offline from 470 ms EEG segments with a 100 ms pre-stimulus baseline period. Baseline correction was performed in relation to the pre-stimulus time.

A monolingual English speaker spoke English to both groups to ensure equal language exposure. A shielded, soundproof booth was used to collect the data. The /ka/ and /ga/ sounds were always standard and deviant stimuli, respectively, and the presentation of these stimuli was pseudo-randomized (i.e., 960 standard and 240 deviant sounds were played). Deviant sounds never occurred consecutively, and at least three consecutive standard sounds preceded them. The final average included 720 standard sounds, excluding those occurring after a deviant. The inter-stimulus interval (offset-to-onset) was 380 ms. Participants were told to focus on the movie and ignore the stimuli. The MMN was calculated by subtracting the response to standard sounds from the response to deviant sounds (deviant minus standard).

3.5. Accepted epoch count in ERP averages

The number of accepted epochs for monolinguals for the standard sound was 712.94 (SD = 26.55) during the English language context and 718.6 (SD = 9.19) during the Spanish language context. For the deviant sound, the number of accepted epochs for monolinguals was 233.77 (SD = 9.19) during the English language context and 235.48 (SD = 7.51) during the Spanish language context. The number of accepted epochs for bilinguals for the standard sound was 710.10 (SD = 25.06) during the English language context and

702.15 (SD = 46.67) during the Spanish language context. For the deviant sound, the number of accepted epochs for bilinguals was 233.40 (SD = 8.86) during the English language context and 230.77 (SD = 15.03) during the Spanish language context.

3.6. Statistical analysis for ERPs

Data-driven analyses tested the presence of the MMN (standard ERP versus deviant ERP) and its amplitude modulation (deviant minus standard) between language contexts. BESA Statistics 2 (BESA GmbH, Gräfelfing, Germany) was used for permutation testing and data clustering to analyze the ERP amplitudes. This multi-step approach assumes that statistical effects observed over extended periods and across adjacent channels are unlikely to be coincidental. Initially, a parametric test identifies time periods with pronounced effects, and the t-values in these regions are summed to form cluster values. Each region with a substantial effect undergoes this process, representing cluster values in both time and space. Thus, a large cluster value indicates a considerable difference in the time domain across numerous surrounding electrodes, whereas a smaller cluster value indicates a significant difference in one or a few neighboring electrodes. This study used a 4.5 cm channel neighbor distance. Subsequently, BESA repeats step 1 using permutation tests (10,000 in this case). This test determines whether the cluster value probabilities across experimental conditions (or participants) are interchangeable. Consequently, all permutations contribute to a cluster value distribution and directly ascertain the α -error of the initial cluster value from step 1. In other words, this process determines whether the initial cluster value obtained in step 1 is as probable as any cluster value from other permutation tests. This type of analysis is performed to control for Type I errors due to the large number of data points in ERP responses (see: Bullmore et al., 1999; Ernst, 2004; Maris & Oostenveld, 2007). Importantly, since ERP data are reported in time and space, the ERP time range in which a substantial cluster is found varies among channels. Therefore, the significant difference observed at electrode "x" will be similar but not identical to that at electrode "y." Unless otherwise specified, we used two-tailed paired and independent t-tests for within- and between-group comparisons.

3.7. Statistical analysis for sLORETA

ERPs were used to calculate neural generators using standardized sLORETA (Pascual-Marqui, 2002; Pascual-Marqui et al., 1994). sLORETA is a distributed source modeling method that makes assumptions regarding the distribution rather than the number of current source densities. Unlike dipole-fitting approaches, sLORETA does not require a priori localization information. sLORETA computes the smoothest possible three-dimensional (3D) current distribution in the brain that generates the observed scalp field. The sLORETA algorithms calculate current density values (in amperes per square meter; A/m²) for 6,239 gray matter voxels of the brain compartment, each with a spatial resolution of 5 mm x 5 mm x 5 mm. Anatomical regions are labeled in accordance with the probabilistic MNI-152 template from the Brain Imaging Center of the Montreal Neurological Institute (MNI; Mazziotta et al., 2001) and the Co-Planar Stereotaxic Atlas of the Human Brain (Lancaster et al., 2000; Talairach & Tournoux, 1988). The validity of sLORETA has been confirmed in several studies, including those combining EEG and fMRI (see Vitacco et al., 2002). We used statistical non-parametric mapping (SnPM) (Nichols & Holmes, 2002) to compute voxel-by-voxel with 10,000 permutations for within- or between-

group comparisons. Comparisons of current density distribution, both between and within groups, were conducted using the log-F-ratio statistic. The results are represented as maps of the log-F-ratio statistics for each voxel corrected for $p < 0.05$. Importantly, the SnPM approach corrects for multiple comparisons and does not require Gaussianity assumptions. Identical protocols were employed in conducting correlations; however, the results are reported in terms of Pearson r -values.

4. Results

Initial analyses confirm the presence of the MMN. We compared standard and deviant ERP responses and MMN polarity inversion between mastoid electrodes and frontal electrodes (Alho, 1995). Both groups exhibited significant MMN responses and polarity inversion in both language contexts. Bilingual ERP figures in both language contexts are shown in the [supplementary section \(Figure S2\)](#) with corresponding analyses (Analysis 1). Monolingual ERP figures in both language contexts are similarly presented in the [supplementary section \(Figure S3 and Analysis 2\)](#).

The analyses of MMN amplitude modulation and sLORETA are organized as follows: Initially, the entire MMN response window (–100 to 470 ms) is analyzed for MMN amplitude modulation in both language contexts, separately for bilinguals and monolinguals. Subsequently, we compare MMN across language contexts for both groups. In the sLORETA analysis, we report MMN time windows showing significant MMN differences. First, sLORETA is compared between language contexts for each group, and then sLORETA comparisons are made across language contexts for both groups. Finally, the sLORETA values are correlated with bilinguals' language dominance shifts.

4.1. MMN responses in different language contexts

4.1.1. Comparison of bilinguals' MMN response between language contexts

Two clusters appeared in the MMN comparisons of the bilinguals. Cluster 1's cluster value was -2284.83 between 138 and 333 ms following stimulus onset for electrodes Fz, F3, F7, FC5, FC1, C3, T7, CP5, Cz, C4, FC2, F4, AF7, AF3, F1, F5, FT7, FC3, C1, C5, CP3, CPz, C2, and FC4. The cluster value exhibited a significantly different probability distribution between the language contexts ($p = 0.02$). This finding strongly suggests a more negative MMN response in the English language context (mean = $-.405 \mu\text{V}$, SD = $.351$) than in the Spanish language context (mean = $-.021 \mu\text{V}$, SD = $.354$). Cluster 1 data-driven analysis is illustrated on the left side of [Figure 2A](#). The voltage map shows a negative voltage distribution in the left frontal and central electrodes.

Cluster 2 was observed between 355 and 469 ms after stimulus onset, with a cluster value of 1832 for electrodes Fz, FC1, Cz, C4, T8, FT10, FC6, FC2, F4, F8, AFz, F1, CP4, C6, C2, FC4, FT8, F6, AF8, AF4, and F2. The cluster value exhibited a significantly different probability distribution between the language contexts ($p = 0.03$). Cluster 2 shows that bilinguals' different waveform in the investigated time range were more positive in the English language context (mean = $.143 \mu\text{V}$, SD = $.455$) than in the Spanish language context (mean = $-.314 \mu\text{V}$, SD = $.416$). The data-driven examination of bilinguals' difference waveforms between the language contexts is depicted on the right side of [Figure 2A](#). The voltage map shows a positive voltage distribution in the right-frontal and central electrodes.

4.1.2. Comparison of monolinguals' MMN response between language contexts

No significant differences in cluster values between language contexts were observed. [Figure 2B](#) presents a visualization of the data-driven analysis of monolinguals' MMN between language contexts.

4.2. Comparison between bilinguals' and monolinguals' MMN responses in different language contexts

4.2.1. English-language context

A cluster value of -3518.01 was found between 186 and 454 ms. The cluster value showed a different probability distribution between bilinguals and monolinguals ($p = .026$) for electrodes F3, F7, FC5, FC1, C3, T7, CP5, CP1, Pz, CP2, Cz, C4, FC2, F4, AF3, F1, F5, FT7, FC3, C1, C5, CP3, CPz, C2, FC4, and F2. This strongly suggests that bilinguals' MMN was more negative (mean = $-.530 \mu\text{V}$, SD = $.332$) than monolinguals' MMN (mean = $-.110 \mu\text{V}$, SD = $.300$). The visualization of the data-driven analyses for bilinguals and monolinguals in the English language context is depicted on the left-hand side of [Figure 3](#). The voltage map shows negative voltage in the left-frontal and central electrodes.

4.2.2. Spanish-language context

A cluster value of 2768.9 was observed between 237 and 469 ms. The cluster value showed a different probability distribution between bilinguals and monolinguals ($p = .026$) for electrodes FP1, F3, F7, FT9, FC5, AF7, AF3, F5, FT7, and C5. This strongly suggests that bilinguals' difference waveform in the investigated time range was more positive (mean = $.329 \mu\text{V}$, SD = $.417$) than monolinguals' difference waveform (mean = $-.133 \mu\text{V}$, SD = $.307$). The right side of [Figure 3](#) presents the visualization of the data-driven analysis for bilinguals and monolinguals in the Spanish language context. The voltage map shows a positive voltage distribution in the left frontal electrodes.

4.3. sLORETA between language contexts

We analyzed the current source density of the MMN using sLORETA. We explored the MMN time regions that showed significant differences between language contexts and between groups.

4.3.1. Bilinguals

Bilinguals' MMN responses in different language contexts showed two clusters. Cluster 1 was observed in the MMN time window between 138 and 333 ms. sLORETA analysis showed a significant difference in the averaged time region between 200 and 333 ms following stimulus onset. The MMN response involved three left frontal brain areas with more cortical activity during the English language context than during the Spanish language context. These areas were the Superior Frontal Gyrus (SFG) (BA 11), Medial Frontal Gyrus (MFG) (BA 10), and Orbital Gyrus (BA 10). [Figure 4A](#) displays the coordinates with the highest log-F-ratio values (1.23, $p < .05$; 1.34, $p < .01$). Please refer to [Table S5](#) in the [supplementary materials](#) for the full list of MNI and Talairach coordinates.

Cluster 2 was observed in the MMN time window between 355 and 469 ms. sLORETA showed a significant difference in the average time region between 355 and 465 ms after stimulus onset. The MMN response engaged one left frontal brain area with more cortical activity in the English language context than in the Spanish language context. The brain region was located in the left SFG (BA 11). [Figure 4B](#) shows the coordinates with the highest log-F-ratio values (1.35, $p < .05$; 1.47, $p < .01$). Please

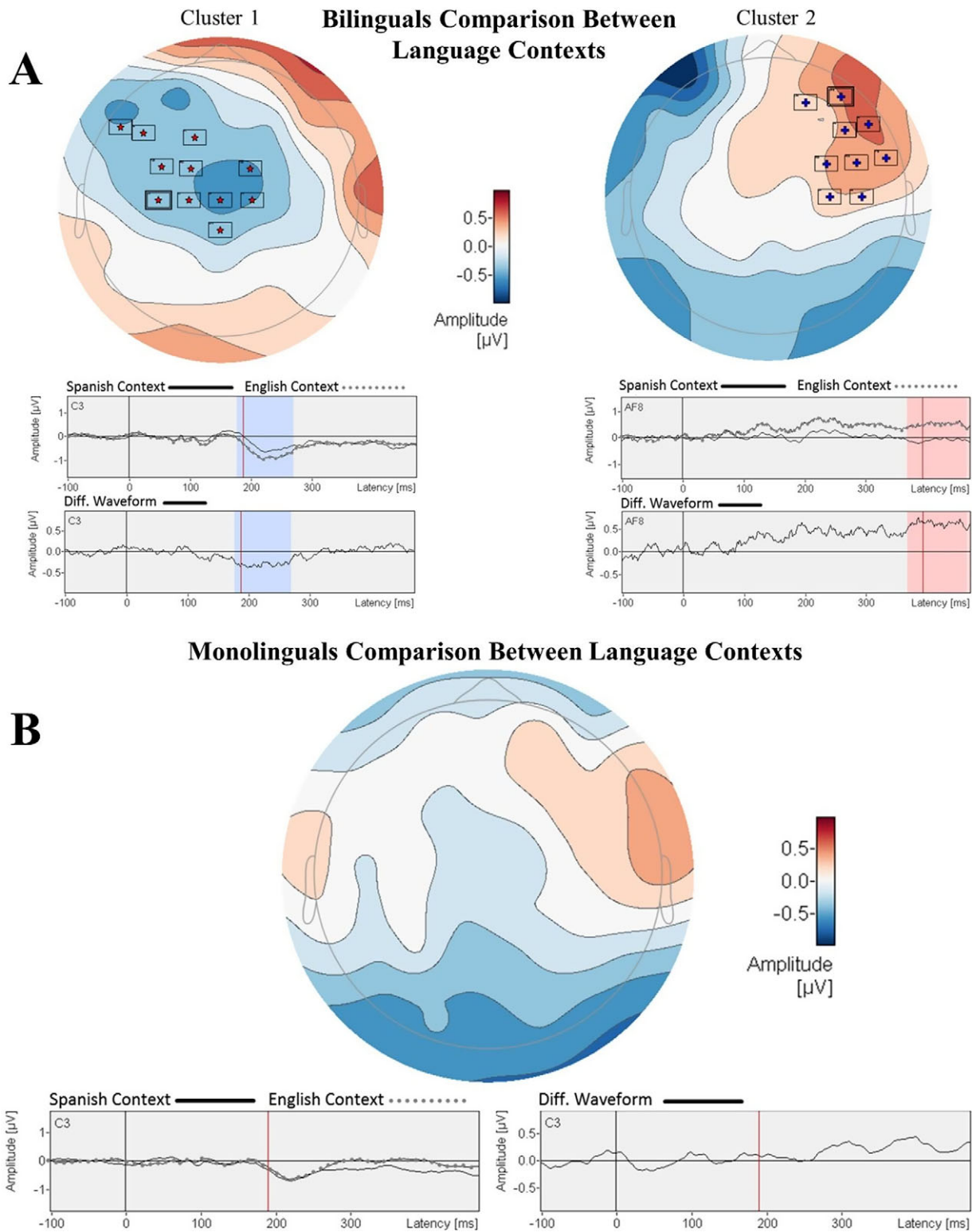


Figure 2. Visualization of the data-driven analysis between language contexts for bilinguals (A) and monolinguals (B). The left side of Section A shows bilinguals' first cluster. The blue shaded areas represent the time intervals where the English mismatch negativities (MMNs) showed a more significant negative amplitude compared to the Spanish MMN. Bilinguals' second cluster is shown on the right side of Section A. The red shaded areas represent the time intervals where the English MMNs showed a more significant positive amplitude compared to the Spanish MMN. The electrodes showing significant differences are displayed with rectangular boxes ($*p < .02$ for cluster 1 and $+p < .03$ for cluster 2). The voltage maps presented in section A represent voltage fluctuations for the difference between the MMNs obtained in both language contexts (English MMN minus Spanish MMN) at approximately 200 ms for cluster 1 (red line) and at approximately 400 ms for cluster 2. The data-driven analysis did not show significant differences for monolinguals (2-B).

Group Comparison in English Language Contexts

Group Comparison in Spanish Language Contexts

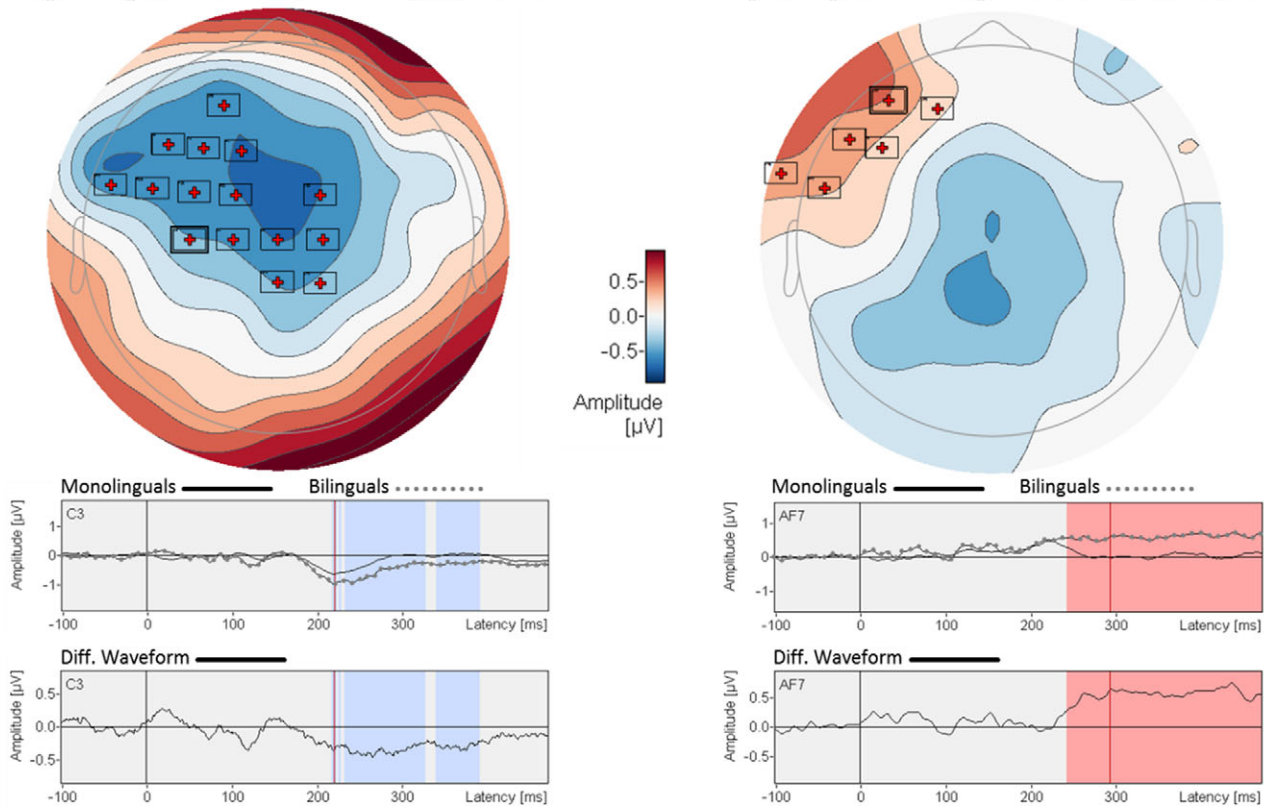


Figure 3. Data-driven analyses comparing both groups' mismatch negativities (MMNs) between language contexts. The English language context (left side) shows a larger MMN for bilinguals when compared with monolinguals. The voltage maps represent voltage fluctuations for the difference between both groups' MMNs (bilinguals minus monolinguals) at approximately 200 ms (red line). The voltage map shows negative values in central and left frontal electrodes. The Spanish language context (right side) shows a larger MMN for bilinguals when compared with monolinguals. The voltage maps represent voltage fluctuations for the difference between both groups' MMNs (bilinguals minus monolinguals) at approximately 300 ms (red line). The voltage map shows positive values in left frontal electrodes.

refer to [Table S6](#) in the [supplementary materials](#) for a full list of MNI coordinates and Talairach coordinates.

4.3.2. Monolinguals

Although monolinguals did not exhibit significant MMN amplitude differences between language contexts, we proceeded to conduct sLORETA analyses in the region where MMN is typically observed (from 200 to 300 ms post-stimulus onset). The results showed no significant difference. This indicates that the brain regions involved in speech processing exhibited similar levels of activation in both language contexts.

4.3.3. Bilinguals' versus monolinguals' sLORETA in the English language context

We explored the MMN time intervals with significant differences between the groups, particularly in the 186–454 ms interval. Nevertheless, given our prior findings that only bilinguals exhibited significant differences between language contexts, we executed a one-tailed log-F-ratio statistical analysis for independent groups. We found a significant difference in the average time region between 200 and 260 ms after stimulus onset. The results showed that the current source densities associated with MMNs involved frontal brain structures with more pronounced cortical activation in bilinguals than in monolinguals (MFG; BA 10, $p < .05$). [Figure 5A](#) shows the log-F-ratio values with significant differences (1.37, $p < .05$; 1.52, $p < .01$) and depicts the coordinates with the highest log-F-ratio value. Please refer to [Table S7](#) in the [supplementary](#)

[materials](#) for a full list of MNI coordinates and Talairach coordinates associated with the MFG.

4.3.4. Bilinguals' versus monolinguals' sLORETA in the Spanish language context

We examined the MMN time intervals in which the groups differed significantly (237–469 ms). We detected significant differences between 350 and 450 ms after the stimulus onset. MMN source densities involved frontal and parietal brain areas with higher cortical activation in bilinguals than in monolinguals. These brain regions were the MFG (BA 8), Postcentral Gyrus (BA 3), Precentral Gyrus (BA 4), and SFG (BA 6). [Figure 5B](#) displays the coordinates with the highest log-F-ratio value (1.12, $p < .05$; 1.25, $p < .01$). Please refer to [Table S8](#) in the [supplementary materials](#) for a full list of MNI coordinates and Talairach coordinates.

4.3.5. Correlating sLORETA values with language shift across the lifespan

We investigated the correlation between bilinguals' source densities in speech processing and shifts in language usage from early childhood to the time of the experiment. To measure the reduction in Spanish usage, we calculated independent averages of the reported percentages for speaking and hearing Spanish during ages 0–3 and at the age of participation in the experiment. We then determined the change by computing the difference between the two averages. A positive value signifies a notable reduction in Spanish use from early childhood to the experimental period.

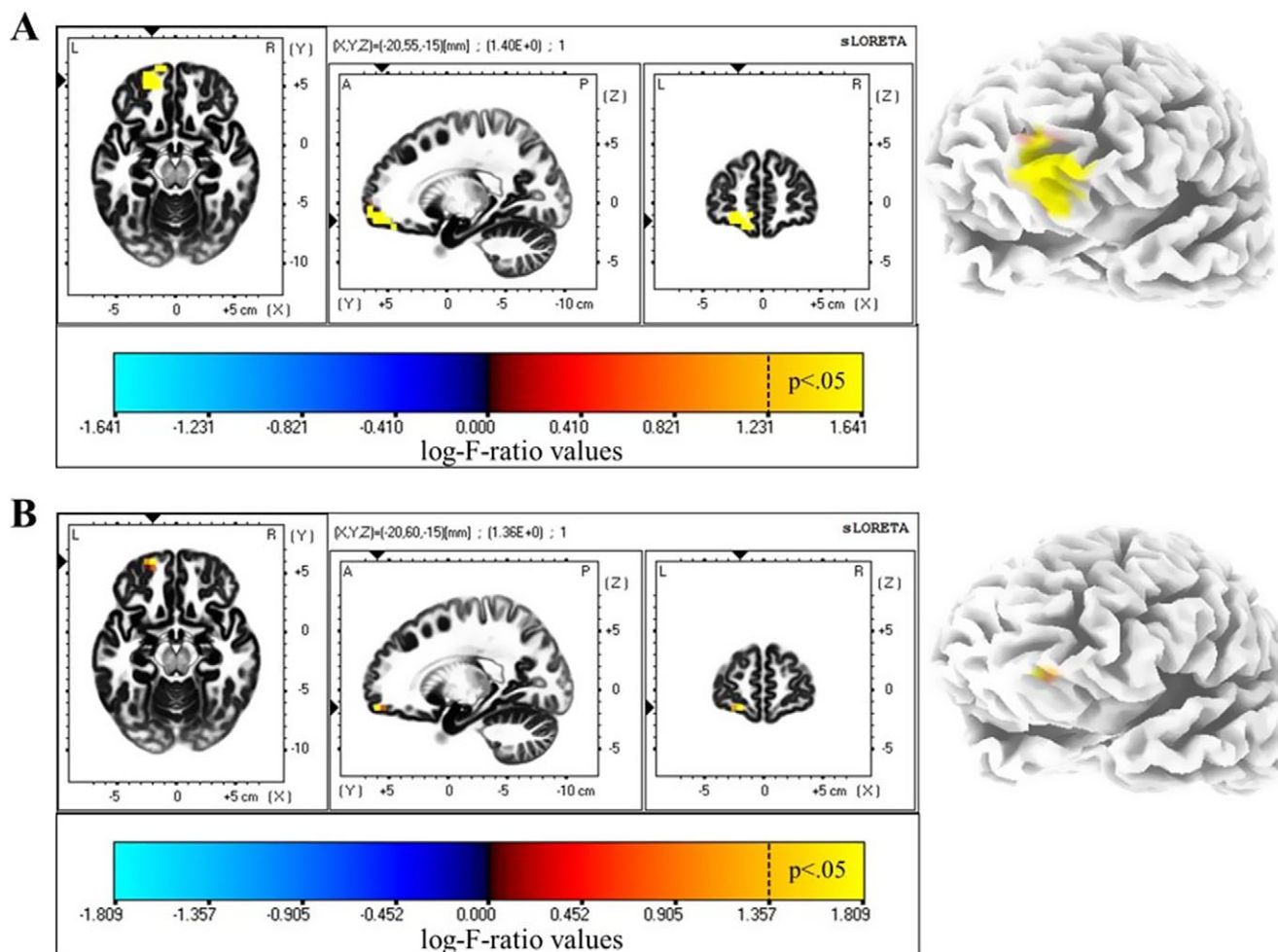


Figure 4. Bilinguals' current source densities between different language contexts. Panel A shows significant differences in BA 10 and 11 in the time frame of 200 to 333 ms post-stimulus onset. Panel B highlights significant differences specifically in BA 11, within the time range of 355 to 465 ms after stimulus onset. The areas of the brain that exhibited statistically higher activation during English language tasks, as compared to Spanish, are indicated in yellow.

Parallel to our prior sLORETA analysis, we examined the MMN time interval (138–333 ms) for significant language context differences in the MMN response. This analysis included 26 bilinguals. The results revealed a positive and significant correlation between reduced Spanish usage and source densities in the left (IFG) (BA 44; MNI -60, 5, 20; Talairach -59, 6, 18, and BA 9; MNI -60, 5, 25; Talairach -59, 6, 23). This correlation was notable between 267 and 305 ms post-stimulus onset and is illustrated in Figure S4 in the Supplementary Materials, which displays the Pearson r -values (.69, $p < .05$). Correlation analysis revealed that bilingual individuals who underwent a significant transition from primarily using Spanish during early childhood to predominantly employing English in young adulthood exhibited more pronounced activation in the IFG during the English language context than in the Spanish context.

5. Discussion

The present study investigated how bilinguals perceive speech sounds with competing phonemic representations across languages, such as the same sound representing different phonemes in Spanish and English. By presenting these sounds in both Spanish and English language contexts, we aimed to determine whether control mechanisms, such as those related to the ECN, were

involved in their perception, even when only one language was being used. Using ERPs to measure brain activity and sLORETA to localize the brain regions associated with these processes, this study sought to provide insight into how phonemic competition is managed in bilinguals. While behavioral research has demonstrated that language context influences phonemic categorization, no study had previously investigated whether active language control mechanisms are engaged during the perception of speech sounds with competing phonological representations in the absence of lexical meaning. This investigation aimed to clarify the role of neural control mechanisms in managing competing phonological representations across languages.

5.1. Brain regions involved in phonetic categorization between language contexts

Our sLORETA results demonstrate that bilinguals show significantly greater activation in the left MMN frontal generators (Brodmann areas 10 and 11: frontopolar cortex and orbitofrontal cortex, respectively) during the English language context compared to the Spanish context. The observed activation in the left frontal cortex aligns with previous studies linking MMN prefrontal generators to a comparator-based mechanism (Giard et al., 1990; Gomes et al., 2000; Maess et al., 2007). This mechanism plays a

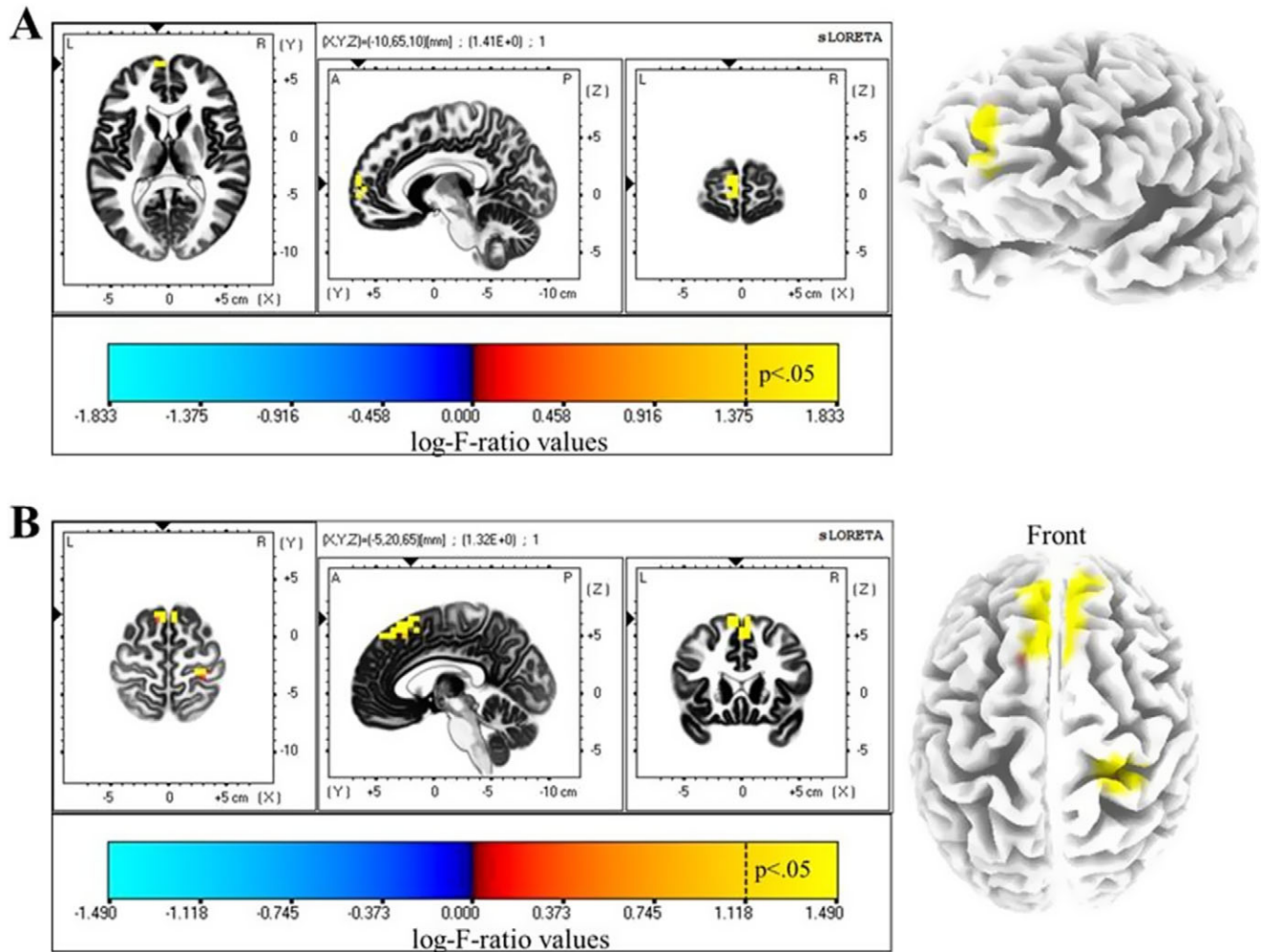


Figure 5. (A) Current source densities between bilinguals and monolinguals during the English language context. Significant differences were found in the averaged time region between 200 and 260 ms after stimulus onset for BA 10. (B) Current source densities between bilinguals and monolinguals during the Spanish language context. Significant differences were found in the averaged time region between 350 and 450 ms after stimulus onset. The frontal activation represents BA 6 and BA 8. The right posterior activation represents BA 3 and BA 4. Yellow coloring depicts brain structures with statistically larger activation in bilinguals when compared to monolinguals.

critical role in modulating the deviance detection system in the temporal cortices, thereby enhancing auditory change detection and control (Doeller et al., 2003; Garrido et al., 2009). In contrast, monolinguals did not show a significant difference in brain activation between language contexts, suggesting that this heightened activation in bilinguals may reflect the additional demands of managing two phonological systems during speech perception. We propose that the activation of the frontal MMN generators is linked to the involvement of the ECN in regulating language-specific processing.

Language selection models vary in their assumptions about the mechanisms underlying language selection. Some models propose that the non-target language must be inhibited (Green, 1998; Green & Abutalebi, 2013; Kroll et al., 2008; van Heuven et al., 2008), while others suggest raising the threshold of the selected language (Blanco-Elorrieta & Caramazza, 2021; Grosjean, 2001) or maintaining distinct resting levels of activation for each language (Dijkstra & van Heuven, 2002; Dijkstra et al., 1998). The heightened activity observed in the left frontal brain area of bilinguals in our study can be interpreted in several ways. It may stem from an inhibitory

process controlled by frontal areas or function as a mechanism to elevate the English phonetic category (Blanco-Elorrieta & Caramazza, 2021; Grosjean, 2001).

However, the frontal activation of Brodmann areas 10 and 11 falls within the domain of executive control, which serves to minimize interference from the non-target language. The prefrontal cortex (PFC), a part of the ECN, is involved in various executive functions such as working memory (Smith & Jonides, 1999), controlled semantic retrieval (Badre et al., 2005; Brian et al., 2006; Gold & Buckner, 2002), phonological retrieval (Poldrack et al., 1999), inhibition of automatic responses, attentional control, planning, and cognitive flexibility to switch between different goals (see Niendam et al., 2012). The ECN comprises the left middle and superior frontal gyri, inferior frontal and orbitofrontal gyri, superior and inferior parietal regions, angular gyri, precuneus, inferior and middle temporal gyri, left thalamus, and right crus (Botvinick et al., 2001; Ridderinkhof et al., 2004; Shen et al., 2020; Koechlin et al., 2003; Miller & Cohen, 2001; Rodriguez-Fornells et al., 2002). Given these perspectives, we posit that the left frontal MMN mechanisms observed are linked to the ECN.

5.2. Integrating the executive control network into speech perception models of second language acquisition

Our findings extend, rather than contradict, existing models of second language speech perception, particularly the SLM (Flege, 1995; Flege et al., 2003) and the L2 Linguistic Perception (L2LP) model (Escudero, 2005). While both models address the interaction between a learner's L1 and L2, they differ in how they conceptualize language activation and contextual influences. The SLM emphasizes the coexistence of L1 and L2 phonetic categories within a shared phonetic space (Lindblom, 1990), where assimilation or dissimilation processes help accommodate new sounds while maintaining phonetic distinctions (Flege et al., 1992; Flege & Eefting, 1988; Flege et al., 2003; Mack, 1990). However, it does not explicitly account for how linguistic context modulates the activation levels of each language.

In contrast, the L2LP model introduces the Language Mode Activation Hypothesis (Escudero & Yazawa, 2024), which builds on Grosjean's Language Mode Framework (1998, 2001, 2008). This hypothesis posits that L2 learners can activate both L1 and L2 perceptual systems to varying degrees depending on contextual demands, supporting a flexible, dynamic interplay between the two languages. The present investigation empirically demonstrates that bilinguals rely on control mechanisms (i.e., ECN) even when processing speech sounds without lexical information, supporting the L2LP framework. These findings confirm that language control occurs at both pre-lexical and lexical stages, reinforcing the idea that bilingual selection operates dynamically, independent of word meaning.

Further supporting this perspective, Van Leussen and Escudero (2015) distinguished between pre-lexical (processing) and lexical (representational) stages in speech perception, arguing that shifts in phonetic categorization occur at the processing level rather than through permanent changes in representation. This aligns with findings from Spivey and Marian (1999), who showed that spoken input in one of a bilingual's languages can automatically activate both mental lexicons in parallel, even in monolingual contexts. Similarly, Ju and Luce (2004) found that bilinguals adjust their processing pathways dynamically, activating lexical representations from both languages when encountering language-specific VOTs.

Expanding on this, Marian and Spivey (2003) proposed that bilingual language processing involves simultaneous activation of both languages, allowing bilinguals to flexibly manage activation levels and competition effects. Their model suggests that processing pathways adapt dynamically based on both bottom-up (stimulus-driven) and top-down (contextual or task-related) factors. This flexibility is further explained by the Adaptive Control Hypothesis (Green & Abutalebi, 2013), which outlines a spectrum of cognitive control processes, such as goal maintenance, selective inhibition, and task switching. These processes operate across multiple levels, from sub-lexical phonetic elements to full lexical representations, ensuring that bilinguals can suppress interference from the non-target language when needed.

5.3. Brain regions involved in phonetic categorization between groups and language contexts

5.3.1. English language context

The sLORETA analysis comparing bilinguals and monolinguals in the English language context revealed that bilinguals exhibited stronger cortical activation. Specifically, bilinguals showed significantly greater activation in the MFG between 200 and 260 ms after

stimulus onset. This suggests that bilinguals recruit more neural resources in frontal regions when processing English speech sounds, reflecting the engagement of the ECN. Such enhanced activation in bilinguals, compared to monolinguals, has been reported in previous studies and is thought to reflect increased cognitive effort (Kovelman et al., 2008a; Kovelman et al., 2008b; Parker Jones et al., 2012; Palomar-García et al., 2015; Román et al., 2015; Wang et al., 2011).

This pattern of activation is particularly noteworthy, as it aligns with concepts from the SLM (Flege, 1995, 2003), which posits a shared phonetic space from which assimilation or dissimilation can occur. The greater activation in the frontal regions of bilinguals compared to monolinguals in the English language context may reflect the phenomenon of dissimilation, where a new phonetic category for an L2 sound is fully established to maintain phonetic contrast with L1. This separation, driven by the tendency of phonetic categories to minimize overlap (Lindblom, 1990), enhances the distinction between L2 sounds and their L1 counterparts. Crucially, the heightened ECN activation observed in bilinguals aligns with Flege's concept of overshoot, in which bilinguals may exaggerate or overrealize certain L2 phonetic contrasts as they work to differentiate them from L1 categories (Flege & Eefting, 1988; Mack, 1990). This increased neural engagement suggests that bilinguals recruit additional cognitive mechanisms to reinforce phonetic distinctions, further supporting the idea that dissimilation is not solely a perceptual or articulatory phenomenon but also engages executive control processes at a neural level.

Our findings align with Kovelman et al. (2008b), who reported greater activation in frontal regions, such as the DLPFC and inferior frontal cortex (IFC), in bilinguals compared to monolinguals. Their study, which examined highly proficient Spanish-English bilinguals, demonstrated that bilinguals recruit additional cognitive control mechanisms to navigate dual-language contexts. This "bilingual signature" supports our observed MFG activity, reinforcing the idea that bilinguals actively engage additional neural resources to process speech sounds with competing phonological representations. These findings also complement the SLM (Flege, 1995, 2003), particularly the notion of dissimilation, where distinct phonetic categories emerge to minimize overlap between languages.

5.3.2. Spanish language context

The comparison between bilinguals and monolinguals in the Spanish language context is particularly significant, as it highlights the brain regions involved in within-category perception for bilinguals and between-category perception for monolinguals. This comparison is crucial for determining whether (1) the ECN provides the necessary flexibility to sustain and regulate parallel activation and adjustments in processing pathways, or (2) the ECN is responsible for managing competition between overlapping phonological representations.

The results differed significantly from those observed in the English language context. In the Spanish context, sLORETA analyses revealed enhanced brain activity in bilinguals compared to monolinguals in several regions: SFG (BA 6), MFG (BA 8), Pre-central Gyrus (PreCG; BA 4), and Postcentral Gyrus (PCG; BA 3). Notably, while the activation sites differ from those observed in the English language context, the regions identified in the Spanish context are integral to the ECN. Specifically, SFG activation, a component of the ECN, has been documented in bilingual language control studies (Abutalebi et al., 2007; Abutalebi et al., 2011; Garbin et al., 2011; Liu et al., 2021; Marian et al., 2017; Marian et al., 2014;

Perani et al., 2003; Rodríguez-Pujadas et al., 2013; Shen et al., 2020; Sulpizio et al., 2020; Geng et al., 2023). Similarly, MFG involvement in bilingual language control and decision-making processes has been established (Garbin et al., 2011; Perani et al., 2003; Shen et al., 2020; Sulpizio et al., 2020). Moreover, Burzynska et al. (2012) found that cortical thickness in the executive network, particularly in left and right MFG (right BA 9 and 46; left BA 8 and 9), the right PreCG (BA 4 and 6), and the left and right PCG (BA 2), among other regions, was a significant predictor of executive function. This was measured using the computerized Wisconsin Card Sorting Test, developed by Heaton et al. (1993). De Sanctis et al. (2009), using source localization analysis, also identified PreCG (BA 4), MFG (BA 6), and SFG (BA 6), among other brain regions, as being associated with the preservation of high levels of executive functioning. Furthermore, PCG activation has been observed in visual word paradigms. For example, Righi et al. (2010) explored the neural underpinnings of phonological onset competition using an eye tracking paradigm combined with fMRI, finding enhanced brain activation in typical executive control areas, including the left PCG (BA 3, 40, and 22), during the target versus competitor condition.

We propose that the differing patterns of brain activation observed between groups in the Spanish language context are due to the ECN engaging distinct processes in response to varying perceptual demands. Specifically, when perceptual elements need to be enhanced for more contrastive perception or diminished for less contrastive perception, the ECN may employ different processing strategies. These findings suggest that bilinguals are capable of dynamically adjusting processing pathways based on the linguistic context. And hence the results favor the idea that the ECN provides the necessary flexibility to maintain and modulate parallel activation and adjustments in processing pathways. If the ECN's function were limited solely to managing phonological representations, we would expect bilinguals and monolinguals to exhibit similar brain activation patterns in the Spanish language context, since their processing would parallel that of monolinguals who do not experience competition between overlapping phonological systems. In other words, if bilinguals did not need to manage multiple language systems simultaneously or adjust their brain's processing to accommodate different languages, there would be less demand for flexibility or additional engagement of the ECN. Therefore, if bilinguals and monolinguals exhibited similar patterns of brain activity, it would indicate that bilinguals were not experiencing the added cognitive challenge of handling sounds from both languages concurrently.

Evidence indicates that distinct brain regions within the ECN are selectively activated by various tasks, such as updating, inhibition, switching, and dual-tasking (Saylik et al., 2022). Similarly, Geng et al. (2023) demonstrate that bilingual individuals engage overlapping yet functionally distinct neural populations across their native and second languages. Geng and colleagues' findings suggest that different but overlapping neural patterns are recruited in response to specific task demands and the language being processed. This differential engagement of brain networks, which varies according to linguistic or perceptual contrasts, aligns with our findings that bilinguals can adjust processing pathways based on the linguistic context. Overall, our results support the concept of dynamic neural engagement and adaptability driven by linguistic context and task-specific demands.

Overall, we propose that differences in brain activation patterns across language contexts arise from the ECN engaging distinct processes in response to varying perceptual demands. Specifically,

when perceptual elements require enhancement for increased contrast or reduction for less contrastive perception, the ECN appears to deploy different strategies. These findings show that bilinguals can dynamically adapt to linguistic context, indicating that the ECN provides the flexibility needed to maintain and modulate parallel activation. Moreover, they extend and support existing models, such as the SLM and the L2LP model, by emphasizing the role of cognitive control in distinguishing language-specific phonemic categories.

5.4. Bilinguals double phonemic representation

As established in the introduction, the concept of bilinguals' double phonemic boundary posits that bilingual individuals maintain dual phonemic representations for the same speech sounds. Although much of the supporting evidence comes from behavioral measures, it has not definitively identified which brain regions underpin the auditory discrimination of two phonemic categories with competing representations. Our findings show that ECN activation persists across both language contexts, indicating that both languages remain active even when only one is ostensibly in use, in line with prior research on bilingual speech processing (Abutalebi et al., 2007; Abutalebi et al., 2011; Hernandez et al., 2001; Ju & Luce, 2004; Marian & Spivey, 2003; Spivey & Marian, 1999). Rather than reflecting a static "double representation," these results point to a context-driven, dynamically recalibrated process at the level of phonemic perception, orchestrated by the ECN.

5.5. Language proficiency and its impact on results

It is essential to consider the influence of language proficiency and usage on the observed results. Although these variables inevitably affected the findings, it is crucial to determine how these effects are most likely reflected in the results. For example, the reduced MMN in the Spanish language context, compared to the English context, can be attributed to decreased language control in a less proficient or less frequently used language (Green & Abutalebi, 2013). However, our findings do not necessarily support this interpretation for two reasons. First, the present study replicated previous findings, showing the expected amplitude pattern in MMN responses, with larger MMN in the English context indicating greater phonemic contrast and reduced MMN in the Spanish context suggesting less contrastive perception of the sounds (García-Sierra et al., 2012). This consistency implies that the bilingual participants were proficient enough in Spanish to adjust their perception of VOT across contexts. Second, the comparison with monolinguals in the Spanish context highlights the degree of perceptual flexibility demonstrated by bilinguals. In the following section, we discuss the specific brain regions associated with these perceptual adjustments and their correlation with language use at the time of the experiment, providing deeper insights into the neural mechanisms involved.

5.6. Language use shifts across the lifespan: patterns and influences

Bilingual language activation is complex. For instance, research on bilingual language production indicates that when bilinguals must either strictly adhere to or dynamically switch between both languages, cognitive effort increases (Green, 1998; Hernandez et al., 2001; Abutalebi et al., 2011). However, when bilinguals select their language without strict adherence or dynamic switching, cognitive effort does not significantly increase (Blanco-Elorrieta &

Pylkkänen, 2017; Kleinman & Gollan, 2016; Zhang et al., 2015; Zhu et al., 2022). While this indicates that a control mechanism may not always be essential for language selection (Costa & Santesteban, 2004; Costa et al., 2006), there are real-life scenarios where bilinguals must maintain the use of only one language. In these scenarios, it is proposed that highly proficient bilinguals exhibit lower levels of inhibitory control than less proficient bilinguals (Green, 1998; Green & Abutalebi, 2013). Still, other researchers propose that once bilinguals learn to activate their language in a language-specific manner, they can utilize it in language-switching tasks, regardless of the proficiency levels of the languages involved (Costa et al., 2006).

The variations often seen in research on bilingualism may partly stem from the diverse methods used to assess bilingual proficiency and the interplay between the initially learned language (Birdsong, 1999; Johnson & Newport, 1989) and the frequency of language use (Dufour & Kroll, 1995; Schreuder & Weltens, 1993). Numerous studies have demonstrated the involvement of the IFG in lexico-semantic processing and lexico-control as a result of increased attentional and verbal working memory demands for dual language processing and cross-linguistic integration of semantic information (Gabrieli et al., 1998; Kovelman et al., 2008a; Kovelman et al., 2008b; Petrides, 2005). Relevantly, IFG often reveals greater activation for L2 compared to L1. However, in the case of early bilinguals, many studies have reported increased activation for L1 compared to L2 (see for a comprehensive review, Sulpizio et al., 2020). Language proficiency, age of acquisition, frequency of use, and the specifics of the language task may all have an impact on this variation in IFG activation, which reflects the complexity of language processing in bilingual individuals.

Our study's methodology was designed to simulate a real-life scenario: watching a movie in either Spanish or English. Specifically, bilingual participants adhered strictly to one language by passively attending to a movie without generating responses related to experimental tasks. Regarding language background, participants exhibited the well-documented dominance shift in heritage bilinguals (Kohnert et al., 1999; Valdés, 2005). In this context, bilinguals are primarily exposed to Spanish during early childhood, with English becoming more prominent later in both academic and non-academic contexts. The frequency of English use in daily activities was confirmed through digital recorders worn by participants for two days. Concerning language proficiency, their age-adjusted PPVT scores were within the normal range for both languages, indicating potential ceiling effects.

Given the minimal variability in language proficiency between Spanish and English and the predominant use of English at the time of the experiment, we chose to explore the relationship between the well-documented language dominance shift in heritage bilinguals and brain activation during speech discrimination. Our findings revealed that bilinguals who experienced a notable shift from predominantly using Spanish in early childhood to primarily using English in adulthood exhibited stronger IFG activation in the English language context compared to the Spanish context. This heightened activation likely reflects increased engagement of executive control mechanisms within the IFG to manage contrastive phonemic distinctions and optimize speech processing in the dominant language through adjustments in processing pathways.

Overall, the correlational results highlight the complexities of language proficiency and brain activation in heritage bilinguals. By examining bilinguals who transitioned from using Spanish (L1) to English (L2) as their dominant language, we found that despite high proficiency in both languages, there was stronger activation in the

IFG during English tasks. This aligns with existing research on the role of IFG in lexico-semantic processing, particularly in bilingual contexts, where language use and proficiency, along with the age of acquisition, influence brain activation patterns (Sulpizio et al., 2020). Our results contribute to the understanding of the dynamic nature of bilingual language processing and the neurological underpinnings of shifts in language dominance. Therefore, our study not only reinforces the current understanding of the role of IFG in bilingual lexico-semantic processing but also enhances knowledge about the neurocognitive processes involved in language dominance shifts.

6. Conclusion

This study provides novel insights into how bilinguals perceive and process speech sounds that have competing phonemic representations across languages. By using ERPs and sLORETA to measure and localize brain activity, we found that bilinguals exhibit greater activation in regions associated with the ECN when processing these sounds, especially in different language contexts. Specifically, increased activation in the left frontal cortex during the English context suggests that the ECN plays a crucial role in adjusting processing pathways to accommodate language-specific phonemic contrasts.

Our findings extend existing models like the SLM and the L2LP model by emphasizing the importance of cognitive control mechanisms in differentiating language-specific phonemic categories. The ability of bilinguals to dynamically adjust their processing pathways based on linguistic context underscores the flexibility and adaptability of the ECN in managing parallel activation across languages.

Additionally, the observed shifts in language dominance among heritage bilinguals highlight the complex interplay between language proficiency, usage, and neural activation patterns. The stronger activation in the IFG during English tasks reflects the neurocognitive adjustments associated with changes in language dominance over time.

Overall, our study enhances the understanding of the neural mechanisms underlying bilingual speech perception. It emphasizes the pivotal role of the ECN in enabling bilinguals to navigate between languages efficiently, thereby contributing to the broader knowledge of bilingual language processing and cognitive control.

Supplementary material. The supplementary material for this article can be found at <http://doi.org/10.1017/S1366728925000148>.

Data availability statement. The data supporting this study is available upon request by contacting the main author.

Acknowledgments. We wish to thank for data collection assistance to Noelle Wig, Eilis Welsh, Calli Smith, Sarah Polcaro, Lexi Arcomano, Molly Barnett, Sydney Bates, Christine Cammisa, Kaleigh Constantine, Leah Cutkomp, Tayla Duntz, Kristen Fagan, Crystal Flores, Lina Kane, Ashley Lombardi, Alondra Marmolejos, Amy O'Rourke, and Allison Tozzi.

Competing interest. The authors declare no competing interests exist.

References

- Abramson, A. S., & Lisker, L. (1967). Discriminability along the voicing continuum: Cross-language tests. In *Proceedings of the sixth international congress of phonetic sciences* (pp. 1–10). Prague.
- Abutalebi, J., Annoni, J.-M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., Lazeyras, F., Cappa, S. F., & Khateb, A. (2007). Language control and lexical competition in bilinguals: An event-related fMRI study. *Cerebral Cortex*, *18*(7), 1496–1505. <https://doi.org/10.1093/cercor/bhm182>

- Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., Cappa, S. F., & Costa, A. (2011). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex*, *22*(9), 2076–2086. <https://doi.org/10.1093/cercor/bhr287>
- Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear & Hearing*, *16*(1), 38–51. <https://doi.org/10.1097/00003446-199502000-00004>
- Alho, K., Winkler, I., Escera, C., Huottilainen, M., Virtanen, J., Jääskeläinen, I. P., Pekkonen, E., & Ilmoniemi, R. J. (1998). Processing of novel sounds and frequency changes in the human auditory cortex: Magnetoencephalographic recordings. *Psychophysiology*, *35*(2), 211–224. <https://doi.org/10.1111/1469-8986.3520211>
- Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2010). Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2. *Journal of Phonetics*, *38*(4), 640–653. <https://doi.org/10.1016/j.wocn.2010.09.004>
- Antoniou, M., Tyler, M. D., & Best, C. T. (2012). Two ways to listen: Do L2-dominant bilinguals perceive stop voicing according to language mode? *Journal of Phonetics*, *40*(4), 582–594. <https://doi.org/10.1016/j.wocn.2012.05.005>
- Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*, *47*(6), 907–918. <https://doi.org/10.1016/j.neuron.2005.07.023>
- Birdsong, D. (1999). Second language acquisition and the critical period hypothesis. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 1–22). Lawrence Erlbaum Associates Publishers. <https://doi.org/10.4324/9781410601667>
- Blanco-Elorrieta, E., & Caramazza, A. (2021). A common selection mechanism at each linguistic level in bilingual and monolingual language production. *Cognition*, *213*, 104625. <https://doi.org/10.1016/j.cognition.2021.104625>
- Blanco-Elorrieta, E., & Pyllkkänen, L. (2017). Bilingual language switching in the laboratory versus in the wild: The spatiotemporal dynamics of adaptive language control. *The Journal of Neuroscience*, *37*, 9022–9036. <https://doi.org/10.1523/JNEUROSCI.0553-17.2017>
- Boersma, P. (1998). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. Universiteit van Amsterdam.
- Bohn, O.-S., & Flege, J. E. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83). Cambridge University Press. <https://doi.org/10.1017/9781108886901.002>
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. <https://doi.org/10.1037/0033-295X.108.3.624>
- Brian, T. G., David, A. B., Sara, J. J., David, K. P., Charles, D. S., & Anders, H. A. (2006). Dissociation of automatic and strategic lexical-semantics: Functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *The Journal of Neuroscience*, *26*(24), 6523. <https://doi.org/10.1523/JNEUROSCI.0808-06.2006>
- Bullmore, E. T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., & Brammer, M. J. (1999). Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Transactions on Medical Imaging*, *18*(1), 32–42. <https://doi.org/10.1109/42.750253>
- Burzynska, A. Z., Nagel, I. E., Preuschhof, C., Gluth, S., Bäckman, L., Li, S. C., Lindenberger, U., & Heekeren, H. R. (2012). Cortical thickness is linked to executive functioning in adulthood and aging. *Human Brain Mapping*, *33*(7), 1607–1620. <https://doi.org/10.1002/hbm.21311>
- Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *The Journal of the Acoustical Society of America*, *54*(2), 421–428. <https://doi.org/10.1121/1.1913594>
- Casillas, J. V., & Simonet, M. (2018). Perceptual categorization and bilingual language modes: Assessing the double phonemic boundary in early and late bilinguals. *Journal of Phonetics*, *71*, 51–64. <https://doi.org/10.1016/j.wocn.2018.07.002>
- Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491–511. <https://doi.org/10.1016/j.jml.2004.02.002>
- Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(5), 1057–1074. <https://doi.org/10.1037/0278-7393.32.5.1057>
- Davidian, R. D., & Flege, J. E. (1984). Transfer and developmental processes in adult foreign language speech production. *Applied Psycholinguistics*, *5*(4), 323–347. <https://doi.org/10.1017/S014271640000521X>
- De Sanctis, P., Gomez-Ramirez, M., Sehatpour, P., Wylie, G. R., & Foxe, J. J. (2009). Preserved executive function in high-performing elderly is driven by large-scale recruitment of prefrontal cortical mechanisms. *Human Brain Mapping*, *30*(12), 4198–4214. <https://doi.org/10.1002/hbm.20839>
- Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(03), 175–197. <https://doi.org/10.1017/S1366728902003012>
- Dijkstra, T., Van Jaarsveld, H., & Brinke, S. T. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, *1*(1), 51–66. <https://doi.org/10.1017/S1366728998000121>
- Doeller, C. F., Opitz, B., Mecklinger, A., Krick, C., Reith, W., & Schröger, E. (2003). Prefrontal cortex involvement in preattentive auditory deviance detection: Neuroimaging and electrophysiological evidence. *Neuroimage*, *20*(2), 1270–1282. [https://doi.org/10.1016/S1053-8119\(03\)00389-6](https://doi.org/10.1016/S1053-8119(03)00389-6)
- Dufour, R., & Kroll, J. F. (1995). Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory & Cognition*, *23*(2), 166–180. <https://doi.org/10.3758/BF03197219>
- Dunn, L. M. (1997). *PPVT-III: Peabody picture vocabulary test* (3rd ed.). American Guidance Service.
- Elman, J. L., Diehl, R. L., & Buchwald, S. E. (1977). Perceptual switching in bilinguals. *Journal of the Acoustical Society of America*, *62*(4), 971–974. <https://doi.org/10.1121/1.381570>
- Ernst, M. D. (2004). Permutation methods: A basis for exact inference. *Statistical Science*, *19*(4), 676–685. <https://doi.org/10.1214/088342304000000396>
- Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. Utrecht University.
- Escudero, P. (2009). The linguistic perception of similar L2 sounds. In B. Paul & H. Silke (Eds.), *Phonology in perception* (pp. 151–190). De Gruyter Mouton. <https://doi.org/10.1515/9783110219234.151>
- Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*(4), 551–585. <https://doi.org/10.1017/S0272263104040021>
- Escudero, P., Benders, T., & Lipski, S. C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics*, *37*(4), 452–465. <https://doi.org/10.1016/j.wocn.2009.07.006>
- Escudero, P., & Yazawa, K. (2024). The second language linguistic perception model. In M. Amengual (Ed.), *The Cambridge handbook of bilingual phonetics and phonology* (pp. 173–195). Cambridge University Press. <https://doi.org/10.1017/9781009105767.009>
- Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*, *89*(1), 395–411. <https://doi.org/10.1121/1.400473>
- Flege, J. E. (1992). The intelligibility of English vowels spoken by British and Dutch talkers. In *Intelligibility in speech disorders* (pp. 1–22). John Benjamins. <https://doi.org/10.1075/sspl.1.06fl>
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–273). York.
- Flege, J. E., & Eefting, W. (1987a). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, *15*(1), 67–83. [https://doi.org/10.1016/S0095-4470\(19\)30538-8](https://doi.org/10.1016/S0095-4470(19)30538-8)
- Flege, J. E., & Eefting, W. (1987b). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, *6*(3), 185–202. [https://doi.org/10.1016/0167-6393\(87\)90025-2](https://doi.org/10.1016/0167-6393(87)90025-2)
- Flege, J. E., & Eefting, W. (1988). Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation.

- Journal of the Acoustical Society of America*, **83**(2), 729–740. <https://doi.org/10.1121/1.396115>
- Flége, J. E., Munro, M. J., & Skelton, L. (1992). Production of the word-final English /t/–/d/ contrast by native speakers of English, Mandarin, and Spanish. *Journal of the Acoustical Society of America*, **92**(1), 128–143. <https://doi.org/10.1121/1.404278>
- Flége, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, **40**(4), 467–491. [https://doi.org/10.1016/S0167-6393\(02\)00128-0](https://doi.org/10.1016/S0167-6393(02)00128-0)
- Gabrieli, J. D., Poldrack, R. A., & Desmond, J. E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the National Academy of Sciences of the United States of America*, **95**(3), 906–913. <https://doi.org/10.1073/pnas.95.3.906>
- Garbin, G., Costa, A., Sanjuan, A., Forn, C., Rodriguez-Pujadas, A., Ventura, N., Belloch, V., Hernandez, M., & Ávila, C. (2011). Neural bases of language switching in high and early proficient bilinguals. *Brain and Language*, **119**(3), 129–135. <https://doi.org/10.1016/j.bandl.2011.03.011>
- García-Sierra, A., Diehl, R. L., & Champlin, C. (2009). Testing the double phonemic boundary in bilinguals. *Speech Communication*, **51**(4), 369–378. <https://doi.org/10.1016/j.specom.2008.11.005>
- García-Sierra, A., Ramírez-Esparza, N., Silva-Pereyra, J., Siard, J., & Champlin, C. A. (2012). Assessing the double phonemic representation in bilingual speakers of Spanish and English: An electrophysiological study. *Brain and Language*, **121**(3), 194–205. <https://doi.org/10.1016/j.bandl.2012.03.008>
- García-Sierra, A., Schifano, E., Duncan, G. M., & Fish, M. S. (2021). An analysis of the perception of stop consonants in bilinguals and monolinguals in different phonetic contexts: A range-based language cueing approach. *Attention, Perception, & Psychophysics*, **83**(4), 1878–1896. <https://doi.org/10.3758/s13414-020-02183-z>
- Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology*, **120**(3), 453–463. <https://doi.org/10.1016/j.clinph.2008.11.029>
- Geng, S., Guo, W., Rolls, E. T., Xu, K., Jia, T., Zhou, W., Blakemore, C., Tan, L.-H., Cao, M., & Feng, J. (2023). Intersecting distributed networks support convergent linguistic functioning across different languages in bilinguals. *Communications Biology*, **6**(1), 99. <https://doi.org/10.1038/s42003-023-04446-5>
- Gershon, R. C., Cook, K. F., Mungas, D., Manly, J. J., Slotkin, J., Beaumont, J. L., & Weintraub, S. (2014). Language measures of the NIH Toolbox Cognition Battery. *Journal of the International Neuropsychological Society*, **20**(6), 642–651. <https://doi.org/10.1017/S1355617714000411>
- Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance: A topographic event-related potential study. *Psychophysiology*, **27**(6), 627–640. <https://doi.org/10.1111/j.1469-8986.1990.tb03184.x>
- Gold, B. T., & Buckner, R. L. (2002). Common prefrontal regions coactivate with dissociable posterior regions during controlled semantic and phonological tasks. *Neuron*, **35**(4), 803–812. [https://doi.org/10.1016/S0896-6273\(02\)00800-0](https://doi.org/10.1016/S0896-6273(02)00800-0)
- Gomes, H., Molholm, S., Ritter, W., Kurtzberg, D., Cowan, N., & Vaughan, H. G. (2000). Mismatch negativity in children and adults, and effects of an attended task. *Psychophysiology*, **37**(6), 807–816. <https://doi.org/10.1111/1469-8986.3760807>
- Gonzales, K., & Lotto, A. J. (2013). A Bafri, un Pafri: bilinguals' Pseudoword identifications support language-specific phonetic systems. *Psychological Science*, **24**(11), 2135–2142. <https://doi.org/10.1177/0956797613486485>
- Gonzales, K., Byers-Heinlein, K., & Lotto, A. J. (2019). How bilinguals perceive speech depends on which language they think they're hearing. *Cognition*, **182**, 318–330. <https://doi.org/10.1016/j.cognition.2018.08.021>
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, **1**(2), 67–81. <https://doi.org/10.1017/S1366728998000133>
- Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, **25**(5), 515–530. <https://doi.org/10.1080/20445911.2013.796377>
- Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, **1**(02), 131–149. <https://doi.org/10.1017/S136672899800025X>
- Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1–22). Blackwell.
- Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press. <https://doi.org/10.1017/S0022226709990089>
- Hazan, V. L., & Boulakia, G. (1993). Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech*, **36**, 17–38. <https://doi.org/10.1177/002383099303600102>
- Heaton, R. K., Chelune, C., Talley, J., Kay, G. G., & Curtiss, G. (1993). *Wisconsin card sorting test manual – Revised and Expanded*.
- Hernandez, A. E., Dapretto, M., Mazziotta, J., & Bookheimer, S. (2001). Language switching and language representation in Spanish-English bilinguals: An fMRI study. *Neuroimage*, **14**(2), 510–520. <https://doi.org/10.1006/nimg.2001.0810>
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, **8**(5), 393–402. <https://doi.org/10.1038/nrn2113>
- Holt, L. L., & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *Journal of the Acoustical Society of America*, **119**(5), 3059–3071. <Go to ISI>://000237459500049
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, **21**(1), 60–99. <https://www.jstor.org/stable/1422957>
- Ju, M., & Luce, P. A. (2004). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science*, **15**(5), 314–318. <https://doi.org/10.1111/j.0956-7976.2004.00675.x>
- Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, **67**(3), 971–990. <https://doi.org/10.1121/1.383940>
- Kleinman, D., & Gollan, T. H. (2016). Speaking two languages for the price of one: Bypassing language control mechanisms via accessibility-driven switches. *Psychological Science*, **27**(5), 700–714. <https://doi.org/10.1177/09567976166634633>
- Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, **302**(5648), 1181–1185. <https://doi.org/10.1126/science.1088545>
- Kohnert, K., Bates, E., & Hernandez, A. E. (1999). Balancing bilinguals: Lexical-semantic production and cognitive processing in children learning Spanish & English. *Journal of Speech, Language, and Hearing Research*, **42**(6), 1400–1413. <https://doi.org/10.1044/jslhr.4206.1400>
- Kovelman, I., Baker, S. A., & Petitto, L. A. (2008a). Bilingual and monolingual brains compared: A functional magnetic resonance imaging investigation of syntactic processing and a possible “neural signature” of bilingualism. *Journal of Cognitive Neuroscience*, **20**(1), 153–169. <https://doi.org/10.1162/jocn.2008.20011>
- Kovelman, I., Shalinsky, M. H., Berens, M. S., & Petitto, L.-A. (2008b). Shining new light on the brain's “bilingual signature”: A functional near infrared spectroscopy investigation of semantic processing. *Neuroimage*, **39**(3), 1457–1471. <https://doi.org/10.1016/j.neuroimage.2007.10.017>
- Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica*, **128**(3), 416–430. <https://doi.org/10.1016/j.actpsy.2008.02.001>
- Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., Kochunov, P. V., Nickerson, D., Mikiten, S. A., & Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*, **10**(3), 120–131. [https://doi.org/10.1002/1097-0193\(200007\)10:3<120::AID-HBM30>3.0.CO;2-8](https://doi.org/10.1002/1097-0193(200007)10:3<120::AID-HBM30>3.0.CO;2-8)
- Levänen, S., Ahonen, A., Hari, R., McEvoy, L., & Sams, M. (1996). Deviant auditory stimuli activate human left and right auditory cortex differently. *Cerebral Cortex*, **6**(2), 288–296. <https://doi.org/10.1093/cercor/6.2.288>
- Li, F., Menon, A., & Allen, J. B. (2010). A psychoacoustic method to find the perceptual cues of stop consonants in natural speech. *Journal of Acoustical Society of America*, **127**(4), 2599–2610. <https://doi.org/10.1121/1.3295689>
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Springer Netherlands. https://doi.org/10.1007/978-94-009-2037-8_16

- Lisker, L., & Abramson, A. S. (1964). A cross-linguistic study of voicing in initial stops: Acoustical measurements. *Word*, *20*, 384–422. <https://doi.org/10.1080/00437956.1964.11659830>
- Liu, C., Jiao, L., Li, Z., Timmer, K., & Wang, R. (2021). Language control network adapts to second language learning: A longitudinal rs-fMRI study. *Neuropsychologia*, *150*, 107688. <https://doi.org/10.1016/j.neuropsychologia.2020.107688>
- Lozano-Argüelles, C., Fernández Arroyo, L., Rodríguez, N., Durand López, E. M., Garrido Pozú, J. J., Markovits, J., Varela, J. P., de Rocafiguera, N., & Casillas, J. V. (2021). Conceptually cued perceptual categorization in adult L2 learners. *Studies in Second Language Acquisition*, *43*(1), 204–219. <https://doi.org/10.1017/S0272263120000273>
- Luck, J. S. (2014). *An introduction to the event-related potential technique* (2nd ed.). The MIT Press.
- Mack, M. (1990). Phonetic transfer in a French-English bilingual child. In P. H. Nelde (Ed.), *Language attitudes and language conflict*. Dümmler.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86. [https://doi.org/10.1016/0010-0285\(86\)90015-0](https://doi.org/10.1016/0010-0285(86)90015-0)
- Maess, B., Jacobsen, T., Schröger, E., & Friederici, A. D. (2007). Localizing pre-attentive auditory memory-based comparison: Magnetic mismatch negativity to pitch change. *Neuroimage*, *37*(2), 561–571. <https://doi.org/10.1016/j.neuroimage.2007.05.040>
- Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, *6*(2), 97–115. <https://doi.org/10.1017/S1366728903001068>
- Marian, V., Bartolotti, J., Rochanavibhata, S., Bradley, K., & Hernandez, A. E. (2017). Bilingual cortical control of between- and within-language competition. *Scientific Reports*, *7*(1), 11763. <https://doi.org/10.1038/s41598-017-12116-w>
- Marian, V., Chabal, S., Bartolotti, J., Bradley, K., & Hernandez, A. E. (2014). Differential recruitment of executive control regions during phonological competition in monolinguals and bilinguals. *Brain and Language*, *139*, 108–117. <https://doi.org/10.1016/j.bandl.2014.10.005>
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., ... Mazoyer, B. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *356*(1412), 1293–1322. <https://doi.org/10.1098/rstb.2001.0915>
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. <https://doi.org/10.1146/annurev.neuro.24.1.167>
- Näätänen, R. (1992). *Attention and brain function*. Lawrence Erlbaum Associates, Publishers.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huottilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*, 432–434. <https://doi.org/10.1038/385432a0>
- Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, *15*(1), 1–25. <https://doi.org/10.1002/hbm.1058>
- Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(2), 241–268. <https://doi.org/10.3758/s13415-011-0083-5>
- Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, *23*(3), 299–325. <https://doi.org/10.1017/S0140525X00003241>
- Obleser, J., Scott, S. K., & Eulitz, C. (2005). Now you hear it, now you don't: Transient traces of consonants and their nonspeech analogues in the human brain. *Cerebral Cortex*, *16*(8), 1069–1076. <https://doi.org/10.1093/cercor/bhj047>
- Opitz, B., Mecklinger, A., Friederici, A. D., & von Cramon, D. Y. (1999). The functional neuroanatomy of novelty processing: Integrating ERP and fMRI results. *Cerebral Cortex*, *9*(4), 379–391. <https://doi.org/10.1093/cercor/9.4.379>
- Palomar-García, M., Bueichekú, E., Ávila, C., Sanjuán, A., Strijkers, K., Ventura-Campos, N., & Costa, A. (2015). Do bilinguals show neural differences with monolinguals when processing their native language?. *Brain Language*, *142*, 36–44. <https://doi.org/10.1016/j.bandl.2015.01.004>
- Parker Jones, O., Green, D. W., Grogan, A., Pliatsikas, C., Filippopolitis, K., Ali, N., Lee, H. L., Ramsden, S., Gazarian, K., Prejawa, S., Seghier, M. L., & Price, C. J. (2012). Where, when and why brain activation differs for bilinguals and monolinguals during picture naming and reading aloud. *Cerebral Cortex*, *22*(4), 892–902. <https://doi.org/10.1093/cercor/bhr161>
- Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details. *Methods Findings in Experimental and Clinical Pharmacology*, *24*(Suppl D), 5–12.
- Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1994). Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain. *International Journal of Psychophysiology*, *18*(1), 49–65. [https://doi.org/10.1016/0167-8760\(84\)90014-X](https://doi.org/10.1016/0167-8760(84)90014-X)
- Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., & Fazio, F. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human Brain Mapping*, *19*(3), 170–182. <https://doi.org/10.1002/hbm.10110>
- Petrides, M. (2005). Lateral prefrontal cortex: Architectonic and functional organization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 781–795. <https://doi.org/10.1098/rstb.2005.1631>
- Poeppel, D., Idsardi, W. J., & van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 1071–1086. <https://doi.org/10.1098/rstb.2007.2160>
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *NeuroImage*, *10*(1), 15–35. <https://doi.org/10.1006/nimg.1999.0441>
- Ramírez-Esparza, N., Jiang, S., García-Sierra, A., Skoe, E., & Benítez-Barrera, C. R. (2024). Effects of cultural dynamics on everyday acoustic environments. *The Journal of the Acoustical Society of America*, *156*(3), 1942–1951. <https://doi.org/10.1121/10.0028814>
- Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*(5695), 443–447. <https://doi.org/10.1126/science.1100301>
- Righi, G., Blumstein, S. E., Mertus, J., & Worden, M. S. (2010). Neural systems underlying lexical competition: An eye tracking and fMRI study. *Journal of Cognitive Neuroscience*, *22*(2), 213–224. <https://doi.org/10.1162/jocn.2009.21200>
- Rinne, T., Alho, K., Ilmoniemi, R. J., Virtanen, J., & Näätänen, R. (2000). Separate time behaviors of the temporal and frontal mismatch negativity sources. *NeuroImage*, *12*(1), 14–19. <https://doi.org/10.1006/nimg.2000.0591>
- Rodríguez-Fornells, A., Rotte, M., Heinze, H.-J., Nössel, T., & Münte, T. F. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature*, *415*(6875), 1026–1029. <https://doi.org/10.1038/4151026a>
- Rodríguez-Pujadas, A., Sanjuán, A., Ventura-Campos, N., Román, P., Martín, C., Barceló, F., Costa, A., & Ávila, C. (2013). Bilinguals use language-control brain areas more than monolinguals to perform non-linguistic switching tasks. *PLOS ONE*, *8*(9), e73028. <https://doi.org/10.1371/journal.pone.0073028>
- Roland, P. E. (1981). Somatotopical tuning of postcentral gyrus during focal attention in man: A regional cerebral blood flow study. *Journal of Neurophysiology*, *46*(4), 744–754. <https://doi.org/10.1152/jn.1981.46.4.744>
- Roland, P. E. (1982). Cortical regulation of selective attention in man: A regional cerebral blood flow study. *Journal of Neurophysiology*, *48*(5), 1059–1078. <https://doi.org/10.1152/jn.1982.48.5.1059>
- Román, P., González, J., Ventura-Campos, N., Rodríguez-Pujadas, A., Sanjuán, A., & Ávila, C. (2015). Neural differences between monolinguals and early bilinguals in their native language during comprehension. *Brain Language*, *150*, 80–89. <https://doi.org/10.1016/j.bandl.2015.07.011>

- Saylik, R., Williams, A. L., Murphy, R. A., & Szameitat, A. J. (2022). Characterising the unity and diversity of executive functions in a within-subject fMRI study. *Scientific Reports*, *12*(1), 8182. <https://doi.org/10.1038/s41598-022-11433-z>
- Scherg, M., Vajsar, J., & Picton, T. W. (1989). A source analysis of the late human auditory evoked potentials. *Journal of Cognitive Neuroscience*, *1*(4), 336–355. <https://doi.org/10.1162/jocn.1989.1.4.336>
- Schreuder, R., & Weltens, B. (1993). *The bilingual lexicon*. John Benjamins Publishing Company. <https://doi.org/10.1075/sibil.6>
- Shen, K.-k., Welton, T., Lyon, M., McCorkindale, A. N., Sutherland, G. T., Burnham, S., Fripp, J., Martins, R., & Grieve, S. M. (2020). Structural core of the executive control network: A high angular resolution diffusion MRI study. *Human Brain Mapping*, *41*(5), 1226–1236. <https://doi.org/10.1002/hbm.24870>
- Shtyrov, Y., & Pulvermüller, F. (2007). Language in the mismatch negativity design: Motivations, benefits, and prospects. *Journal of Psychophysiology*, *21*(3), 176–187. <https://doi.org/10.1027/0269-8803.21.3.176>
- Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*(5408), 1657–1661. <https://doi.org/10.1126/science.283.5408.1657>
- Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, *10*(3), 281–284. <https://doi.org/10.1111/1467-9280.00151>
- Sulpizio, S., Del Maschio, N., Fedeli, D., & Abutalebi, J. (2020). Bilingual language processing: A meta-analysis of functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, *108*, 834–853. <https://doi.org/10.1016/j.neubiorev.2019.12.014>
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-Dimensional proportional system: An approach to cerebral imaging*. Thieme Medical Publishers Inc.
- Tervaniemi, M., Medvedev, S. V., Alho, K., Pakhomov, S. V., Roudas, M. S., Van Zuijen, T. L., & Näätänen, R. (2000). Lateralized automatic auditory processing of phonetic versus musical information: A PET study. *Human Brain Mapping*, *10*(2), 74–79. [https://doi.org/10.1002/\(SICI\)1097-0193\(200006\)10:2<74::AID-HBM30>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1097-0193(200006)10:2<74::AID-HBM30>3.0.CO;2-2)
- Tiitinen, H., May, P., Reinikainen, K., & Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature*, *372*(6501), 90–92. <https://doi.org/10.1038/372090a0>
- Valdés, G. (2005). Bilingualism, heritage language learners, and SLA research: Opportunities lost or seized? *The Modern Language Journal*, *89*(3), 410–426. <https://doi.org/10.1111/j.1540-4781.2005.00314.x>
- van Heuven, W. J. B., Schriefers, H., Dijkstra, T., & Hagoort, P. (2008). Language conflict in the bilingual brain. *Cerebral Cortex*, *18*(11), 2706–2716. <https://doi.org/10.1093/cercor/bhn030>
- Van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology*, *6*(1000). <https://doi.org/10.3389/fpsyg.2015.01000>
- Vitacco, D., Brandeis, D., Pascual-Marqui, R., & Martin, E. (2002). Correspondence of event-related potential tomography and functional magnetic resonance imaging during language processing. *Human Brain Mapping*, *17*(1), 4–12. <https://doi.org/10.1002/hbm.10038>
- Wang, Y., Xiang, J., Vannest, J., Holroyd, T., Narmoneva, D., Horn, P., Liu, Y., Rose, D., deGrauw, T., & Holland, S. (2011). Neuromagnetic measures of word processing in bilinguals and monolinguals. *Clinical Neurophysiology*, *122*(9), 1706–1717. <https://doi.org/10.1016/j.clinph.2011.02.008>
- Wig, N., & García-Sierra, A. (2020). Matching the mismatch: The interaction between perceptual and conceptual cues in bilinguals' speech perception. *Bilingualism: Language and Cognition*, *24*(3), 467–480. <https://doi.org/10.1017/S1366728920000553>
- Williams, L. (1977). The perception of stop consonant voicing by Spanish-English bilinguals. *Perception & Psychophysics*, *21*(4), 289–297. <https://doi.org/10.3758/BF03199477>
- Yazawa, K., Whang, J., Kondo, M., & Escudero, P. (2020). Language-dependent cue weighting: An investigation of perception modes in L2 learning. *Second Language Research*, *36*(4), 557–581. <https://doi.org/10.1177/0267658319832645>
- Yoshida, K. A., Iversen, J. R., Patel, A. D., Mazuka, R., Nito, H., Gervain, J., & Werker, J. F. (2010). The development of perceptual grouping biases in infancy: A Japanese-English cross-linguistic study. *Cognition*, *115*(2), 356–361. <https://doi.org/10.1016/j.cognition.2010.01.005>
- Zhang, Y., Wang, T., Huang, P., Li, D., Qiu, J., Shen, T., & Xie, P. (2015). Free language selection in the bilingual brain: An event-related fMRI study. *Scientific Reports*, *5*(1), 11704.
- Zhu, J. D., Blanco-Elorrieta, E., Sun, Y., Szakay, A., & Sowman, P. F. (2022). Natural vs forced language switching: Free selection and consistent language use eliminate significant performance costs and cognitive demands in the brain. *NeuroImage*, *247*, 118797. <https://doi.org/10.1016/j.neuroimage.2021.118797>