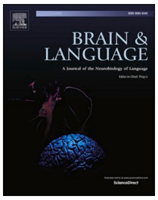
Brain & Language 264 (2025) 105553

Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l



Native prosodic structures constrain L2 word recognition: Evidence from Bengali-English bilinguals

Isabella Fritz *, Aditi Lahiri

Language and Brain Laboratory, Faculty of Linguistics, Philology and Phonetics, University of Oxford, UK

ARTICLE INFO

Keywords:
Word recognition
ERP
Phonology
Bilingualism
L2
Bengali
Metrical feet

ABSTRACT

Bilingual word recognition is assumed to be modulated by a word's segmental and meaning similarity across languages, labelled *cognate* in psycholinguistics, usually conflating borrowed and inherited words. We conducted an ERP fragment priming study with Bengali-English bilinguals. English words *borrowed* into Bengali (*doctor,* Bengali: [ˈɖaktar]) were compared with those which were not (*river*). The stimuli varied in fine metrical details, one-foot (*dóctor*) or two-feet (*éxpèrt)* whilst stress placement was kept constant. Crucially, two-feet English words are always one-foot in Bengali [ˈeksparʈ]. Behavioural results (RTs) showed that although loan status did not affect priming, mismatch in feet significantly reduced the effect. In the ERP data, only one-foot words elicited significant priming effects. Furthermore, different ERP components were modulated depending on loan type. Thus, loan status alone is not sufficient to understand L2 word processing; the influence of the native metrical structure (preference for a single foot) constrains processing of all words.

1. Introduction

A central question in the study of bilingualism concerns the architecture of the bilingual mental lexicon and its consequences for word recognition and processing. Much of what we know about the bilingual mental lexicon is based on experimental studies on *cognate* vs. *non-cognate* recognition and production (Sánchez-Casas & García-Albea, 2005). In the psycholinguistic literature, the term *cognate* refers to translation equivalents that overlap in meaning, form and, depending on the language combination, orthography (Dijkstra & van Heuven, 2002) and thus conflating the linguistic concepts of inherited words (i.e., cognates) and loanwords.

Studies have found that cognate (loan) status impacts processing. The general idea is that if two items are orthographically and/or phonologically similar across languages, lexical access is enhanced. This is evident, for example, in faster reaction times and fewer errors across different experimental paradigms (e.g., lexical decision task) investigating word recognition. This *cognate facilitation effect* is taken as evidence that bilingual lexical access is non-selective meaning that cross-language activation occurs even in tasks where only one language is required (Dijkstra & van Heuven, 2002; Dijkstra et al., 2019). However, several factors modulate this facilitation effect or may even reverse it leading to inhibition. Such factors include the degree of orthographic and phonological overlap and whether the two languages share a script (see Lijewska, 2020 for an overview). Moreover, experimental findings on cognate facilitation are mainly based on visual word processing. Even studies specifically looking at the relationship between phonology and word recognition, often present their stimuli through the visual modality (Comesaña et al., 2015; Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010).

To date, it is still unclear to what extent and on which phonological level the L1 phonology impacts L2 'cognate' processing. One reason may be the lack of a systematic approach to disentangle the type of similarity/difference of 'cognates' that impacts L2 word recognition. For example, there is no consensus on how to quantify phonological similarity across studies. Some researchers ask bilinguals to rate the similarity of a given word pair (e.g., Muntendam, van Rijswijk, Severijnen, & Dijkstra, 2022), others use algorithms to calculate similarity based on phonetic (e.g., ALINE algorithm Kondrak (2003) used in Frances, Navarra-Barindelli, & Martin, 2021) or phonemic differences (e.g., Levenshtein Distance used in Mulder, Wloch, Boves, ten Bosch, & Ernestus, 2021). Only more recent studies have investigated specific sub-phonemic aspects and its role in L2 processing including word stress placement (Mulder et al., 2021; Muntendam et al., 2022).

We argue that one way to disentangle how different phonological levels impact bilingual word recognition, is to design experiments that

---

* Corresponding author.
E-mail address: isabella.fritz@ling-phil.ox.ac.uk (I. Fritz).

https://doi.org/10.1016/j.bandl.2025.105553
Received 21 May 2024; Received in revised form 12 January 2025; Accepted 17 February 2025

are guided by phonological theory and historical developments. Relevant for the current study, is the distinction between cognates (inherited words) and loanwords. Usually, in psycholinguistic experiments that include ’cognates’ as stimuli, no distinction is made between cognates (inherited words) and loanwords which have been borrowed into both/all languages being investigated (see Shetty et al., 2022 for an exception). Within linguistics, however, a word is a cognate in L1 and L2 if they are inherited from the same source language: Bengali [gou]/[goru], English *cow* [kʰaʊ], German *Kuh* [kʰu]. In this example, the Indo-European reconstructed form would be *gʷṓws, and the current word in the three modern languages have undergone regular phonological rules to attain this form. However, Bengali [nɔmbor], English *number,* German *Nummer* are also originally from Indo-European but these are not cognates. That is, these words did not undergo regular processes of the individual languages. English borrowed the word from Old French, German borrowed it from Italian and Bengali borrowed it from English. Although present-day speakers do not have knowledge of the history of their own native language, grammars are pertinacious and past developments leave their mark on the synchronic system in systematic ways, which must be processed by the present-day speaker. In the present study, we focus on loans which conform to the native phonology at the time of borrowing. Thus, depending on their borrowing history, shared loans across languages are often phonologically more closely related than cognates which, especially for distantly related languages, may not be considered as such from a naive speaker’s perspective.

Moving on from loans to the general language processing literature, several studies have shown that syllables and stress placement play a crucial role in word recognition in many languages (e.g., Cutler & Jesse, 2021; Mehler, Dommergues, Frauenfelder, & Segui, 1981; Zwitserlood, Schriefers, Lahiri, & van Donselaar, 1993). Cutler & Carter (1987) argued in support of a *Metrical Segmentation Strategy* that is part of a listener’s processing competence and which claims that different languages employ different prosodic segmentation strategies (Cutler, Mehler, Norris, & Segui, 1983). Importantly, listeners then also apply this native processing strategy when processing an L2 (Cutler, Mehler, Norris, & Segui, 1989). Thus, besides cognate and loanword recognition, more general phonological aspects need to be considered to enhance our understanding of bilingual word processing.

When considering the role of syllables in speech segmentation, it is the syllable type (weak or strong) which matters for recognition. Whilst strong syllables contain full vowels that can potentially carry stress, weak syllables include a reduced vowel or schwa (Cutler & Norris, 1988). However, if we consider strong syllables, it does not automatically suggest that it bears main stress; cf. *bamboo* [ˌbæmˈbuː] vs. *reptile* [ˈrɛpˌtʌɪl] where we have a combination of two strong syllables. Thus, a stressed (strong) syllable alone does not determine the unit which governs the metrical pattern. To understand stress placement in English, we need to consider an additional level related to the organisation of stressed and unstressed syllables into a larger phonological unit, namely the metrical foot. For example, the syllable [mæn] bears main stress in both *mandate* and *mantle,* but the latter constitutes a single foot while the former has two feet. Thus, the stressed syllable [ˈmæn] is a foot by itself in *mandate* but is part of a foot in *mantle.*

One of the critical questions we are asking is whether the foot itself plays a role in lexical access. Foot structure has hardly been investigated in word recognition studies. Initial investigations using metrical violation paradigms suggest a *psychological reality* for the concept of metrical feet (Domahs, Klein, Huber, & Domahs, 2013; Magne et al., 2007; Marie, Magne, & Besson, 2011). Further, in a previous study we found that the metrical foot structure of someone’s first language (German) impacts word recognition in L2 (English). The current study is an elaboration of this study, testing the effect of the L1 metrical structure in more distantly related languages with cross-script bilinguals (i.e., Bengali-English). In Fritz et al. (2023) participants heard the first syllable cut out of a disyllabic word that either matched the following visually presented target or was phonologically unrelated. Event-related potentials (ERP) time-locked to the onset of the target word and presented in the participants’ L2 English were modulated differently depending on both the word’s loan status (i.e., shared vs. absent) and the word’s metrical structure. Half the stimuli we used were Romance loans that were borrowed into both English and German (shared) but varied in their metrical structure, the other half were only borrowed into English (absent) but had no phonologically related counter-part in German. They, however, either had the preferred metrical structure for German words (e.g., strong–weak syllables: pigeon [ˈpɪdʒɨn]) or a metrical structure that barely exists in German disyllabic words (strong-strong with initial stress: curfew [ˈkɜːˌfjuː]). In words with two strong syllables, German almost always places stress on the final syllable as in *Reptil* [rɛpˈtiːl] (reptile).We only found a significant N400 priming effect in loans which were not borrowed into German and there only for words with the preferred German metrical structure (i.e., strong–weak). Thus, participants processed words that consisted of the preferred metrical structure for German disyllabic words more easily, i.e., one foot with initial stress as in all inherited Germanic words (Tochter [ˈtɔxtɐ] “daughter”; Finger [ˈfɪŋɐ] “finger”). This indicates that the listener’s metrical knowledge of their L1 is transferred to L2 processing. Importantly, in this study we do not suggest that the foot structure of any given word is a stored unit as part of a lexical entry, i.e. as prelexical units in perception as suggested for syllables by Mehler and colleagues (1981). Rather, we hypothesise that speakers are sensitive to the type of metrical foot that their native language prefers.

In the current study, we further investigate whether the L1 metrical foot structure constrains L2 word recognition and processing. The chosen language combination (i.e., Bengali-English) and the experimental stimuli were set-up in a way that possible differences across conditions cannot be explained by the combination of strong and weak syllables alone. The languages tested were only distantly related, however, the metrical structure of the chosen words was more similar compared to the words we included in our previous study with German-English bilinguals. Most importantly, in our previous study (Fritz et al., 2023) stress always differed between English and German in shared loans. Thus, we cannot rule out that the differences that we found were driven by stress placement alone rather than foot structure differences. We will first outline the linguistic concept of metrical feet based on the two languages relevant for the present study before we will link metrical phonology back to word recognition in bilinguals.

Although Bengali and English are both Indo-European languages as evident in many cognates (e.g., E *two* B [d̪ui]; E *brother,* B [bʰrat̪a]), historical lineage is many thousand years apart. In this study, our focus is not on cognates but on loans from English into Bengali since the 18th century. The influx of English loans into Bengali began when the country came under formal British rule. In the course of two centuries, many ordinary English words were integrated into Bengali in a way that speakers did not overtly realise that these words were not a part of their native lexicon. Words like [ˈekspaɽʈ] *expert* and [ˈɖaktar] *doctor,* are very much part of the Bengali native vocabulary.

A major difference between Bengali and English is the assignment of metrical stress. Although initial stress is the dominant stress pattern in English disyllabic words, particularly nouns, words with final stress also exist, e.g., canteen [kanˈtiːn]. However, in Bengali, all words bear initial stress. Thus, loans from English always carry initial stress in Bengali regardless of the original stress pattern; both *dóctor* and *cantéen* are adapted with initial stress.

Although in explanatory terms stress is described as being assigned to particular syllables (e.g., *gazette* [gəˈzɛt] stress is on the second syllable; *pigeon* [ˈpɪdʒɨn] stress is on the initial syllable), syllables alone are not sufficient to account for stress assignment. Syllables differ in weight. Generally speaking, syllables with long vowels and closed with a consonantal coda are considered *heavy*, while all other syllables are *light.* Syllables are then grouped into feet (trochees and iambs) and languages differ with respect to the type of feet they prefer. Trochees are left headed (strong side is on the left) while iambs are right-headed.

Amongst the trochees, one can distinguish between a *moraic* trochee and a *syllabic* trochee. In English, the preferred foot type is a moraic trochee where two light syllables can be grouped into one foot, or alternatively a single heavy syllable can also be a foot. Contrasting the words *pigeon* and *reptile* illustrates these two options. Although stress falls on the first syllable in both words, they differ in terms of their foot structure. *Pigeon* consists of two light syllables and has only one foot (thus only one stress) while *reptile* can have two stresses and thus consists of two feet. Consequently, a light syllable which is the head of a foot can bear stress (e.g. *pigeon*) while if a word has two feet, then potentially it can bear two stresses with the main stress on the initial syllable and secondary stress on the final syllable (Lahiri, 2001).

This is where English and Bengali differ. Bengali has heavy syllables with coda consonants but does not distinguish vowel length, i.e., all vowels are full vowels (i.e. no schwa) which do not differ in weight and quantity. Weight is irrelevant in building feet in Bengali and thus, the foot structure is a *syllabic trochee*, where two syllables are grouped together regardless of weight and the first syllable bears stress. We illustrate this with two English loans borrowed into Bengali (see Fig. 1).

The word *expert* has two feet in English, with the first foot bearing main stress. The borrowed version in Bengali also has two full syllables, but only one syllabic trochaic foot. Thus, it is not just the main word stress that is critical, but also the structure of the feet which indicates the possibility of more than one stress. Less frequently, English words with two feet also allow main stress to fall on the second foot such as in *càntéen.*

2. The present study

The question that we are addressing in this study is the consequence of the native Bengali phonology on accessing and processing British English disyllabic words. All words were stressed on the initial syllable some of which had the possibility of bearing secondary stress on the final syllable in English.

Half of the words had been absorbed into the native lexicon of Bengali. We conducted an ERP fragment priming study where participants were auditorily presented with the first syllable of the target word followed by the target being visually displayed. Priming paradigms are widely used in psycholinguistics to investigate questions related to lexical access and processing. Generally speaking, priming effects refer to the facilitation or interference in the processing of a target item which is preceded by a related (semantically of phonologically) compared to an unrelated prime (Bölte & Coenen, 2002; Zwitserlood, 1996). Previous studies have shown that the presentation of word onset fragments can prime the target word (Friedrich et al., 2004). Fragments may only consist of the initial syllable cut out from the target word. For example, in English [ˈrɛp] primes *reptile* meaning that RTs are faster compared to *reptile* being preceded by a prime that is phonologically unrelated (Friedrich, Lahiri, & Eulitz, 2008; Marslen-Wilson, 1990; Scharinger & Felder, 2011).

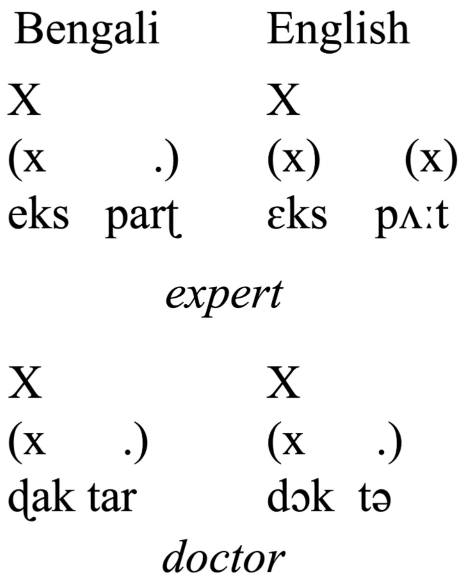


Fig. 1. Example foot structures for English (one foot and two feet words) and their Bengali counterparts. The X indicates the word level whilst the (x) refers to the head of the foot. For both words in Bengali (*expert* and *doctor*), we have a left headed branching foot indicated by the (x.). The two (x) (x) in English *expert* indicate two separate feet.

Fragment priming studies have also been utilized to explore the role of lexical stress during word recognition. Across various languages, research has shown that words preceded by stress-congruent primes facilitate lexical access (i.e., faster reaction times) compared to stress-incongruent primes (cf. for English: Cooper, 2002) whilst stress incongruent primes may inhibit lexical access (Soto-Faraco, Sebastián-Gallés, & Cutler, 2001; van Donselaar, Koster, & Cutler, 2005). Several EEG studies have also used the cross-modal fragment priming paradigm to investigate online word recognition. Friedrich and colleagues (2004) used the stress-congruent vs. incongruent paradigm introduced above, e.g. the prime [ˈre] followed by [ˈreːgl̩] (Regel “rule”) = congruent vs. [ˈre] followed by [reˈgaːl] (Regal “shelf”) = incongruent. In addition to a behavioural priming effect (RTs), the authors also found a reduced ERP amplitude for congruently stressed primes compared to stress incongruent prime-target pairs 300–400 ms post target word onset, and with a left-anterior distribution. They attributed this effect to the P350 which is thought to reflect a phonological mapping between prime fragment and target words. Thus, if the fragment does not match the target, a more positive deflection for phonologically incongruent fragments is expected compared to congruent prime-target pairs.

Generally, the P350 component has been found across several fragment priming studies conducted with L1 speakers (Friedrich, Kotz, Friederici, & Gunter, 2004; Friedrich, Schild, & Röder, 2009). These studies suggest that the P350 is modulated in a rather short time-window (i.e., 300–400 ms post stimulus onset) whilst more recent studies (both with L2 speakers), found a prolonged P350 (i.e., 300–700 ms post target-word onset) (Fritz et al., 2023; Kobor et al., 2018).

In addition, these ERP fragment priming studies reported negativities that were usually elicited between 300–600 ms post stimulus onset and with a broad centro-parietal distribution. For form fragment priming studies, this means that a target is easier to access and integrate for a prime-target match. Hence, a less negative deflection is elicited compared to a prime-target mismatch (e.g., Scharinger & Felder, 2011). Friedrich and colleagues (2004) linked this effect to the phonological N400 which indexes phonological expectancies. Based on previous studies, we take the modulation of the P350 vs. N400 as indicator for different processing/integration mechanisms. Whilst the P350 exclusively reflects matching on a phonological level, the N400 also reflects ease of lexical access and integration processes.

Finally, Fritz and colleagues (2023) found a late positivity (P600) in a condition where the presented target word (i.e., in English) was metrically closest to the German counterpart, i.e., only stress placement differed as in English *reptile* [ˈrɛpˌtʌɪl] and German *Reptil* [rɛpˈtiːl]. Such a posteriorly distributed P600 was also found in a lexical decision task where “identical cognates” (i.e., translation equivalents with the exact same orthography across L1 and L2) were compared to non-cognates and non-identical cognates (Peeters, Dijkstra, & Grainger, 2013). The authors interpreted these results as a conflict resolution process created by an L1 reading of the target (i.e., French) which was presented in English (L2). In Fritz et al. (2023), we argued along the same lines, i.e., the P600 was triggered because of the phonological and orthographical similarity between the presented English word and the German (L1) counterpart.

Our predictions are based on findings reported in our previous study where the same paradigm was used when testing English-German bilinguals as well as metrical theory of English and Bengali as outlined above. In our previous study, the behavioural priming effect (RTs) was significantly larger for words which did not exist in the speakers' L1. We attributed this to interference from the L1 counterpart leading to a reduction of priming for the shared loans. Thus, for the current study, we also expect differences in the extent of priming between shared loans and loans that have not been borrowed into Bengali, with larger priming effects for the ones not existing in Bengali. As for the metrical structure, we would expect that a foot structure match (i.e., one foot) elicits larger behavioural priming effects compared words comprised of two feet which is an *impossible* foot structure in Bengali disyllabic words. As for the ERP data, we expect a modulation of the N400 for the conditions including words that have not been borrowed into Bengali since there is no lexical competition from the L1 counterparts. This is in line with what we found for our German L1 speakers (Fritz et al., 2023). Given the lexical competition, we then predict no significant N400 effect for the shared loans but rather a P350 indicating more shallow mapping mechanisms between the prime and the target.

As for the impact of foot structure, Bengali does *not* allow disyllabic words with two feet. If a language's metrical system constrains processing in an L2, we should see both RTs and ERPs modulated by our *Foot* structure (i.e., one foot vs. two feet) manipulation in a way that two feet words exhibit more processing effort for Bengali speakers. This may either be by the reduction of certain ERP effects or by the modulation of different ERP components (i.e., the N400 or P350).

In Fritz et al. (2023), we attributed the significant P600 effect to reanalysis costs because this effect was elicited in the conditions where the German and English shared loans were orthographically and phonologically most similar. Based on these findings, we formulate two possible outcomes for the present study: If this P600 was driven by phonological similarity, we would expect a P600 effect for the *one foot* items where the foot structure is the same as in Bengali but not for the *two feet* items. If, however, the P600 effect in our previous study was driven by orthographic similarities as suggested by Peeters et al. (2013), we would not expect any modulation of this component due to the different scripts across the two languages.

3. Methods

3.1. 1. Participants

30 Bengali native speakers participated in the study. They were graduate students at Jadavpur University in Kolkata studying either for a degree in English or Linguistics. Two participants had to be excluded due to low performance in the Lexical Decision Task (i.e., below 50 % accuracy). One more participant was excluded due to a very noisy signal for the reference electrodes (TP9/TP10). Thus, leaving us with 27 participants (mean age = 25.6 years; age range = 22–35, 17 female) for the analyses. All participants had normal or corrected-to-normal vision and were right-handed. Furthermore, they did not report any hearing or other cognitive impairments. All participants gave written informed consent and the study was reviewed by CUREC (ethics approval code: R77464/RE001). All students were Bengali native speakers; that is both parents spoke Bengali as their L1. We ensured that no participant had a second primary Indian language spoken at home such as Hindi or Assamese. All participants were educated in Kolkata in English medium schools. This means that their medium of instruction was in English from nursery through university. All students were required to have both English and Bengali as languages for their school leaving exams. Kolkata is an inherently multilingual city, where the educated middle class are equally comfortable in Indian English and Bengali. The English 'proficiency' of the participants is directly correlated with their schooling. Moreover, at home and at university, they would speak a mixture of Bengali and Indian English because code-switching is almost automatic (cf. Mitra & Dutta (2023) for Bengali-English code-switching). Thus, *quantifying* language use, as it is often asked for in standardised questionnaires, is challenging. Given this language situation in Kolkata, we concluded that the participants' school and university background, along with their family history, was a more accurate assessment of their English proficiency compared to a standardised test and is similar to the information reported in Shetty et al. (2022) for Malayalam-English bilinguals tested in the South of India.

3.2. Materials

Our stimuli set consisted of 96 disyllabic English words which all carried initial stress. Half of the words within these two categories (i.e., 1FOOT vs. 2FEET) where English words which have been borrowed into Bengali (SHARED/ABSENT in Bengali). Table 1 illustrates the foot structure for each condition in English and their Bengali counterparts. The English words across all conditions were matched on frequency (pair-wise comparisons all $p < 0.1$). For assessing frequency, we used the Zipf values from the SUBTLEX-UK corpus (van Heuven, Mandera, Keuleers, & Brysbaert, 2014). In terms of the number of phonemes and graphemes, it is an inherent characteristic that the 1Foot words are shorter than the 2Feet ones and thus not allowing us to match these conditions on these two dimensions. Within the 1Foot (present vs. absent) and 2Feet (present vs. absent) categories, both the number of phonemes and the number of graphemes were matched.

Norming values are summarised in Table 1. For the Lexical Decision Task, we also included 96 pseudowords which were created in a way that they either consisted of 1FOOT (e. g., *tilder, *pollet) or 2Feet (e.g., *lampile, *wellow) to mirror the target words in the experimental conditions. Pseudowords were following English phonotactics and were constructed in a way that they are phonologically similar to the Romance loans which were used as experimental items. Although our pseudowords mirrored the types of loans used in our experiment, we did not base them off existing words and then change a certain number of phonemes/letters. For the recordings of the pseudowords, the linguistically trained speaker was told to always place stress on the initial syllable and to make sure to maintain the 1Foot and 2Feet distinction in their pronunciation as indicated by a category label.

We also had 192 control primes (one for each pseudoword and one for each target) which had the same foot structure in English as the target words (either 1FOOT or 2FEET) and were disyllabic, monomorphemic words with initial stress mirroring the target items. The fragment cut out from the disyllabic control-primes had no phonemic overlap with the target word.

All items (primes, targets and pseudowords) were recorded by a male speaker of Southern British English in a sound-proof booth and the recoding was digitised at 44.1 kHz (16 bit, mono). We then cut the recordings in Praat (Boersma & Weenink, 2009) at the offset of the first

Table 1
Norming values for frequencies and number of phonemes (Nphoneme) and graphemes (Ngrapheme) across the conditions plus an example target word. Note, participants were presented with the English word. Bengali counterparts are shown where applicable.

Type	*Foot*	*English*	*Bengali*	*Zipf Value*	*Nphoneme*	*Ngrapheme*
ABSENT	1FOOT	doctor [ˈdɒktə]	ডাক্তার [ˈɖaktar]	3.93	4.79	5.71
	2FEET	expert [ˈɛksˌpʌːt]	এক্সপার্ট[ˈekspaɽt]	3.91	5.79	6.67
SHARED	1FOOT	temple [ˈtɛmpl]	–	4.02	4.67	5.92
	2FEET	reptile [ˈrɛpˌtʌɪl]	–	3.72	5.46	6.42

syllable (e.g., coral [ˈkɒ.rəl], montage [ˈmɒn.ˌtɑːʒ]). The cut recordings were normalised in Praat so that all have a mean amplitude of 70 dB. We then paired each of the targets (word and pseudoword targets) with an experimental prime and a control prime.

The experiment consisted of two blocks with 192 trials each. Within the blocks, each target word appears once either with an experimental or a control prime. The primes were counter-balanced across blocks in a way that participants saw each target (words and pseudowords) twice both with an experimental and a control prime. However, the order in which participants saw the different prime types was counter-balanced. The blocks were pseudorandomised: no more than two trials from the same condition followed each other and no more than five pseudowords were presented in a row. The experiment was performed using Presentation® software.

3.3. Procedure

Participants were tested at Jadavpur University in Kolkata (India). Before starting the experiment, participants read the information sheet and filled in and signed the consent form. Participants were seated approximately 100 cm away from the screen in an air-conditioned booth. Participants were told that they would hear fragments of a word through headphones followed by a letter string appearing on the screen. Their task was to indicate as quickly and as accurately as possible whether what they saw was a real word in English or not. Importantly, participants were instructed to only respond to what they saw but not to what they heard. After explaining the task, participants completed a practice round including 16 trials. If the accuracy in these trials was very low, the practice was repeated, and the participant had the chance to ask any (additional) clarification questions.

All targets were shown in lowercase letters in a white 40-point Arial font against a black background. Trials started with a fixation cross shown for 500 ms followed by the fragment prime (auditory). The ISI between the prime and the target item was 250 ms. The target was presented for 500 ms and participants had a total of 1500 ms to respond to the target before the next trial started. Responses were collected with a game controller. After 36 trials, the screen turned blank and the participants heard a beep. This indicated a short break of five seconds to reduce fatigue during the experiment. After the first block participants could take a longer break. The whole experiment, including the break between lists, took about 25 mins to complete.

4. Data analyses

4.1. Behavioural data

Prior to running the statistical analyses on the RT data, we removed RT responses that were ± 3 standard deviations away from a subject's mean as well as response times shorter than 200 ms. This led to the exclusion of 123 trials (i.e., 1.19 %). We further excluded all incorrect responses and pseudowords. Behavioural responses (RTs) were analysed by fitting linear mixed-effects models in R by using the lmer function of the lme4 package (Bates, Kliegl, Vasishth, & Baayen, 2015). We first tested whether there are significant main effects and/or interactions of the variables *Prime, Type and Feet.* For that, we fitted a model including log-transformed RTs as dependent variable and *Prime* (experimental/control), *Type* (shared, absent), *Feet* (1Foot, 2Feet) and their interactions including the three-way interaction, as fixed factors. For the random effects structure, we treated subject and item as random effects. We further fitted a random slope structure for subject including *Prime*, *Feet* and *Type* but not their interactions. This random-effects structure was design-driven rather than determined via likelihood ratio tests, i.e., the random effect structure was selected prior to our analysis. This approach is in line with Barr, Levy, Scheepers, & Tily (2013) arguing in favour of a theoretically defendable parsimonious linear model which is often seen in opposition to a data driven approach favoured by the Matuschek, Kliegl, Vasishth, Baayen, & Bates (2017). Thus, within this approach, variables are not eliminated because they do not improve the fit of the model. To derive p-values (i.e., *Prime*, *Type* and *Feet*), we used the summary() function. More specifically, in these model summaries, p-values were derived from the t-values in the lmer output by using the Satterthwaite approximation which is implemented in the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017). This method was used because it allowed us to extract outputs for both fixed and random effects.[1]

In a next step, we tested whether there was a significant priming effect within each condition, i.e., whether RTs differed significantly between the experimental and control items. For that, we ran a planned pair-wise comparison on a model that included only the two fixed effects *Prime*Condition* (four levels: 1Foot shared, 2Feet shared, 1foot absent, 2feet absent). We then used the testInteractions() function which performs a Wald χ^2 test that is implemented in the *phia* package (De Rosario-Martinez, 2015). For the random effects structure, we treated subject and item as random effects. Further, we fitted a random slope structure for subject including *Prime*, but due to convergence issues not for *Condition.* If we find significant priming effects (i.e., significant differences between experimental and control items) in more than one condition, we also use the summary() function to compare the strength of the priming effect across conditions. Importantly, this function provides parameter-specific p-values which is useful for factors including more than two levels. More specifically, this allows us to statistically test for possible differences in the strength/extent of a priming effect across our four conditions, i.e., the reference level is compared to all other levels separately even for interactions (e.g., *Prime* * *1foot shared* vs. *Prime* * *2feet shared*) (cf. Luke, 2016 for evaluating significance in linear mixed-effects models).

4.2. EEG recording, processing and analyses

Continuous EEG was recorded with 32 active electrodes (ActiCap, Brain Products, Inc., Gilching, Germany). Electrodes were placed according to the international 10–10 system. During data acquisition, FCz was used as reference channel and the data was sampled at 1000 Hz. Impedance was kept below 25 kOhm wherever possible. All EEG channels were then re-referenced offline to the mastoids (TP9, TP10) which were placed outside of the elastic cap on the left and right mastoid respectively. This was done because the cap positions of TP9/TP10 only approximately met left and right mastoids. A further modification to the standard cap layout was made to the two electrodes FP9 and FP10 of which one was used as the reference in FCz and the second placed in CPz, one of the electrode sites where we expect the N400 component to be strongest.

EEG data was analysed using the EEGlab (Delorme & Makeig, 2004) and ERPlab (Lopez-Calderon & Luck, 2014). Data were filtered offline using a band-pass filter of 0.1–30 Hz. EEG triggers were time-locked to the onset of the target word and data was epoched from 200 ms prior to word onset to 800 ms post word onset using a baseline correction. Eyeblinks were removed by running an Independent Component Analysis (ICA) using the analysis implemented in EEGlab and following the procedure described by Nunez and colleagues (2016). After eye-blink correction, trials were rejected according to the following criteria: We utilized an automatic artifact rejection procedure with a 200 ms sliding window and 50 ms steps. If any of the electrodes exceeded a ± 100 μV threshold within this time-window, the epoch was excluded from the analyses. This pre-processing pipeline led to the exclusion of 194 trials which comprised 1.87 % of the data across the four conditions (not considering pseudowords). Very noisy electrodes were interpolated by

---

[1] Obtaining p-values with Type III Wald chi-square tests (ANOVAs) yielded the same results for main effects and interactions for both the behavioural and the ERP data.

using the surrounding electrodes to reconstruct the signal. Across the whole dataset only three electrodes which were later used in the analyses were interpolated.

As described earlier, previous ERP fragment priming studies reported different ERP components to be affected by the experimental manipulations. These are the N400, the P350 and the P600. Especially for the P350 different time-windows were reported for the effect. Studies with L1 speakers found short effects from 300-400 ms post stimulus onset (Friedrich, Kotz, Friederici, & Alter, 2004; Friedrich, Kotz, Friederici, & Gunter, 2004) whilst L2 studies found prolonged P350 lasting throughout the epoch (Fritz et al., 2023; Kobor et al.,2018). To capture these three possible ERP effects, we used the same statistical approach as for the behavioural data (i.e., LMMs) running models in 100 ms increments from 200-700 ms post-stimulus onset. As for the regions of interest (ROIs), the topographical distributions of the three ERP components of interest are well-attested. Whilst the N400 and P600 usually show a centro-parietal distribution (N400: Kutas & Federmeier, 2011; P600 for conflict detection: Kuperberg, Brothers, & Wlotko, 2020), the P350 is sought to be strongest over left-anterior sites (Friederich et al., 2009). Moreover, in Fritz et al. (2023) we used cluster mass statistic which is suitable for broadly distributed ERP effects reflecting higher level cognitive processes (Groppe et al., 2011). This data-driven approach was used because fragment priming studies with L2 speakers were sparse (only Kobor et al., 2018). Thus, it was largely unknown whether the time-course and topographical distribution of ERP effects of interest (i.e., N400, P350, P600), would align with the previous studies. Given that the present study is methodologically and stimuli-wise very similar to our previous study and that the effects found there aligned with the well attested distributions of the P350, N400 and P600, we limited our analyses to two ROIs with seven electrodes each: left-anterior (FCz, Fz, FC1, FC5, F3, F7, Fp1) and centro-parietal (Cz, CPz, Pz, P4, P3, CP1, CP2).

We calculated the mean voltages for both ROIs. For each 100 ms time-window, we ran liner mixed-effects models with the averaged EEG voltages as dependent variable. Since ROI was not a factor of interest in our study, we ran separate models for each of the two ROIs. We fitted the same models as for the behavioural data outlined above, i.e., a two-step analysis, starting with a model including *Prime, Type and Feet* and their interactions, followed by fitting a model including *Prime * Condition* from which we get planned pair-wise comparisons of experimental vs. control items across the four conditions. So, we can test for significant ERP priming effects. For the *Prime * Condition* models, a by-participant random slope was only included for the factor *Prime*. This was because of convergence problems and to ensure a consistent random effects structure across all time windows (cf. Tsang & Zou, 2022).

For the topographical maps and brainwaves, we grand-averaged the ERPs across subjects for each condition.

5. Results

5.1. Behavioural

Participants' responses showed a very high accuracy rate across all conditions. With the lowest (absent 2feet) being at 92 %. For pseudowords, accuracy also reached over 80 % indicating that participants performed overall very well in the Lexical Decision Task. Since the accuracies in the word conditions are for most participants close to ceiling, we did not run any statistical test on that data. Descriptive statistics are summarised in Table 2 including the accuracies and RTs for pseudowords. Fig. 2 shows the priming effect (in ms) across conditions.

In a first step, we tested for effects of *Prime*, *Type* and *Feet* and their interactions. The results revealed a significant main effect of *Prime* ($\beta = 6.30$, $t = 228.04$, $p < 0.001$) as well as a significant *Prime * Feet* interaction ($\beta = 0.03$, $t = 2.21$, $p = 0.027$). None of the other interactions including the factor *Prime* reached significance (*Prime * Type*: $\beta = 0.01$, $t = 0.49$, $p = 0.626$; *Prime * Type * Feet*: $\beta = -0.03$, $t = -1.26$, $p = 0.208$). The full summary output can be found in Table 3.

In a next step, we fitted a model with *Prime*Condition* as fixed factors on which we then ran the planned pair-wise comparison using the testInteraction() function in the phia package (Wald χ^2 test). By doing so, we tested for significant priming effects (experimental vs. control items) across the four conditions. Table 4 shows the $\chi2$ statistics for each of the four pair-wise comparisons. Since all of the pair-wise comparisons were significant, we also looked at the model's summary output which provides parameter-specific p-values. The only significant difference in the priming strength was found between *Prime * shared 1foot* (32 ms) and *Prime * absent 2feet* (14 ms) ($\beta = 0.032$, $SE = 0.014$, $t = 2.232$, $p = 0.026$). The full model output including the comparison of priming strength across conditions can be found in the supplementary materials (Table 1).

5.2. ERP results

Topographic plots and brainwaves across conditions are shown in Fig. 5. As for the behavioural (RT) data, we conducted a two-step analysis beginning with a model that includes *Prime * Type * Feet* as fixed effects. For the ERP results, we report interactions involving the factor *Prime* across the five time-windows investigated, i.e., in 100 ms increments from 200 to 700 post-stimulus onset for both ROIs (centro-parietal and left-anterior). Significant main effects and interactions for both ROIs across all time-windows are reported in Table 5. The full summary outputs for the 300–400 ms time-window and for both ROIs can be found in Appendix B and the remaining time-windows in the supplementary materials.

For the centro-parietal ROI, we found a significant *Prime*Type* interaction across four time-windows (200–300 ms: ($\beta = 2.05$, $t = 2.62$, $p = 0.009$; 300–400 ms: $\beta = 2.37$, $t = 2.72$, $p = 0.007$; 400–500 ms: $\beta = 2.42$, $t = 2.49$, $p = 0.013$; 500–600 ms: $\beta = 2.41$, $t = 2.35$, $p = 0.019$). The difference between experimental and control items is larger for the absent compared to the shared loans (i.e., more negative for control items), and even reversed for the 1foot shared loans (for example in the 200–300 ms time-window). Fig. 3 illustrates the difference in mean EEG amplitude for control vs. experimental items across conditions.

However, in our planned pair-wise comparisons of experimental and control items, only the absent 1Foot condition showed a significant priming effect in the 200–300 ms ($\chi2$ (1) = 8.44, $p = 0.015$) and 300–400 ms ($\chi2$ (1) = 12.36, $p = 0.002$) time-windows. Similarly, with an early onset, topographical plots also suggest a significant N400 effect

Table 2
Accuracies (%) and Reaction Times (ms) in the experimental and control conditions including their standard deviation (SD). Priming effect (ms) in each of the four conditions plus for pseudowords (pseudo).

Type	Feet	Control Accuracy	Experimental Accuracy	Control RTs	Experimental RTs	Priming Effect
shared	1FOOT	92.7 (9.1)	95.8 (5.8)	558 (86)	526 (66)	32
shared	2FEET	93.5 (7.2)	94.1 (6.8)	550 (72)	536 (65)	14
absent	1FOOT	95.2 (6.5)	94.8 (6.6)	549 (82)	521 (72)	28
absent	2FEET	90.0 (11.2)	92.1 (9.4)	562 (78)	541 (79)	21
Pseudo	1FOOT	86.7 (12.4)	80.9 (16.6)	648 (108)	653 (108)	−5
Pseudo	2FEET	86.6 (10.7)	83.6 (13.5)	660 (111)	657 (104)	3

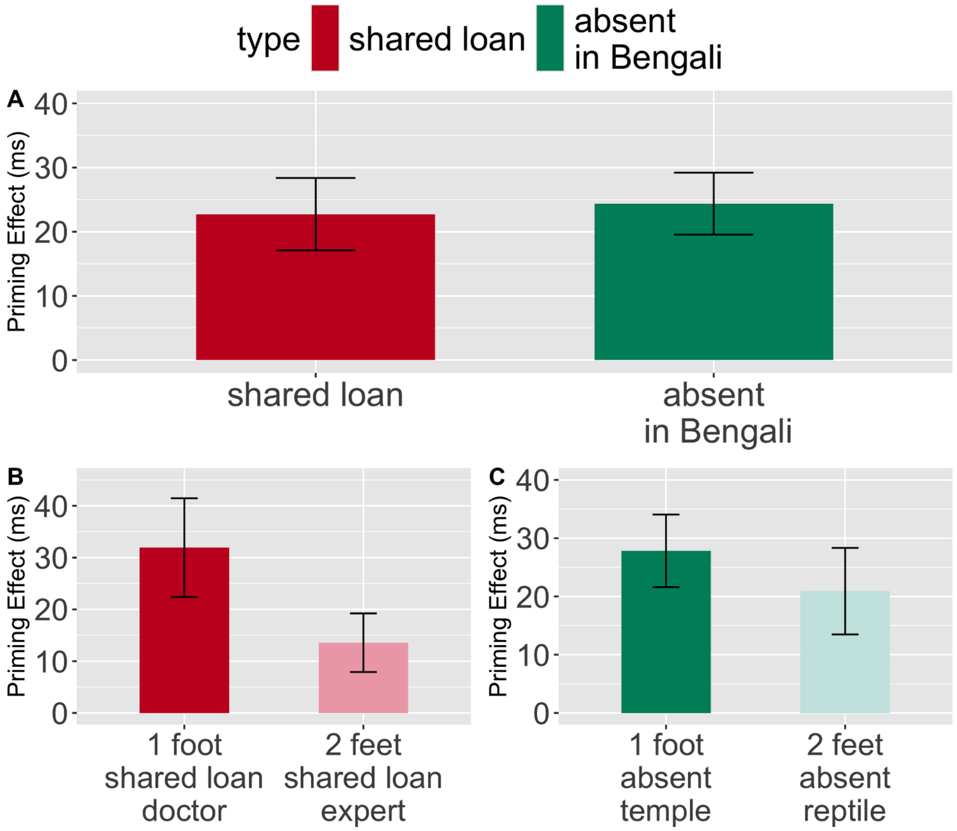


Fig. 2. **A.** Priming effects for the two Word Types collapsed over One Foot and Two Feet items (RTs of control items – RTs of experimental items). **B-C.** Each of the two Types split based on Foot Type. Error bars represent the standard error of the mean.

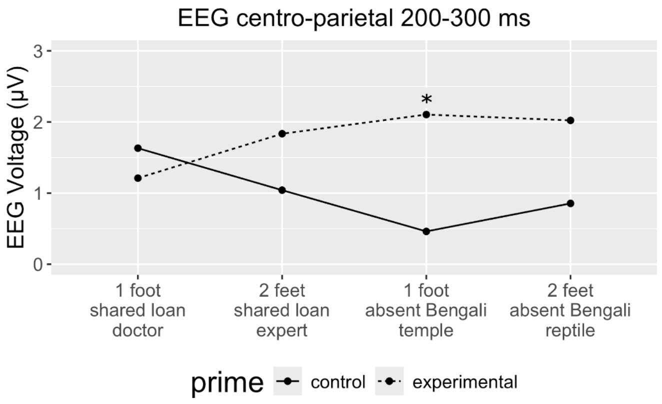


Fig. 3. Mean EEG amplitude comparing control and experimental items across conditions in the centro-parietal ROI, 200–300 ms post-stimulus onset. Asterisk indicates significant pair-wise comparison (experimental vs. control prime).

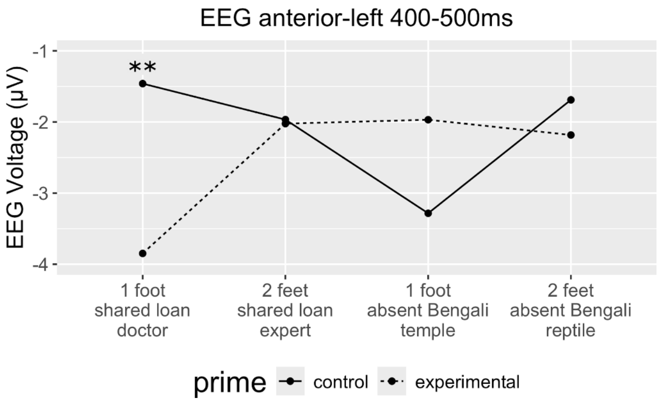


Fig. 4. Mean EEG amplitude comparing control and experimental items across conditions in the anterior-left ROI, 400–500 ms post-stimulus onset. Asterisks indicate significant pair-wise comparison (experimental vs. control prime).

in the ABSENT 2FEET CONDITION. However, in none of the pair-wise comparisons did we find such a significant effect (200–300 ms: $\chi2\ (1) = 4.39$, $p = 0.109$; 300–400 ms: $\chi2\ (1) = 0.081$, $p = 0.837$). A summary of significant pair-wise comparisons in both ROIs are illustrated in Table 6. Statistics of all pair-wise comparisons can be found in Appendix B.

For the models run on the left-anterior data, we found significant three-way interactions (i.e., *Prime* * *Type* * *Feet*) across all time-windows strongest between 300–500 ms (300–400 ms: $\beta = -3.39$, $t = -2.85$, $p = 0.004$; 400–500: $\beta = -4.14$, $t = -3.22$, $p = 0.001$). In addition, the *Prime* * *Type* interaction reached significance across all time-windows, i.e., 200–300 ms: $\beta = 2.36$, $t = 3.05$, $p = 0.002$; 300–400 ms: $\beta = 2.82$, $t = 3.35$, $p = 0.001$; 400–500 ms: $\beta = 3.67$, $t = 4.03$, $p < 0.001$; 500–600 ms: $\beta = 3.32$, $t = 3.45$, $p = 0.001$; 600–700 ms: $\beta = 3.00$, $t = 2.96$, $p = 0.003$). Finally, the *Prime* * *Feet* interaction was significant in three different time-windows (200–300 ms: $\beta = 1.66$, $t = 2.14$, $p = 0.032$; 400–500 ms: $\beta = 2.35$, $t = 2.58$, $p = 0.010$; 500–600 ms: $\beta = 2.08$, $t = 2.15$, $p = 0.032$). Fig. 4 illustrates the impact of the factor *Prime* on the mean EEG amplitudes across conditions in the 400–500 ms time-window.

Planned pair-wise comparisons of experimental and control items, revealed a significant effect in the SHARED 1FOOT condition in all but the early time-window (i.e., 200–300) and was strongest in the 400–500 ms time-window (300–400 ms: $\chi2\ (1) = 6.35$, $p = 0.047$; 400–500 ms: $\chi2\ (1) = 12.64$, $p = 0.002$; 500–600 ms: $\chi2\ (1) = 11.33$, $p = 0.003$; 600–700 ms: $\chi2\ (1) = 6.91$, $p = 0.034$). Given both the topography and the latency, we interpret this effect as a prolonged P350 as found in

Table 3

Summary of the RT full model for the behavioural data. Square brackets indicate the factor level comparisons. * <.05, *** <.001; Reference levels were *control* for the factor *Prime*; Shared for the factor *Type*; 1foot for the factor *Feet*.

	RTsLog			
Predictors	*Estimates*	*CI*	*t-value*	*p*
(Intercept)	6.30	6.25 – 6.36	228.04	**<0.001** ***
prime [experimental]	−0.06	−0.08 – -0.04	−5.44	**<0.001** ***
type [absent]	−0.02	−0.06 – 0.02	−0.92	0.355
foot [2feet]	−0.01	−0.05 – 0.03	−0.46	0.643
prime [experimental] × type [absent]	0.01	−0.02 – 0.03	0.49	0.626
prime [experimental] × foot [2feet]	0.03	0.00 – 0.06	2.21	**0.027** *
type [absent] × feet [2feet]	0.04	−0.01 – 0.10	1.48	0.138
prime [experimental] × type [absent] × feet [2feet]	−0.03	−0.06 – 0.01	−1.26	0.208
Random Effects				
σ^2	0.03			
$\tau_{00\ \mathrm{item}}$	0.00			
$\tau_{00\ \mathrm{subject}}$	0.02			
$\tau_{11\ \mathrm{subject.primeexperimental}}$	0.00			
$\tau_{11\ \mathrm{subject.typeabsent}}$	0.00			
$\tau_{11\ \mathrm{subject.foot2feet}}$	0.00			
$\rho_{01\ \mathrm{subject.primeexperimental}}$	−0.38			
$\rho_{01\ \mathrm{subject.typeabsent}}$	0.59			
$\rho_{01\ \mathrm{subject.foot2feet}}$	−0.34			
N_{subject}	27			
N_{item}	96			
Observations	4827			
Marginal R^2 / Conditional R^2	0.023 / NA			

Syntax: RTsLog ~ prime * type * feet + (1 + prime + type + feet|subject) + (1|item).

Table 4

Planned pair-wise comparison of control-experimental items across the four conditions. (Wald χ^2 test). * <.05, ***<.0001.

Type	Feet	Value	SE	DF	Chisq	Pr(>Chisq)
shared	1FOOT	0.058	0.011	1	30.22	**<.0001** ***
shared	2FEET	0.051	0.011	1	6.26	**0.012** *
absent	1FOOT	0.051	0.011	1	23.89	**<.0001** ***
absent	2FEET	0.045	0.011	1	17.65	**<.0001** ***

previous fragment priming studies with L2 speakers (i.e., Fritz et al., 2023; Kobor et al., 2018). In the shared 2Feet condition, pair-wise comparisons did not reach significance in any of the tested time-windows (300–400 ms: $\chi2$ (1) = 0.024, p = 0.876; 400–500 ms: $\chi2$ (1) = 0.000, p = 0.992; 500–600 ms: $\chi2$ (1) = 0.091, p = 1; 600–700 ms: $\chi2$ (1) = 0.337, p = 1); rather the elicited EEG amplitude between experimental and control items were very similar. The same applies for the two absent conditions. Outputs of all pair-wise comparisons are listed in Appendix B. Since no more than one pair-wise comparison (i.e., control-experimental) was significant in each time-window and ROI, we did not compare the strength of the priming effect across conditions as we did for the behavioural (RT) data.

6. Discussion

We set out to test whether L1 metrical phonology impacts L2 word recognition across distantly related languages. For that, we employed an ERP fragment priming paradigm in English with L1 speakers of Bengali. The experimental paradigm included two manipulations. First, we varied the type of foot structure of the presented words (i.e., 1foot vs. 2feet). Second, half of the English items were borrowed into Bengali (i.e., shared) whilst the other half had a translation equivalent which was unrelated in form (phonology) to the Bengali counterpart (i.e., absent).

Our results indicate that both manipulations (i.e., *foot structure, word type*) affected the processing of the target words but in different ways. Although we found significant priming effects in our RT data across all four conditions, those effects were larger for items consisting of 1foot. Further, it did not matter whether the English word was borrowed into Bengali or not. This finding suggests that Bengali L1 speakers' processing of L2 English words is guided by their L1 metrical structure which does not allow disyllabic words with two feet.

This preference for 1foot words is also evident in our ERP data. Here, the 2feet conditions did not elicit any significant ERP priming effects when comparing control and experimental items. For the 1foot conditions, we did find significant effects but our *word type* manipulation (i.e., shared vs. absent) modulated ERPs differently. In the absent condition, we found an N400 effect with the experimental items eliciting a less negative response indicating ease of integration when the prime matched the target. In the shared condition, however, the P350 was modulated when comparing experimental with control items. These findings suggest a degree of interference when the prime in the shared condition activates the Bengali cohort more strongly due to a close phonological match across the languages. This interference may also have blocked the elicitation of an N400 effect in the shared conditions. Although pair-wise comparisons are not significant in the 2feet conditions, significant interactions (*Type*Prime*) showed that the absent and the shared items regardless of foot structure elicited similar ERP patterns, i.e. a modulation of the P350 for shared items and a modulation of the N400 for absent items. We did not find any late effects (P600) in our data. We will now discuss the implications of these results for the word recognition and the bilingual literature also drawing from results reported in our previous ERP fragment priming study (Fritz et al., 2023).

6.1. *Metrical phonology*

Phonology does not just include segmental levels but also higher-level structures such as syllables and feet. Metrical phonology is what accounts for grouping syllables into feet to account for the assignment of stress in any given languages. Thus, the metrical structure of L1 has equal potential to influence recognition of L2 words similar to that of segments and syllables. Our results do indicate that the parsing of a word is influenced by

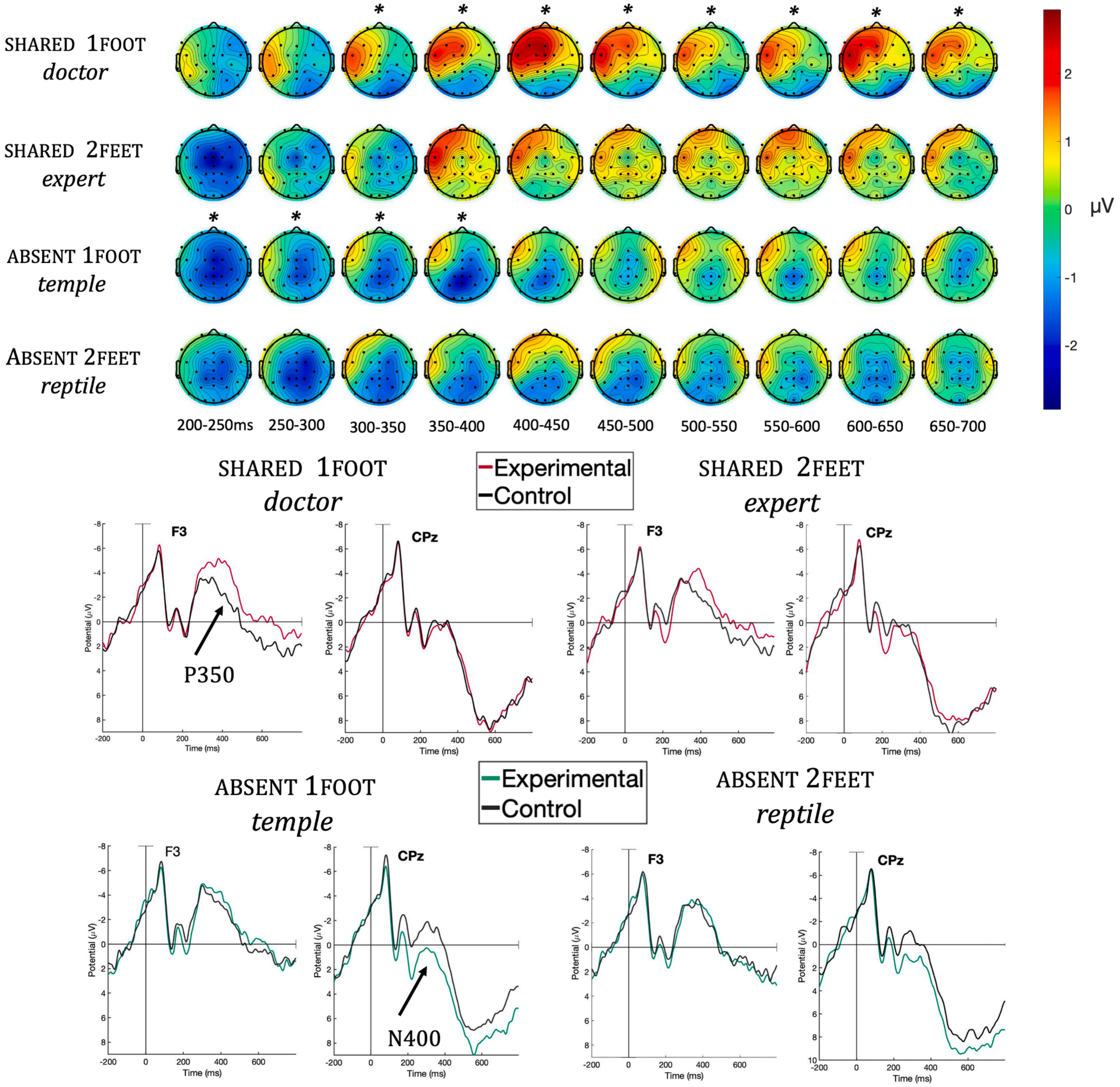


Fig. 5. Top: Topographic plots illustrating differences between the control minus the experimental items for each of the four conditions between 200–700 ms post stimulus onset averaged in 50 ms bins. Asterisks (*) indicate significant pair-wise comparisons of the experimental and the control items. Bottom: Brainwaves time-locked to the target-word onset showing experimental vs. control conditions for SHARED and ABSENT items across 1FOOT and 2FEET. Electrodes are taken from the two ROIs used in the analyses.

the speakers' L1 metrical phonology. Although previous studies have shown the importance of syllables in language processing (Mehler, Norris, & Segui, 1983; Cutler & Norris, 1988), our results cannot be explained by only looking at the combination of strong and weak syllables. All items in our study bore initial stress in both English and Bengali. The difference across the conditions were in foot structure and how the language groups disyllabic words into either one or two feet. Participants always heard a stressed syllable. What we manipulated was whether the syllable was a head of a single foot word or a complete foot taken from a two feet word.

Our results suggest that when encountering a stressed fragment, the word is parsed according to its foot structure. Although Bengali and English both favour a trochaic foot, Bengali allows for syllabic trochees (non-weight sensitive) while English prefers moraic trochees. Thus, for Bengali all syllables have the same weight regardless of the syllable type. In much longer words, the beats could be rhythmic based on syllabic grouping but in English, syllable weight is critical to build feet. Thus, in disyllabic words, Bengali words are always comprised of one-foot and thus one stress but English may or may not have two feet and thereby two stresses.

The impact of Bengali's preferred metrical structure was evident in both smaller behavioural priming effects (RTs) and no significant modulation of any ERP components for two-feet English words. Together with results reported in our previous paper (Fritz et al., 2023), we showed that the preferred L1 foot structure does not only impact the recognition of shared loans but L2 word processing more generally. This is an important contribution to the bilingual literature since most bilingual word recognition studies treated non-cognates (non-shared loans) exclusively as a control group comparing them to cognates (shared loans) with varying degrees of phonological/orthographic overlap across languages (cf. Dijkstra, 2007 for an overview). Our results, however, suggest a more complex picture, i.e. the L1 phonology guides word recognition regardless of the word's counterpart in a speaker's L1. Here, we focussed on metrical phonology, however, similar experiments could be conducted testing

Table 5
Results from the Prime * Type * Feet LMM for each of the five time-windows and two ROIs (centro-parietal (top), anterior left (bottom)). For full model outputs, see Appendix B (time-window 300–400) and supplementary materials for the remaining time-windows. * <.05, **<.01, *** <.001.

ROI: Centro-parietal	Time-window				
Effect	200–300	300–400	400–500	500–600	600–700
Prime					
Type		*			
Feet					
Prime × Type	**	**	*	*	
Prime × Feet					
Type × Feet					
Prime × Type × Feet					
ROI: Left anterior	Time-window				
Effect	200–300	300–400	400–500	500–600	600–700
Prime	*	*	***	***	**
Type	*	*	*	*	*
Feet					
Prime × Type	**	**	***	***	**
Prime × Feet	*		**	*	
Type × Feet	*	*			*
Prime × Type × Feet	*	**	**	*	*

whether or to what extent L1 segmental phonology impacts L2 processing of words that have usually been used as controls (i.e., non-cognates or non-shared loans).

6.2. Loan ('cognate') status

In this study, our manipulation of loanword status (SHARED vs. ABSENT) did not impact our behavioural priming effects (RT) but modulated different ERP components depending on whether an English word was borrowed into Bengali or not, i.e. the P350 for shared loans and the N400 for words that were never borrowed into Bengali. In our study with German L1 speakers, we did find a significantly larger priming effect for non-shared loans. We attributed smaller effects for shared loans to interference from the German counterparts which did not exist for words which were not borrowed into German. Given the lack of such an effect for Bengali speakers, we speculate that the interference in the German L1 study stemmed from the high orthographic similarity or even identical orthography between the Romance loans in German and English (e.g., identical: signal – Signal; high overlap: reptile – Reptil) whilst being metrically dissimilar. Note, this is the opposite effect as found in other studies comparing cognates with non-cognates, i.e. *facilitation effect* (cf. Lijewska, 2020). For the current study, we suspect that same stress placement and/or the different scripts across languages *reduced* interference from the L1 counterpart and thus no significant RT priming differences were observed. Based on the ERP results, however, the fragment prime presented in the speakers' L2 still co-activated the L1 counterpart as also shown in previous studies with bilinguals from cross-script languages (Degani, Prior, & Hajajra, 2018; Voga & Grainger, 2007). Thus, our results provide further evidence that both languages are co-activated from phonological input alone.

Further, we interpret the lack of an N400 effect for shared loans as interference from the L1 counterpart. Within a priming context, the modulation of the N400 has been related to phonological expectancies derived from the cohort activated by the prime. Evidence for this interpretation comes from Friedrich and colleagues (2009) where they found that the length of the fragment prime modulated the size of the N400 effect, i.e., longer fragments showed a larger reduction of the N400 compared to shorter fragments. This was not the case for the P350 which is thought to reflect modality-independent activation of word form representations (Friederich et al., 2009; Kobor et al., 2018) that only relies on orthographic and/or phonological matching. Consequently, the P350 is not sensitive to the co-activated neighbours whilst the N400 takes the cohort into account.

The current study provides further evidence that in the ABSENT conditions the L1 cohort would not be activated by the auditory prime and thus, the number of competing items is lower compared to the SHARED conditions. Consequently, we see a significant modulation of the N400 component in the 1FOOT ABSENT condition. Descriptively, we found a similar modulation in the 2FEET ABSENT condition which, however, did not reach significance. As outlined above, we conclude that the results can only be interpreted if we take into account Bengali's native metrical phonology which does not permit disyllabic words to have two feet, thus preventing a secondary stress. Thus, this *structural mismatch* between L1 and L2 prevents a stronger facilitation of the matching prime-target items.

6.3. Models of bilingual word recognition

Zooming into the L2 literature, our results are in line with recent accounts supporting an integrated bilingual lexicon (e.g., Dijkstra & van Heuven, 2001 for BIA + and Dijkstra et al., 2019 for Multilink) as our ERP data was modulated by whether the stimulus word was borrowed into Bengali or not. Importantly, these models assume orthographically-driven co-activation which is based on research with same-script languages. There is, however, evidence for phonology-based co-activation in different-script bilinguals when tested in a monolingual task with auditory stimuli (e.g., Moon & Jiang, 2012; Spivey & Marian, 1999). We assume that similar mechanisms must have been at play in our study, i.e., the fragment prime activated the shared loans in both languages. As pointed out by Jiang (2019), if Multilink was to include phonology-based co-activation, it would have to replace language-specific phonological representations since, in the current version of the model, the phonological representation of any given lexeme is linked to only one language.

This idea of an at least partially-integrated phonological system is evidenced, for example, in adult L2 learners who have not developed L2 specific phonetic categories (yet). Rather they map L2 sounds onto the closest counterparts in their L1. This phenomenon is well attested in the speech perception and developmental literature and has been referred to as *perceptual assimilation* (cf. Best & Tyler, 2007 for an overview). Even if new L2 categories were created, this does not mean that co-activation diminishes. For bilingual word recognition, this means that a "slight mismatch" in phonetic categories between L1 and L2 is not sufficient to eliminate competitors (Dijkstra & Van Heuven, 2002, p. 194). What needs to be tested empirically, however, is to what extent and on which levels such *mismatches* can occur. Relevant for our current study is Jiang's (2019) proposal that phonology driven co-activation in a monolingual context occurs *wherever* there is phonological overlap and

Table 6
Overview of planned pair-wise comparisons between control vs. experimental items across our four conditions, for the five time-windows and across the two ROIs. * <.05, **<.01, *** <.001. The (+) and (−) refer to the polarity of the effect with the control items as reference level. For full statistical outputs, see Appendix B.

	200–300		*300–400*		*400–500*		*500–600*		*600–700*	
	ABSENT	*SHARED*	*ABSENT*	*SHARED*	*ABSENT*	*SHARED*	*ABSENT*	*SHARED*	*ABSENT*	*SHARED*
1FOOT										
Left anterior				* (+)		** (+)		** (+)		* (+)
Centro-parietal	* (−)		** (−)							
2 FEET										
Left anterior										
Centro-parietal										

consequently is not limited to cognates and cross-language homophones.

Within this discussion of phonology driven co-activation, our study makes two contributions. First, and as outlined above, we show that the L1 phonology not only has an impact on shared loans (cognates) but on L2 word processing more generally. In particular, we propose that more general operations are at play that result in transfers from the L1 phonology to L2 word recognition and processing. This is in line with a (partially) integrated phonological system in bilinguals and also supported by earlier studies on language specific segmentation strategies (Cutler et al., 1983, 1989). Although these mechanisms need to be tested further, such general transfers should be incorporated in bilingual word recognition models. For our study, this means that L2 English speakers with L1 Bengali process English words via their preferred L1 metrical structure. As mentioned before, these transfers may not be limited to metrical structure as in the present study and in Fritz et al. (2023), but could be empirically tested on other phonological levels.

Second, in order to test phonological co-activation systematically and test which phonological level(s) are crucial for co-activation, experiments guided by historical developments and separating cognates from loans can be useful. Our argument for such an approach relates to the systematicity of both sound change and loan adaption processes. This more finely-grained approach that is driven by phonological theory, allows to test step-by-step which types of phonological *mismatch* between L1 and L2 are permitted so that phonological co-activation is still achieved; an important research agenda as already pointed out by Dijkstra & Van Heuven (2002). For the present study, we chose shared loans because they are phonologically more similar across English and Bengali than most cognates would have been. This is due to the distant relationship of the two languages. Cognates, on the other hand, could be used to test whether or to what extent specific systematic phonemic differences between languages impact processing, especially in closely related languages. For example, in German, as a result of the Second Sound Shift, word final voiceless stops have become fricatives such as /p/ > /f/. Thus, *all* inherited words with a final /p/ in English will correspond to /f/ in German as in ship [ʃɪp] and Schiff [ʃɪf]. This is an alternative to current approaches in the bilingual literature where phonological similarity is usually calculated by an algorithm weighting different types of phonemic, sub-phonemic (i.e., features) and supra-segmental differences in cross-language word pairs as either the same (e.g., normalized Levensthein Distance: Schepens et al., 2013) or by criteria that have not been experimentally tested (ALINE: Kondrak (2003)).

CRediT authorship contribution statement

Isabella Fritz: Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Aditi Lahiri:** Writing – review & editing, Project administration, Methodology, Funding acquisition, Conceptualization.

Funding

This work was supported by the Engineering and Physical Sciences Research Council UK.

PI: A. Lahiri. Grant number: EP/X026035/1 (Pertinacity).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We are extremely grateful to the faculty, students and staff of Jadavpur University, School of Languages and Linguistics, India particularly Professor Indranil Dutta and Professor Samir Karmakar. We would also like to thank our research assistant in Jadavpur, Abhilasha Thakuria and the research assistant in Oxford Jiarui Zhang, who provided invaluable assistance. We would also like to thank members of the Language & Brain Laboratory in particular Joshua Booth and Hilary Wynne.

Appendix A

Stimuli List

1FOOT SHARED		2FEET SHARED		1FOOT ABSENT	2FEET ABSENT
album	এলবাম	colleague	কলিগ	angel	ambush
atom	এটম	congress	কংগ্রেস	barber	archive
bandage	ব্যানডেজ	contact	কনট্যাক্ট	battle	athlete
cabin	ক্যাবিন	coupon	কুপন	coral	bistro
cable	কেবিল	curfew	কারফিউ	cousin	costume
centre	সেন্টার	depot	ডিপো	digit	exile
circus	সার্কাস	duty	ডিউটি	legend	hero
collar	কলার	expert	এক্সপার্ট	mammal	impulse
diet	ডায়েট	garage	গ্যারেজ	melon	insect
doctor	ডাক্তার	glucose	গ্লুকোজ	monster	journey
license	লাইসেন্স	hygiene	হাইজিন	moral	laundry
locket	লকেট	migraine	মাইগ্রেন	mustard	mandate
medal	মেডেল	perfume	পারফিউম	organ	montage
method	মেথড	picnic	পিকনিক	panther	mushroom
number	নম্বর	profile	প্রোফাইল	portal	nomad
passage	প্যাসেজ	programme	প্রোগ্রাম	rival	octave
pedal	পেডাল	protein	প্রোটিন	river	process
pilot	পাইলট	radar	রেডার	scandal	refuge
powder	পাউডার	statue	স্ট্যাচু	silence	reptile
salad	স্যালাড	textile	টেক্সটাইল	spinach	template
sandal	স্যান্ডেল	transplant	ট্রান্সপ্ল্যানট	temper	termite
senate	সেনেট	trophy	ট্রফি	temple	tribute
siren	সাইরেন	vaccine	ভ্যাকসিন	tiger	turmoil
surgeon	সার্জেন	venue	ভেনু	villain	value

Appendix B

Table 1

Model output summary(model) for single-trial ERP data in **the centro-parietal ROI 300–400 ms** post-stimulus onset. * <.05, **<.01, *** <.001. Square brackets indicate the factor level comparisons. Reference levels were *control* for the factor *Prime*; Shared for the factor *Type*; 1foot for the factor *Feet.*

	EEG			
Predictors	*Estimates*	*CI*	*t-values*	*p*
(Intercept)	0.96	−0.85 – 2.77	1.04	0.300
prime [experimental]	−0.12	−1.37 – 1.14	−0.18	0.855
type [absent]	−1.64	−3.18 – -0.09	−2.08	0.038 *
feet [2feet]	−0.06	−1.61 – 1.50	−0.07	0.943
prime [experimental] × type [absent]	2.37	0.66 – 4.08	2.72	0.007 **
prime [experimental] × feet [2feet]	0.81	−0.91 – 2.52	0.92	0.356
type [absent] × feet [2feet]	1.48	−0.69 – 3.66	1.34	0.181
(prime [experimental] × type [absent]) × feet [2feet]	−2.42	−4.85 – 0.00	−1.96	0.050
Random Effects				
σ^2	119.86			
τ_{00} item	2.78			
τ_{00} subject	14.86			
τ_{11} subject.primeexperimental	0.85			
τ_{11} subject.typeabsent	0.22			
τ_{11} subject.foot2feet	0.32			
ρ_{01} subject.primeexperimental	0.32			
ρ_{01} subject.typeabsent	−0.61			
ρ_{01} subject.foot2feet	0.78			
ICC	0.14			
N subject	27			
N item	96			
Observations	5017			
Marginal R^2 / Conditional R^2	0.003 / 0.146			

Table 2

Model output summary(model) for single-trial ERP data in the **anterior-left ROI 300–400 ms** post-stimulus onset. * <.05, **<.01, *** <.001. Square brackets indicate the factor level comparisons. Reference levels were *control* for the factor *Prime*; Shared for the factor *Type*; 1foot for the factor *Feet.*

	EEG			
Predictors	*Estimates*	*CI*	*t-values*	*p*
(Intercept)	−3.62	−5.57 – -1.67	−3.65	<0.001 ***
prime [experimental]	−1.58	−2.81 – -0.35	−2.52	0.012 *
type [absent]	−1.80	−3.28 – -0.32	−2.39	0.017 *
feet [2feet]	−0.35	−1.82 – 1.12	−0.47	0.637
prime [experimental] × type [absent]	2.82	1.17 – 4.46	3.35	0.001 **
prime [experimental] × feet [2feet]	1.48	−0.17 – 3.13	1.76	0.079
type [absent] × feet [2feet]	2.50	0.45 – 4.56	2.39	0.017 *
(prime [experimental] × type [absent]) × feet [2feet]	−3.39	−5.72 – -1.05	−2.85	0.004 **
Random Effects				
σ^2	110.96			
τ_{00} item	2.32			
τ_{00} subject	19.32			
τ_{11} subject.primeexperimental	1.14			
τ_{11} subject.typeabsent	0.59			
τ_{11} subject.foot2feet	0.37			
ρ_{01} subject.primeexperimental	0.11			
ρ_{01} subject.typeabsent	−0.03			
ρ_{01} subject.foot2feet	−0.65			
N subject	27			
N item	96			
Observations	5017			
Marginal R^2 / Conditional R^2	0.004 / NA			

Table 3

Planned pair-wise comparisons of control-experimental items across the four conditions for the ERP data. (Wald $\chi2$ test). * <.05, **<.001,***<.0001.

200–300 ms centro-parietal ROI

Type	Feet	Value	SE	DF	Chisq	Pr(>Chisq)	ERP effect
shared	1FOOT	0.430	0.554	1	0.603	0.437	–
shared	2FEET	0.823	0.559	1	2.172	0.281	–
absent	1FOOT	−1.618	0.557	1	8.444	**0.015** *	**N400**
absent	2FEET	−1.165	0.556	1	4.385	0.109	–

200–300 ms left-anterior ROI

Type	Feet	Value	SE	DF	Chisq	Pr(>Chisq)	ERP effect
shared	1FOOT	1.081	0.548	1	3.896	0.145	–
shared	2FEET	−0.576	0.551	1	1.089	0.593	–
absent	1FOOT	−1.277	0.550	1	5.392	0.081	–
absent	2FEET	−0.261	0.550	1	0.226	0.634	–

300–400 ms centro-parietal ROI

Type	Feet	Value	SE	DF	Chisq	Pr(>Chisq)	ERP effect
shared	1FOOT	0.115	0.639	1	0.032	0.858	–
shared	2FEET	−0.697	0.644	1	1.172	0.837	–
absent	1FOOT	−2.257	0.642	1	12.359	0.002 **	**N400**
absent	2FEET	−0.644	0.641	1	1.007	0.837	–

300–400 ms left-anterior ROI

Type	Feet	Value	SE	DF	Chisq	Pr(>Chisq)	ERP effect
shared	1FOOT	1.577	0.626	1	6.347	**0.047** *	**P350**
shared	2FEET	0.098	0.630	1	0.024	0.876	
absent	1FOOT	−1.238	0.628	1	3.878	0.147	
absent	2FEET	0.668	0.628	1	1.133	0.574	

400–500 ms centro-parietal ROI

Type	Feet	Value	SE	DF	Chisq	Pr(>Chisq)	ERP effect
shared	1FOOT	0.806	0.719	1	1.256	0.525	–
shared	2FEET	−0.662	0.724	1	0.836	0.525	–
absent	1FOOT	−1.615	0.722	1	5.001	0.101	–
absent	2FEET	−1.147	0.721	1	2.525	0.336	–

400–500 ms anterior-left ROI

Type	Feet	Value	SE	DF	Chisq	Pr(>Chisq)	ERP effect
shared	1FOOT	2.365	0.665	1	12.644	**0.002** **	**P350**
shared	2FEET	0.007	0.667	1	0.0001	0.992	–
absent	1FOOT	−1.302	0.668	1	3.801	0.154	–
absent	2FEET	0.483	0.667	1	0.525	0.938	–

500–600 ms centro-parietal ROI

Type	Feet	Value	SE	DF	Chisq	Pr(>Chisq)	ERP effect
shared	1FOOT	1.048	0.737	1	2.020	0.311	–
shared	2FEET	−0.418	0.743	1	0.316	0.574	–
absent	1FOOT	−1.367	0.740	1	3.415	0.258	–
absent	2FEET	−1.238	0.740	1	2.799	0.283	–

500–600 ms anterior-left ROI

Type	Feet	Value	SE	DF	Chisq	Pr(>Chisq)	ERP effect
shared	1FOOT	2.350	0.698	1	11.326	**0.003** **	**P350**
shared	2FEET	0.266	0.703	1	0.143	1.000	–
absent	1FOOT	−0.965	0.701	1	1.894	0.506	–
absent	2FEET	−0.179	0.701	1	0.065	1.000	–

600–700 ms centro-parietal ROI

Type	Feet	Value	SE	DF	Chisq	Pr(>Chisq)	ERP effect
shared	1FOOT	0.700	0.770	1	0.827	0.830	–
shared	2FEET	−0.528	0.776	1	0.464	0.830	–
absent	1FOOT	−1.023	0.774	1	1.749	0.744	–
absent	2FEET	−0.841	0.773	1	1.183	0.830	–

600–700 ms anterior-left ROI

Type	Feet	Value	SE	DF	Chisq	Pr(>Chisq)	ERP effect
shared	1FOOT	2.001	0.763	1	6.912	**0.034** *	**P350**
shared	2FEET	0.231	0.768	1	0.091	1.000	–
absent	1FOOT	−0.999	0.766	1	1.699	0.577	–
absent	2FEET	0.444	0.765	1	0.337	1.000	–

Appendix C. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bandl.2025.105553.

Data availability

Datasets (behavioural and ERP) and analyses scripts that were generated during the current study are available in the OSF repository, https://osf.io/v6qwa/?view_only=bdeaa0fd9d8b4236a767821345dd794b

References

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language., 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Best, C. T., & Tyler, M. D. (2007). Nonnative and second language speech perception: Commonalities and complementaries. In O. S. Bohn, & M. J. Munro (Eds.), *Language experience in second language speech learning* (pp. 13–34). Amsterdam: John Banjamins.

Boersma, P., & Weenink, D. (2009). *PRAAT: Doing phonetics by computer (ver. 5.1. 0)*. Amsterdam: Institut for Phonetic Sciences.

Bölte, J., & Coenen, E. (2002). Is phonological information mapped onto semantic information in a one-to-one manner? *Brain and Language, 81*(1), 384–397. https://doi.org/10.1006/brln.2001.2532

Comesaña, M., Ferre, P., Romero, J., Guasch, M., Soares, A. P., & Garcia-Chico, T. (2015). Facilitative effect of cognate words vanishes when reducing the orthographic overlap: The role of stimuli list composition. *J Exp Psychol Learn Mem Cogn, 41*(3), 614–635. https://doi.org/10.1037/xlm0000065

Cooper, N. C., Anne, & Wales, Roger (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech, 45* (3), 207–228.

Cutler, A., & Jesse, A. (2021). Word stress in speech perception. In J. S. Pardo, L. C. Nygaard, R. E. Remez, & D. B. Pisoni (Eds.), *The Handbook of Speech Perception* (pp. 239–265). Wiley. https://doi.org/10.1002/9781119184096.ch9.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1983). A language-specific comprehension strategy. *Nature, 304*(5922), 159–160. https://doi.org/10.1038/304159a0

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1989). Limits on bilingualism. *Nature, 340* (6230), 229–230. https://doi.org/10.1038/340229a0

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*(1), 113–121. https://doi.org/10.1037/0096-1523.14.1.113

De Rosario-Martinez, H., Fox, J., Team, R. C., & De Rosario-Martinez, M. H. (2015). Package 'phia'. *CRAN Repos.*

Degani, T., Prior, A., & Hajajra, W. (2018). Cross-language semantic influences in different script bilinguals. *Bilingualism: Language and Cognition, 21*(4), 782–804. https://doi.org/10.1017/S1366728917000311

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

Dijkstra, A. (2007). The multilingual lexicon. In M. G. Gaskell (Ed.), *Handbook of Psycholinguistics* (pp. 251–265). Oxford: Oxford University Press.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language, 62*(3), 284–301. https://doi.org/10.1016/j.jml.2009.12.003

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*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., Wahl, A., Buyuenhuijs, F., et al. (2019). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition, 22*(4), 657–679. https://doi.org/10.1017/S1366728918000287

Dijkstra, T. O. N., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2019). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition, 22*(4), 657–679. https://doi.org/10.1017/S1366728918000287

Domahs, U., Klein, E., Huber, W., & Domahs, F. (2013). Good, bad and ugly word stress – fMRI evidence for foot structure driven processing of prosodic violations. *Brain and Language, 125*(3), 272–282. https://doi.org/10.1016/j.bandl.2013.02.012

Frances, C., Navarra-Barindelli, E., & Martin, C. D. (2021). Inhibitory and facilitatory effects of phonological and orthographic similarity on L2 word recognition across modalities in bilinguals. *Sci Rep, 11*(1), 12812. https://doi.org/10.1038/s41598-021-92259-z

Friedrich, C. K., Kotz, S. A., Friederici, A. D., & Alter, K. (2004). Pitch modulates lexical identification in spoken word recognition: ERP and behavioral evidence. *Cognitive Brain Researcs, 20*(2), 300–308. https://doi.org/10.1016/j.cogbrainres.2004.03.007

Friedrich, C. K., Kotz, S. A., Friederici, A. D., & Gunter, T. C. (2004). ERPs reflect lexical identification in word fragment priming. *Journal of Cognitive Neuroscience, 16*(4), 541–552. https://doi.org/10.1162/089892904323057281

Friedrich, C. K., Lahiri, A., & Eulitz, C. (2008). Neurophysiological evidence for underspecified lexical representations: Asymmetries with word initial variations. *Journal of Experimental Psychology: Human Perception and Performance, 34*(6), 1545–1559. https://doi.org/10.1037/a0012481

Friedrich, C. K., Schild, U., & Röder, B. (2009). Electrophysiological indices of word fragment priming allow characterizing neural stages of speech recognition. *Biological Psychology, 80*(1), 105–113. https://doi.org/10.1016/j.biopsycho.2008.04.012

Fritz, I., Lahiri, A., & Kotzor, S. (2023). Shared loanword recognition in German–English bilinguals: The role of metrical phonology. *Journal of Cognitive Neuroscience, 35*(12), 2028–2048. https://doi.org/10.1162/jocn_a_02063

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x

Jiang, N. (2019). Phonology-based bilingual activation among different-script bilinguals? *Bilingualism: Language and Cognition, 22*(4), 693–694.

Kobor, A., Honbolygo, F., Becker, A. B. C., Schild, U., Csepe, V., & Friedrich, C. K. (2018). ERP evidence for implicit L2 word stress knowledge in listeners of a fixed-stress language. *International Journal of Psychophysiology, 128*, 100–110. https://doi.org/10.1016/j.ijpsycho.2018.04.006

Kondrak, G. (2003). Phonetic alignment and similarity. *Computers and the Humanities, 37* (3), 273–291. https://doi.org/10.1023/A:1025071200644

Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A tale of two positivities and the N400: distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience, 32*(1), 12–35. https://doi.org/10.1162/jocn_a_01465

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

Lahiri, A. (2001). In E. König, & M. Haspelmath (Eds.) (pp. 1347–1367). https://doi.org/10.1515/9783110194265-033

Lijewska, A. (2020). Cognate processing effects in bilingual lexical access. In *Bilingual Lexical Ambiguity Resolution* (pp. 71–95). Cambridge University Press.

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

Luke, S. G. (2016). Evaluating significance in linear mixed-effects models in R. *Behav. Res. Methods, 49*, 1494–1502.

Magne, C., Astesano, C., Aramaki, M., Ystad, S., Kronland-Martinet, R., & Besson, M. (2007). Influence of syllabic lengthening on semantic processing in spoken French: Behavioral and electrophysiological evidence. *Cerebral Cortex, 17*(11), 2659–2668. https://doi.org/10.1093/cercor/bhl174

Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience, 23*(2), 294–305. https://doi.org/10.1162/jocn.2010.21413

Marslen-Wilson, W. (1990). Activation, competition, and frequency in lexical access. In *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148–172). Cambridge, MA, US: The MIT Press.

Mitra, A., & Dutta, I. (2023). Mixed language processing increases cross-language phonetic transfer in Bengali–English bilinguals. *Bilingualism: Language and Cognition, 26*, 896–909. https://doi.org/10.1017/S1366728923000159

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Mehler, J., Dommergues, J. Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior, 20*(3), 298–305. https://doi.org/10.1016/S0022-5371(81)90450-3

Moon, J., & Jiang, N. (2012). Non-selective lexical access in different-script bilinguals. *Bilingualism: Language and Cognition, 15*(1), 173–180. https://doi.org/10.1017/S1366728911000022

Mulder, K., Wloch, L., Boves, L., ten Bosch, L., & Ernestus, M. (2021). Cognate status modulates the comprehension of isolated reduced forms. *Language, Cognition and Neuroscience, 37*(5), 576–614. https://doi.org/10.1080/23273798.2021.1995611

Muntendam, A., van Rijswijk, R., Severijnen, G., & Dijkstra, T. (2022). The role of stress position in bilingual auditory word recognition: Cognate processing in turkish and dutch. *Bilingualism: Language and Cognition, 25*(4), 679–690. https://doi.org/10.1017/s1366728922000037

Nunez, M.D., Nunez, P.L., Srinivasan, R., Ombao, H., Linquist, M., Thompson, W., & Aston, J. (2016). Electroencephalography (EEG): Neurophysics, experimental methods, and signal processing. In Ombao, H., Linquist, M., Thompson, W. & Aston, J. (Eds.) Handbook of Neuroimaging Data Analysis (pp. 175- 197), Chapman & Hall/CRC. Advance online publication. doi: 10.13140/rg.2.2.12706.63687.

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

Sánchez-Casas, R., & García-Albea, J. E. (2005). The representation of cognate and noncognate words in bilingual memory. In J. F. Kroll, & A. M. B. D. Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 226–250). Oxford University Press.

Scharinger, M., & Felder, V. (2011). ERP signatures of cross-modal semantic fragment priming: Early context effects in speech perception. *Int. J. Psychophysiol., 80*(1), 19–27. https://doi.org/10.1016/j.ijpsycho.2011.01.006

Schepens, J., Dijkstra, T., Grootjen, F., & Van Heuven, W. J. (2013). Cross-language distributions of high frequency and phonetically similar cognates. *PLoS One, 8*(5), Article e63006.

Shetty, A., Hebbar, S. P., Shenoy, R., et al. (2022). A prime-masked ERP investigation on phonology in visual word processing among bilingual speakers of alphasyllabic and alphabetic orthographies. *Sci Rep, 12*, 9870. https://doi.org/10.1038/s41598-022-13654-8

Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language, 45*(3), 412–432. https://doi.org/10.1006/jmla.2000.2783

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.

Tsang, Y. K., & Zou, Y. (2022). An ERP megastudy of Chinese word recognition. *Psychophysiology, 59*(11), Article e14111. https://doi.org/10.1111/psyp.14111

van Donselaar, W., Koster, M., & Cutler, A. (2005). Exploring the role of lexical stress in lexical recognition. *Q J Exp Psychol A, 58*(2), 251–273. https://doi.org/10.1080/02724980343000927

van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Q J Exp Psychol, 67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Voga, M., & Grainger, J. (2007). Cognate status and cross-script translation priming. *Memory & Cognition, 35*(5), 938–952. https://doi.org/10.3758/BF03193467

Zwitserlood, P. (1996). Form priming. *Language and Cognitive Processes, 11*(6), 589–596. https://doi.org/10.1080/016909696386980

Zwitserlood, P., Schriefers, H., Lahiri, A., & van Donselaar, W. (1993). The role of syllables in the perception of spoken Dutch. *J. Exp. Psychol. Learn. Mem. Cogn., 19*(2), 260–271. https://doi.org/10.1037/0278-7393.19.2.260