

Shared loanword recognition in German-English bilinguals: The role of metrical phonology

Isabella Fritz,¹ Aditi Lahiri¹ & Sandra Kotzor^{1,2}

¹Language and Brain Laboratory, Faculty of Linguistics, Philology and Phonetics, University of Oxford

²Department of British and American Studies, Ludwig-Maximilians-Universität München

Abstract:

The role of phonology in bilingual word recognition has focussed on a phonemic level especially in the recognition of cognates. In this study, we examined differences in metrical structure to test whether L1 (first-language) metrical structure influences the processing of L2 (second-language) words. For that, we used words of Romance origin (e.g., *reptile*, *signal*) which both German and English have borrowed extensively. However, the existing metrical patterns are not identical nor are the borrowed vocabularies the same. Rather those identical words differ systematically in their foot structure.

We conducted a cross-modal form fragment priming EEG experiment (auditory-visual) with German native speakers who were highly proficient in English. Both behavioural and ERP results showed an effect of the native phonology and the loan status, i.e., whether the loan exists only in the speaker's L2 or is shared across languages. Priming effects (RTs) were largest for non-shared loanwords indicating some interference from German (L1). This was also evident in a reduced N400 but only if the metrical structure aligned with German patterns for Germanic words, i.e., two light syllables as in *pigeon*. If the words exist in both languages, metrical structure also mattered shown by the modulation of different ERP components across conditions. Overall, our study indicates that metrical phonology plays a role in loanword processing. Our data shows that the more similar a word is in terms of its metrical phonology across L1 and L2, the more effortful the processing of a word within a priming paradigm indicating interference from the L1 phonology.

Keywords: phonology, foot structure, processing, L2

Funding: AHRC Standard Grant AH/S011323/1 (PIs A. Lahiri & S. Kotzor);

Engineering and Physical Sciences Research Council UKEP/X026035/1 (PI; A. Lahiri)

Introduction

For decades, psycholinguists have been investigating possible units involved in speech segmentation. On a sub-lexical level, the syllable and in particular the role of lexical stress has received much attention (see Cutler & Jesse, 2021 for an overview). Indeed, the syllable is a crucial building block in language processing and in addition they are easy to understand conceptually. Syllables, however, come in different shapes and sizes and the combination of which provides another layer of structure, namely the metrical foot. Thus, the metrical foot is the grouping of syllables into a larger constituent (cf. Lahiri, 2001). Experimental findings for Germanic languages (English, Dutch and German) show that initial strong syllables mark the beginnings of words and are used for the segmentation of words out of a continuous speech stream (Cutler and Norris, 1988). Moreover, stressed syllables play a more important role in word recognition than syllables that don't bear stress (Grosjean and Gee, 1987). These findings provide indirect evidence for the metrical foot being taken into account during processing. Using metrical violation paradigms, a few more recent studies tested whether the metrical structure impacts lexical retrieval and processing directly. Findings across multiple methodologies (i.e., behavioural, EEG and fMRI) all indicate a psychological reality for the concept of metrical feet (Domahs et al., 2013, Marie et al., 2011; Magne et al., 2007).

In the present study, we test how phonological differences in shared loanwords (i.e., L1 German and L2 English) impact processing. As a unit for assessing phonological differences and similarities, we used the metrical foot which best illustrates how identical Romance loans have been adapted differently into German and English. Thus, we also examine those loans from a historical perspective for a better understanding of the loans' integration into the metrical structure of English and German.

In a contact situation, native speakers of one language inevitably borrow from a neighbouring language. Due to centuries of close contact, Romance words (that is words from Latin or the descendants of Latin such as Old French) have been incorporated into all Germanic languages in different periods. Our focus is on the present-day Germanic languages, particularly English and German, and the effect Romance loans have had in these two languages. Even if the languages are related, absorbing loans can be difficult. Often, words with different segmental patterns are difficult to perceive and are invariably pronounced with some variation. For instance, the surname of the famous Dutch painter Vincent van Gogh, is pronounced in several different ways: original Dutch [xɔx], [gou] in American English, [gɔf] in British English, [gɔx] in German. These different pronunciations reflect the respective native phonological systems as well as letter-to-sound correspondences.

Such differences are evident not just in segmental phonology but also with respect to metrical structure and word stress. Inherited words, that is Germanic words which have remained in both English and German, always share the same stress pattern: e.g., G **T**ochter ~ E **d**aughter. Under normal circumstances, a borrowed word conforms to the native phonology of the time of borrowing (Kennard & Lahiri 2020). Our research focuses on Romance loans borrowed into German and English. Many Romance loans have been borrowed at different periods in the two languages and their stress patterns can differ. Table 1 provides some examples of similarities and differences with respect to which syllables in a word bear stress.

Table 1. Examples of disyllabic ($\sigma \sigma$) and trisyllabic ($\sigma \sigma \sigma$) words with similar and dissimilar stress patterns in English and German. Stress on syllables is marked with a diacritic, e.g., $\acute{\sigma}$

English	Syllables	German	Syllables	Stress comparison
1. pánic	$\acute{\sigma} \sigma$	Pánik	$\acute{\sigma} \sigma$	same
2. vendétta	$\sigma \acute{\sigma} \sigma$	Vendétta	$\sigma \acute{\sigma} \sigma$	same
3. inféction	$\sigma \acute{\sigma} \sigma$	Infection	$\sigma \sigma \acute{\sigma}$	different
4. crócodile	$\acute{\sigma} \sigma \sigma$	Krokodíl	$\sigma \sigma \acute{\sigma}$	different

The words in the first two rows in Table 1 have identical stress patterns. But in rows 3 and 4 the words bear stress on different syllables. Our focus is on the processing of words with different metrical patterns with respect to L2 shared loan processing.

From a phonological perspective, word stress is not merely emphasis on a particular syllable. Rather, word stress is determined on the basis of organising syllables (groups of vowels and consonants) into larger units, based on the weight of syllables, 'heavy' or 'light'. A heavy syllable has either a long vowel or a short vowel followed by one or two consonants. In English and German, either a single heavy syllable or two light syllables make up a foot. Thus, English *móral* has one foot with two light syllables, and English *réptile* has two separate feet each on a heavy syllable.

Syllables which end with a consonant (*signal* ['sig.nəl] or contain a long vowel or diphthong (*silence* ['saɪ.ləns]) behave in the same way and are considered to be 'heavy' (designated in the figures with H). The first syllable of *moral* ['mɔ.rəl], however, contains a short vowel and does not end in a consonant and is therefore 'light' (L). A disyllabic word can have two stresses (i.e., two feet) or a single main stress (i.e., one foot). Consequently, although the disyllabic English words

moral and *reptile* both bear main stress on the initial syllable, they have different foot structures because the weight of the initial syllable differs. If we now compare the words *moral* ['mɔ.rəl] and *signal* ['sɪg.nəl], both only have a single foot, but the foot in *moral* is built on two light syllables, while the heavy first syllable in *signal* is sufficient as a foot on its own. Syllable weight is indicated by moras (μ), where a light syllable has a single mora and a heavy syllable carries two moras ($\mu\mu$). In Figure 1a, we provide three examples of disyllabic English words.

Figure 1 Metrical patterns in English (a) and German (b) for disyllabic nouns of Romance origin

a.)	<i>moral</i>	<i>signal</i>	<i>reptile</i>
	X	X	X
	(x)	(x)	(x)
	μ μ	$\mu\mu$ μ	$\mu\mu$ $\mu\mu$
	σ σ	σ σ	σ σ
	[mɔ rəl]	[sɪg nəl]	[rɛp tʰaɪl]
	L L	H L	H H
	one foot	one foot	two feet
b.)	<i>Moral</i>	<i>Signal</i>	<i>Reptil</i>
	X	X	X
	(x)	(x) (x)	(x) (x)
	μ $\mu\mu$	$\mu\mu$ $\mu\mu$	$\mu\mu$ $\mu\mu$
	σ σ	σ σ	σ σ
	[mo ʁa:l]	[zɪg na:l]	[rɛp tʰi:l]
	L H	H H	H H
	one foot	two feet	two feet

Note that the weight of a syllable does not automatically lead to the presence or absence of stress. It is not the syllable, but the foot built on the syllables which predicts the position of stress. The preferred foot in English and German is a moraic trochee which is made up of two light syllables (*moral*) or a single initial heavy syllable (*signal*). The first syllable in *moral* has one mora (short vowel and no consonant closing the syllable) and the second syllable has a reduced vowel. Thus, a single trochaic foot can be built on both syllables with the head of the foot on the left (indicated by 'x' in Figure 1). The word *reptile* has two heavy syllables each forming a foot on its own, but the leftmost carries main stress, indicated by a capital X. Although not relevant for our purposes, the final syllables of English nouns are disregarded for main stress regardless of their weight (cf. Hayes 1995).

German too has borrowed these three words, but their metrical patterns differ because vowel length has been adapted differently in German. All three example words consist of bimoraic heavy final syllables and, according to German metrical rules, main stress falls on the rightmost foot (cf. Wiese, 2000; Zonneveld et al., 1999). Consequently, German *Moral*, *Signal* and *Reptil* are all words with main stress on the final syllable (see Figure 1b).

The focus of this paper is on the processing of Romance loans in English by native German speakers who are highly proficient L2 English speakers. The central query is to what extent the metrical structure of the L1 phonology influences their processing of spoken English words. To investigate this issue, we have chosen sets of words with varying metrical patterns and different degrees of overlap. Investigating the processing of these shared Romance loanwords also enabled us to test more generally how phonological overlap of shared loanwords (often referred to as 'cognates' in the L2 literature) impacts processing. Thus, the present study contributes to both the L2 literature on cognate processing and aims to extend previous psycholinguistic research on the metrical foot as potential processing unit which, until now, is mostly based on experimental evidence from studies using misplaced stress (e.g., Domahs et al., 2008, Henrich et al., 2014). Before discussing English and German foot structure of loanwords in detail, we first touch upon the discussion in the literature regarding L2 processing of loans.

L2 Shared Loanword Recognition

In this section, we discuss the processing of loanwords by L2 speakers as well as the notion of the term cognate in the psycholinguistic literature. Within linguistics, cognate only refers to those words that are inherited from the same ancestor; that is, they are descendants from the same source. Thus, none of the words used in this study are real ‘cognates’ because we focus only on loans. Germanic nouns in English and German have their main stress inevitably on the same position (*‘daughter ~ ‘Tochter; ‘Monday ~ ‘Montag*). The L2 and multilingual literature, however, conflates the terms **loanword** and **cognate** and uses **cognate** as an umbrella term for translation equivalents that are similar in meaning, sound and orthography (cf. Dijkstra, Grainger, & van Heuven (1999). Thus, the literature would assume that the words *reptile* and *Reptil* are ‘cognates’ whereas they have not been inherited from older Germanic but instead have been borrowed from Latin at different periods of time.

From a psycholinguistic perspective, many studies have investigated the processing of shared loanwords and cognates. The consensus is that if a similar word form (cognate or loanword) exists in both languages, lexical access will be facilitated for a bilingual. A large number of studies have found that this facilitation effect is modulated (and sometimes even reversed) by multiple factors including the degree of orthographic and phonological overlap and whether participants are tested in their L1 or L2 (see for Lijewska (2020) an overview). Generally, those facilitation effects are attributed to an integrated mental lexicon assuming that lexical access is language non-selective (e.g., Dijkstra & van Heuven, 2002). Cognate facilitation effects, however, are predominantly based on studies on visual word processing. Most studies do not consider phonological issues, but even those that investigated the influence of phonological similarity of cognates on word processing, the stimuli presentation was often through the visual modality (Comesana et al., 2015; Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010). Only recently, studies have begun to present stimuli auditorily to address questions about ‘cognate’ processing in relation to their degree of phonological overlap (Frances, Navarra-Barindelli, & Martin, 2021; Mulder, Wloch, Boves, ten Bosch, & Ernestus, 2021; Muntendam, van Rijswijk, Severijnen, & Dijkstra, 2022). Importantly, those studies did not employ a priming paradigm but lexical decision tasks without priming.

The research most relevant to the current study is by Muntendam et al. (2022) who investigated whether differences in stress position influence ‘cognate’ recognition. In particular, they were interested whether cross-linguistic differences in stress placement, which reduce phonological similarity, still lead to co-activation and thereby cognate facilitation. Stimuli were disyllabic words extant in both Dutch and Turkish with either matching or mismatching stress positions. That is, final or penult stress in both languages or penult in Dutch and final in Turkish. ‘Cognate’ facilitation effects were found when participants were tested in their dominant language (i.e., Dutch) whilst inhibition arose in their weaker language. In terms of stress position, in both languages, items with final stress were processed faster than when stress was placed on the penultimate syllable regardless of whether stress matched in the language not used in the task. Moreover, ‘cognates’ with penultimate stress were processed slower than non-cognates.

The authors argue that this pattern stems from differences in how competition unfolds for words shared by several languages. ‘Cognates’ with initial stress were co-activated from the onset and throughout because of the overlap in segmental information. For the penult-final stress combination, initial competition was weak because the overlap was smaller. Moreover, previous studies suggested that initial stress can eliminate competing candidates beginning with an unstressed syllable (e.g., Reinisch, Jesse, & McQueen, 2010). Finally, greater facilitation for the condition with final stress in both languages was assumed to be due to the overlap between Dutch and Turkish and because default stress in Turkish falls on the final syllable (cf. Cutler & Jesse, 2021). Although this study showed that stress placement influences ‘cognate’ co-activation, it did not consider whether those facilitation/inhibition effects arose exclusively from differences in stress placement.

The Present Study

In this study, we tested the extent to which the existence of a similar loanword in German (L1) affects processing of the corresponding English (L2) word. Since loans differ in stress patterns across languages, a related question is what role the metrical system plays in the activation of the L2 English lexicon. To address our research questions, we systematically varied the metrical structure of English disyllabic loans in relation to their German counterparts. This resulted in three main conditions: English loans which do not exist in German (NONEX), English and German similar loans with two syllables (DISYLL) and those which are disyllabic in English but trisyllabic in German (TRISYLL). All English words were disyllabic nouns with initial stress. For each of these three types of loans we included one condition where the loans in English consisted of *one foot* (1FOOT) and a second condition with loans made up of *two feet* (2FEET) resulting in six experimental conditions (see Table 2).

Table 2. Example stimulus for each of the six experimental conditions. The English word was presented in the experiment. Table also shows the German counterparts.

Condition	Foot	English	German
NONEX	1FOOT	pigeon ['pɪdʒ(ɪ)n]	-
NONEX	2FEET	curfew ['kəːfjuː]	-
DISYLL	1FOOT	signal ['sɪgnəl]	Signal [zɪ'ɡnaːl]
DISYLL	2FEET	reptile ['rɛp.taɪl]	Reptil [rɛp'tiːl]
TRISYLL	1FOOT	carrot ['kærət]	Karotte [kʰa'rɔtə]
TRISYLL	2FEET	nomad ['nəʊ.məd]	Nomade [no'maːdə]

We conducted a cross-modal ERP study with fragment priming where participants heard the first syllable of the target, which always bore stress, followed by the visual full-word target. Different types of priming paradigms are widely used in psycho- and neurolinguistics when investigating lexical access and lexical representation. Priming refers to the phenomenon observed when a related item (i.e., prime) preceding a target, results in either facilitation or interference in the processing of the target item compared to an unrelated prime preceding the same target item. The relationship of the prime and the target can vary in terms of form or meaning (e.g., Bölte & Coenen, 2002; Zwitserlood, 1996). We used a cross-modal paradigm where the fragment prime was presented auditorily and followed by a visual target.

Word onset fragments presented auditorily have been known to facilitate visual targets. For example, in German ['myn] primes Münze ('coin') ['myntsə] and thus leads to faster RTs compared to Münze being preceded by a semantically or phonologically unrelated prime (Friedrich, Lahiri, & Eulitz, 2008; Marslen-Wilson, 1990; Scharinger & Felder, 2011; Schutz, Schendzielarz, Zwitserlood, & Vorberg, 2007). Fragment priming studies have also been used to investigate the representation of lexical stress and its activation during word recognition. A number of studies across multiple languages found that lexical access is facilitated for words preceded by stress-congruent primes (cf. for English: Cooper et al. (2002). In addition to the facilitation effect, van Donselaar, Koster, and Cutler (2005) also found an inhibition effect for words preceded by stress-incongruent primes (see for Spanish: Soto-Faraco, Sebastián-Gallés, & Cutler, 2001).

This task has also been successfully used in EEG experiments. Friedrich, Kotz, Friederici, and Alter (2004) employed a fragment priming study using EEG to test whether the online processing of target words differs depending on whether they are preceded by stress-matching or stress-mismatching primes. If the stress pattern of prime and target matched, e.g. ['re] followed by ['re:g] (*Regel* 'rule'), the ERP amplitude of the P350 was reduced compared to stress mismatching prime-target pairs, e.g. ['re] followed by [re'ga:l] (*Regal* 'shelf'). The P350 component is thought to reflect the phonological mapping between prime fragment and target words with a more negative deflection for phonologically matching fragments compared to phonologically unrelated fragments. This component has been found in a number of fragment priming studies and usually shows a left-frontal distribution between 300-400ms after target-word onset (Friedrich, Kotz, Friederici, & Gunter, 2004; Friedrich, Schild, & Röder, 2009; Scharinger & Felder, 2011). The only ERP fragment priming study with L2 speakers comparing stress-matching and stress-

mismatching prime target pairs revealed a prolonged P350 lasting from 300ms to the end of the epoch, i.e. 700ms (Kobor et al., 2018).

In addition to the P350 component, cross-modal fragment priming studies also elicited negativities usually within 300 and 600ms post target onset with a centro-parietal distribution. The function of this component is still debated and some authors refer to this component as central negativity. Some studies relate it to the N400 component which is sought to index the ease of lexical access and integration in semantic priming studies (e.g., Brown & Hagoort, 1993; Chwilla & Kolk, 2005). Thus, a target word that is preceded by a phonologically related word is easier to access and integrate and thus elicits a less negative deflection compared to a target word preceded by an unrelated prime (e.g., Scharinger & Felder, 2011). Related to the phonological overlap, Friedrich, Kotz, Friederici, and Alter (2004); Friedrich, Kotz, Friederici, and Gunter (2004) further argue that this phonological N400 indexes phonological expectancies based on the prime that activates the cohort (see also Praamstra, Meyer, & Levelt, 1994 for a similar view on the phonological N400). Based on previous studies, we assume that the P350 exclusively reflects matching on a phonological level whilst the phonological N400 (or central negativity) also reflects lexical access and integration processes.

Taken together, this experimental paradigm allowed us to test whether the fragment (i.e. the initial syllable cut out from the target word) in an L2 (English) also activates the L1 (German) loanword and whether (1) this depends on the degree of similarity in the metrical structure across the two languages and/or (2) whether the presented fragment stems from a word with one or two feet. Thus, although the participants hear a syllable, this method really taps into the processing of the metrical structure in an L2. Crucially, it allowed us to extend previous research on the role of syllables (e.g., Grainger & Jacobs, 1996) and lexical stress (cf. Cutler & Jesse, 2021) on L2 word recognition by bringing in the concept of the metrical foot as a unit in speech processing.

Predictions

Previous studies on ‘cognate’ effects based their predictions on the phonological similarity between the items in the two languages. However, how similarity was assessed differed across studies (Comesana et al., 2012; Dijkstra et al., 2010; Frances et al., 2021; Muntendam et al., 2022). Some studies calculated the similarity of two strings by using the Levenshtein Distance (Levenshtein, 1966; used, for instance, in Mulder et al., 2022) In this approach, the underlying assumption is that each change equally affects the similarity status and that phonological similarity is limited to a phonemic level. Other studies used a more sophisticated approach by calculating similarity using the ALINE algorithm (Kondrak, 2003) which is based on feature-level phonetic similarity (e.g., Frances et al., 2021). This calculation is based on different categories of phonetic segments which are represented as vectors of the following feature values: prosodic (syllabic), place, manner, phonation, vowel length and colour. Although, the ALINE algorithm calculates a much more detailed comparison, it was developed for reconstructing historical linguistic relationships (Downey, Hallmark, Cox, Norquest, & Lansing, 2008) and thus it is unclear whether the scoring scheme reflects differences which are relevant for language processing. The second approach consists of bilinguals rating how similar a pair of words are in their pronunciation. Based on these similarity ratings, ‘cognates’ were then classified as being either of high or low phonological similarity (e.g., Comesana et al., 2012; Frances et al., 2021) or the similarity ratings were added as a factor in the statistical analyses (Dijkstra et al., 2010).

Compared to previous studies, we based our predictions on whether the English word also exists in German and the second, if there is a German counterpart, how similar the word is in terms of its metrical structure. Recall, all English words in the experiment consisted of two syllables with initial stress. In each condition, the words were divided into two categories depending on their foot structure – either with a single foot (1FOOT) or two feet (2FEET). Words with a single foot only carry main stress while those with two feet can bear a main stress and a secondary stress. Additionally, the conditions differed with regard to the corresponding German words as seen in Table 2.

First, we predict that our highly proficient L2 speakers will have no difficulty in accessing English words and thus we expect priming effects in the behavioural data, i.e., faster RTs for items preceded by matching primes compared to unrelated control primes. Nevertheless, we predict that the participants' native phonology will influence their processing to some degree. If they have no competition in their native language (NONEX, Table 2), that is if the words do not exist in German and thus there is no competition from the L1 item, we predict strongest priming effects. Moreover, we also expect the metrical structure to play a significant role for words not existing in German. If the word consists of a single left-headed 1FOOT, as in *pigeon*, this fits best and does not conflict with existing metrical patterns in German. Disyllabic trochaic words ending in a light syllable are the norm (cf. Wiese, 2000). If, however, the word consists of 2FEET, then the tendency in German would be to stress the second foot, which is the second syllable rather than the first. Consequently, words like *curfew*, are expected to cause more processing effort for Germans. Although final stress only holds for loanwords and not for inherited words (i.e. Germanic words such as *Demut* "humility" or *Schicksal* "fate"), overall this tendency still holds as illustrated by a CELEX search: Disyllabic German words comprised of two feet (regardless of their origin) are still predominantly stressed on the final syllable (195 instances for finally stressed words vs. 64 for words with initial stress). In an integrated bilingual lexicon (e.g., Dijkstra & van Heuven, 2002), if the word exists in both languages and their structures differ across the two languages and would thus be treated differently by English and German phonology, we predict reduced facilitation (i.e., smaller priming effects) as this difference in the patterns existing in the two languages would need to be resolved.

For words that do exist in German, we predict that the different types of foot structure will again be reflected in the processing of the English words. If loanwords are disyllabic in both languages, they still have different foot structures across the languages (cf. DISYLL in Table 2). We divided the words in two categories, either they consist of 2FEET (*reptile*) or 1FOOT (*signal*) in English. The German counterparts, in contrast, always have a super-heavy second syllable (Zonneveld et al., 1999) which is either closed with two coda consonants as in *Aspekt* 'aspect' [as'pekt] or with a single coda consonant but a long vowel (*Signal* 'signal' [sig'na:l]). Only one of the stimuli consisted of an open final syllable with a long vowel (i.e., *Klischee* 'cliche' [kli'ʃe:]). Consequently, all German counter-parts were stressed on the final syllable. If native German phonology plays a role, we would expect no difference between those two sets but, overall, they will be more difficult to process than words without a similar counterpart in German. As for the conditions with trisyllabic German counterparts (TRISYLL in Table 2), the English words are of the same structure as those in Condition B, i.e. 2FEET or 1FOOT, but the corresponding German words are all trisyllabic with non-initial stress. Comparing DISYLL and TRISYLL conditions, words in the DISYLL conditions are closer to their German counterparts in terms of the number of syllables. Thus, we predict that the DISYLL are the most difficult to process due to direct conflict. Our rationale here is based on previous psycholinguistic studies using interference paradigms, for example picture naming tasks (e.g., Hoshino & Thierry, 2011) or a word naming tasks (e.g., Jared & Kroll, 2001) where phonological similarity of the non-target language interfered with naming latencies. Most notably, Jared & Kroll (2001) showed that spelling to sound correspondences from both languages are activated also in a non-target language.

As for the ERP data, based on earlier fragment priming studies, we expect modulations of the P350 and N400 components across the conditions reflecting different degrees of phonological integration.

Methods

Participants

24 German native speakers from either Austria or Germany participated in the EEG study which is a comparable number of participants to previous ERP fragment priming studies (cf. Scharinger

& Felder, 2011: 18 participants; Friederich et al., 2009: 22 participants). The study was run in the Language and Brain Laboratory at the University of Oxford and participants received monetary compensation for their time. One participant was excluded from the analysis due to excessive artifacts (fewer than 20 trials in multiple conditions). Thus, 23 participants were included in the analyses (mean age = 26.6, SD = 5.6, 22 right-handed). These were all German native speakers who grew up monolingually and have normal or corrected-to-normal vision. They did not report any hearing or other cognitive impairments. Note that we kept left-handed participants (N = 1) in the analyses to better represent the wider population (cf. Willems, der Haegen, Fisher, & Francks, 2014). All participants gave written informed consent and the study was reviewed by the Central University Research Ethics Committee at the University of OxfordCUREC (ethics approval code: R60564/RE002).

Further selection criteria included the following: Participants had only limited or no knowledge of French, Italian or Spanish since this could influence the processing of our stimuli which were all Romance loanwords. In addition, we made sure that none of the participants knew any of the Scandinavian languages in which most of the loanwords exist as well, adapted to their respective phonologies. We assessed the L2 participants' language history in greater depth by using parts of the Language Experience and Proficiency Questionnaire (LEAP-Q) (Marian, Blumenfeld, & Kaushanskaya, 2007) as well as administering the English version of LexTale (Lemhofer & Broersma, 2012), a lexical decision task which has been found to be a reliable indicator of language proficiency. The mean LexTale score for the German participants was 90.5 % (min = 73.8%, max = 100%, SD = 6.0). According to Lemhofer and Broersma (2012) this score (i.e., 80-100%) translates to an English proficiency score at C1/C2 level in the Common European Framework which refers to upper (C2) and lower (C1) advanced/proficient speakers of English.

The L2 participants' English experience, as reported in the language questionnaire, can be found in Table 3. All participants lived in England at the time of testing and predominately used English in their every-day lives.

Table 3. Summary of participants' language history background questionnaire based on Marian et al. (2007).

Question	Means (SD)
<i>Percentage of time you are currently and on average exposed to each of the languages you listed (percentages should add up to 100%).¹</i>	
English	73.3 %
German	23.0 %
Other	4.0%
<i>Age when you</i>	
began acquiring English	8.8 (2.4) years
became conversationally fluent in English.	15.5 (3.7)
began reading in English.	12.4 (4.2)
became fluent reading in English.	15.9 (4.3)
 Amount of time you have spent in an English-speaking environment (specify years, months)	 4,3 (max = 17 years, min = 4 months)

Materials

In total, our stimulus set consisted of 144 disyllabic monomorphemic English loanwords exclusively consisting of nouns (French and Latin) divided into three main conditions (Table 2): English loans (A) which do not exist in German, (B) which are similar in German with two syllables and (C) which are disyllabic in English but trisyllabic in German. All loans consisted of either one or two feet (see Table 2) resulting in six experimental conditions with 24 items each.

¹ One participant's language use did not add up to 100% but 110% which results in a discrepancy in the overall percentage.

The foot structure for each condition in English plus their German counterparts can be found in Table 2.

All conditions were matched on frequency across conditions using the Zipf values from the SUBTLEX-UK corpus which were introduced by van Heuven, Mandera, Keuleers, and Brysbaert (2014). The Zipf-scale is a standardised word frequency scale with values ranging from 1-7 where the middle of the scale separates the low-frequency words from the high-frequency ones (see Table 5). We then created 144 non-words which were closely modelled after real Romance loanwords in English. All nonwords were orthographically and phonotactically legal in English and were neither homographic nor homophonic with any existing words in German. Moreover, we created the nonwords so that half consist of 1FOOT (e. g., **tilder*, **pollet*) and the other half of 2FEET (e.g., **lampile*, **wellow*).

Of the 288 control primes (one for each nonword and one for each target), half were also Romance loans (e.g., *stipend*, *target*) whilst the other half were of Germanic origin (e.g., *heaven*, *shoulder*) and across all primes about 50% also had counterparts in Modern German (e.g., *crisis/Krise*, *muscle/Muskel*). For words of Germanic origin, this means that the relationship between the cognates (English and German) is semantically transparent from a synchronic perspective alone. All control primes were disyllabic, monomorphemic words with initial stress mirroring the target items. All items (primes and targets) were recorded by a male speaker of Southern British English in a sound-proof booth and the recording was digitised at 44.1 kHz (16 bit, mono). Later, the disyllabic words were cut in Praat (Boersma & Weenink, 2009) at the offset of the first syllable (e.g., *coral* ['kɒ.rəl], *mattress* ['mæ.trəs]). As a last step, all stimuli were normalised in Praat so that all have a mean amplitude of 70 dB. We then paired each of the targets (word and nonword targets) with an experimental prime and a control prime. The experimental prime consisted of the first syllable cut from the recording of the target word whilst the control prime was a syllable phonologically unrelated to the target (see Table 4 for the experimental design).

Table 4. Example stimulus illustrating the target prime relationship for a two feet word and a two feet non-word.

<i>Prime</i>	<i>Fragment</i>	<i>Target</i>
<i>Experimental</i>	['rɛp]	reptile
<i>Control</i>	['mas] (from <i>mascot</i>)	reptile
<i>Experimental</i>	['fab]	fabrile
<i>Control</i>	['wis] (from <i>whisper</i>)	fabrile

Table 5. Norming values for frequencies and number of phonemes (Nphoneme) and graphemes (Ngrapheme) across the conditions.

<i>Condition</i>	<i>Foot structure</i>	<i>Zipf Value</i>	<i>Nphoneme</i>	<i>Ngrapheme</i>
NONEX	1FOOT	3.88	5.09	6.27
	2FEET	3.92	5.42	6.42
DISYLL	1FOOT	3.86	4.96	5.58
	2FEET	3.91	5.58	6.38
TRISYLL	1FOOT	3.88	5.19	5.90
	2FEET	3.57	5.52	6.78

In total, the experiment consisted of 576 trials divided into two separate blocks. Each block contained all targets; half preceded by an experimental prime, the other half by a control prime. The same was true for non-words. Each list was pseudorandomised with no more than two trials from the same condition following each other and no more than five nonwords presented in a row. Across the lists, the number of Germanic/Romance primes, the number of one/two foot nonwords and the number of items per condition were divided equally. The order of the blocks was counterbalanced across participants. Thus, each participant saw each target item twice but the order of whether they saw the control or the experimental prime first was counterbalanced across participants.

After reading the information sheet and giving consent, participants filled in a questionnaire on their handedness and medical history. They then completed the LexTale online and completed the language history questionnaire. The experiment itself took place in a sound-attenuated booth where participants were seated approximately 100 cm away from the screen. All items were shown in lowercase letters in a white 30-point Arial font against a dark grey background. Each item was preceded by a fixation cross which was shown for 500ms followed by the fragment prime. The ISI between the prime and the target item was 350ms. The target was presented for 500ms and participants had a total of 1500ms to respond to the target before the next trial. Participants responded using a game controller. They were told they would hear fragments of a word through headphones followed by a word appearing on the screen. Their task was to indicate as quickly and as accurately as possible whether the word 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. The experiment started with a practice block consisting of 10 trials. 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, took about 30 mins to complete.

Data analysis

Behavioural analysis

We analysed behavioural responses (RTs) by fitting linear mixed-effects models in R by using the `lmer` function of the `lme4` package (Bates, Kliegl, Vasishth, & Baayen, 2015). We obtained p-values when fitting GLMs with `lmer` by using the Satterthwaite approximation which is implemented in the `lmerTest` package (Kuznetsova, Brockhoff, & Christensen, 2017). First, we tested whether there was a significant priming effect (i.e., whether RTs differed significantly between the prime and control) in the six conditions. For that, we fitted a model including log-transformed RTs as dependent variable and *Condition* and *Prime* (experimental/control) as fixed factors. For the random effects structure, we treated subject and item as random effects. For each model, we started with the maximal random effects structure which was then simplified by removing main effects in the order of least variance explained until the model converged. We then ran a pair-wise comparison on the given model using the `testInteractions` function which is implemented in the `phia` package (De Rosario-Martinez, Fox, Team, & De Rosario-Martinez, 2015).

In a next step, we tested whether number of feet (*Foot*) and *Loan Type* had an impact on the RTs. For that, we included *Prime* (control/experimental) and *Loan Type* (DISYLL, TRISYLL, NONEX) and *Foot* (1FOOT, 2FEET) as fixed factors. All factors were dummy coded and the baseline was set to the control condition for the factor *Prime* and NONEX for the *Loan Type* factor and 1FOOT for the *Foot* factor. We then tested for main effects and interactions by a step-wise removal of first the interaction terms followed by main effect using the `anova()` function. Prior to fitting the RT models, we removed data points with RTs that were ± 3 standard deviations away from a subject's mean which led to the exclusion of 214 trials (i.e. 1.7 %). We further excluded all incorrect responses and nonwords.

2.3.2. EEG pre-processing & analysis

Continuous EEG was recorded with 62 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 2 500 Hz. Impedance was kept below 20 kOhm wherever possible. All EEG channels were then re-referenced offline to the linked mastoids (TP9 and TP10) which were placed outside of the elastic cap on the left and right mastoid respectively since the cap positions of TP9/TP10 only approximately left and right mastoids. 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 and downsampled to Hz 500. EEG triggers were time-locked to the onset of the target word and data was epoched from 200ms prior

to word onset (baseline corrected) to 800ms post word onset. Horizontal and vertical eye movements were monitored by means of the fronto-polar Fp1 and Fp2 electrodes which were located just above the eye-brows. Eye-blinks were removed by running an Independent Component Analysis (ICA) using the analysis implemented in EEGLab and following the procedure described in Nunez et al. (2016). After eye-blink correction, trials were rejected according to the following criteria: We utilized an automatic artifact rejection procedure with a 200ms sliding window and 50ms steps. In order to keep as many trials as possible for the analyses, we only included electrodes which were later used in the analyses (i.e. 44 electrodes) and thus excluded all electrodes placed on the most outward circle of the cap with the exception of the frontal left and right peripheral electrodes (AF7, F7, AF8, F8) which in previous studies the P350 effect was located (e.g., Scharinger & Felder, 2011). If any of the 44 electrodes exceeded a $\pm 100 \mu V$ threshold within a 50ms sliding time-window, the epoch was excluded from the analyses. This pre-processing pipeline led to the exclusion of 120 trials which comprised 3.6 % of the data (not considering nonwords). One participant exceeded our threshold of excluded trials due to artifacts (i.e. below 20 trials in more than one of the conditions) and was thus excluded from the analyses. We only interpolated one electrode across all datasets. This was mainly achieved by excluding the peripheral electrodes (e.g., T7 & T8) which were not used for the analyses and are generally prone to large muscle artifacts or enlarged alpha waves.

Previous ERP fragment priming studies reported different ERP components to be affected by the experimental manipulations. These are the N400 (300-500ms) and the P350 (300-400ms). The only ERP fragment priming study with L2 speakers, ran, in addition to those time-windows, ANOVAs in 50ms increments starting from target word onset (Kobor et al., 2018). This time-wise broader analysis, revealed that ERP effects were present also in a post N400 time-window. Thus, in addition to the P350 and N400 time-windows, we tested for any late effects in the 500-650ms time-window. As for the regions of interest (ROI), we included five regions with six electrodes each to cover any of the potential ERP effects: left anterior (AF7, F3, F5, F7, FC3, FC5), right anterior (AF8, F4, F6, F8, FC4, FC6), left posterior (CP3, CP5, P1, P3, PO3), right posterior (CP4, CP6, P2, P4, P6, PO4) and centro-parietal (Cz, C1, C2, CP1, CP2, CPz). For each of the three time-windows, we started with a repeated-measured by subject ANOVA including the averaged EEG as dependent variable and Prime (2 levels: experimental/control), ROI (5 levels), Loan Type (3 levels: NONEX, DISYLL, TRISYLL) and Foot (2 levels: 1FOOT, 2FEET) as fixed factors. With the ANOVA analyses, we are following all previous ERP fragment priming studies (Kobor et al. 2018; Scharinger & Felder, 2011; Friederich et al., 2004a, 2004b).

To further determine the time-course and topographical distribution of any effects elicited by our experimental manipulations, we ran cluster-mass permutation tests (as implemented in ERPlab: Groppe, Urbach, & Kutas, 2011) in addition to ANOVAs as reported in previous studies. We had two main reasons for applying this data-driven approach: (i.) fragment priming studies with non-native speakers are sparse (only Kobor et al., 2018) and thus time-course and topographical distribution of possible effects are largely unknown. (ii.) Even for native speaker participants, the time-course of the P350 effect does not always seem to align with the proposed 300-400ms time-window (for example see Scharinger & Felder, 2011; Figure 2).

The cluster mass statistic is suitable for broadly distributed ERP effects reflecting higher level cognitive processes (Groppe et al., 2011). We ran separate tests comparing the experimental and control items for each of the six conditions. For this comparison, ERPs from the experimental and the control conditions were submitted to a repeated-measures, two-tailed cluster-based permutation test based on the cluster mass statistic (Bullmore et al., 1999). For each test, we used a family-wise alpha level of .05. Furthermore, we included all data points between 250-700ms post target word onset (which is the onset of the earliest effect in Kobor et al., 2018)). In order to decrease the number of comparisons, we down-sampled the data to Hz 125 resulting in test bins of 8ms. We further included all electrodes apart from the outermost circle (already excluded during pre-processing; resulting in 44 electrodes x 51 time bins = 2244 comparisons).

Repeated measures t-tests were performed for each comparison using the original data and 1000 random within-participant permutations of the data. For each permutation, all t-scores corresponding to an uncorrected p-value of $< .05$ were formed into clusters with any neighbouring such t-scores. Spatial neighbours were all electrodes within approximately 5.44 cm of one another. Adjacent time points were considered temporal neighbours. The “mass” of a cluster is the sum of the t-scores in each cluster. The most extreme cluster mass in each of the 1001 sets of tests was recorded and used to estimate the distribution of the null hypothesis (i.e., no difference between conditions). The permutation cluster mass percentile ranking of each cluster from the observed data was used to derive its p-value. The p-value of the cluster was then assigned to each member of the cluster. For accessing the raw data and analyses scripts, please email the lead author.

Results

Behavioural Results

Accuracies for all control and experimental conditions were above 90% with only one condition (control non-existing in German) with correct responses below 95%. Given the near ceiling performance across conditions, we did not run any statistical tests on the accuracy data (for descriptive statistics see Table 6).

Table 6. Accuracies (%) and Standard Deviations (SD) in the experimental and control conditions

Condition	Control Accuracy (SD)	Experimental
<i>Loans not existing in German (NONEX)</i>		
1FOOT (pigeon)	93.1 (12.6)	96.2 (7.5)
2FEET (curfew)	94.3 (7.3)	95.1 (6.1)
<i>German disyllabic (DISYLL)</i>		
1FOOT (signal)	97.1 (5.3)	97.2 (5.4)
2FEET (reptile)	96.4 (6.2)	98.1 (3.1)
<i>German trisyllabic plus stress on penultimate syllable (TRISYLL)</i>		
1FOOT (carrot)	95.9 (7.0)	97.2 (5.4)
2FEET (nomad)	94.9 (6.6)	95.1 (7.0)

Descriptive statistics for the reaction time data are summarised in Table 7 showing that in each of the six conditions the experimental prime led to faster RTs. We first tested whether this priming effect (i.e., difference between control vs. experimental items) reached significance in the six conditions. For that, we ran pair-wise comparisons on a model including *Prime* and *Condition* as fixed factors. Priming effects were significant across all conditions and are summarised in Table 7 and Figure 2.

Table 7. Reaction Time Data (ms) by item and Standard Deviations (SD) for Control and Experimental (Exp.) items as well as the extend of the priming effect in each condition. Statistics show the pair-wise comparisons of experimental and control items testing the priming effect across conditions

Condition	Control RTs (SD)	Exp.	Priming Effect	χ^2	p-value
<i>NONEX</i>					
1FOOT (pigeon)	667 (71)	612 (47)	55 (83)	66.6	<0.001***
2FEET (curfew)	667 (55)	623 (42)	44 (68)	44.7	<0.001***
<i>DISYLL</i>					
1FOOT (signal)	627 (23)	591 (32)	36 (39)	40.3	<0.001***
2FEET (reptile)	627 (36)	605 (48)	22 (59)	19.0	<0.001***
<i>TRISYLL</i>					
1FOOT (carrot)	648 (57)	622 (46)	26 (72)	12.7	<0.001***
2FEET (nomad)	662 (55)	634 (50)	28 (73)	16.3	<0.001***

Syntax model: $RTsLog \sim \text{prime} * \text{condition} + (1 + \text{prime} | \text{subject}) + (1 | \text{item})$; optimizer: nlminbwrap

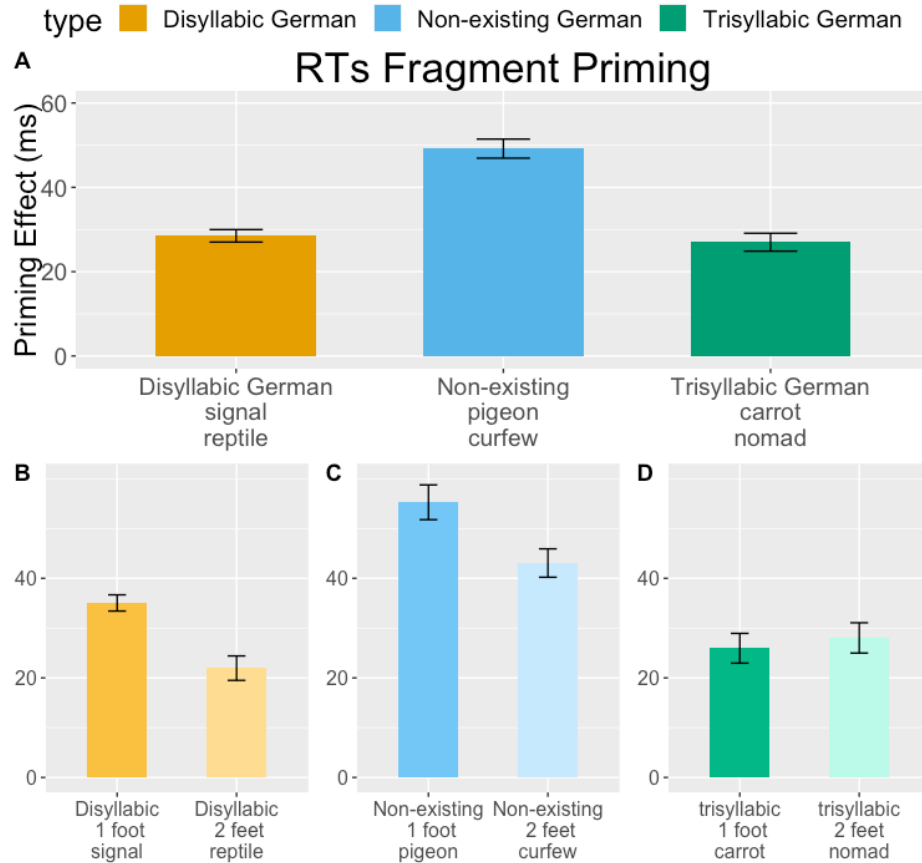


Figure 2. A. Priming effects in the three *Loan Types* collapsed over *One Foot* and *Two Feet* items. B-D. Each of the three *Loan Types* split based on *Foot Type*. Error bars represent the standard error of the mean.

We then tested whether there is a significant difference of the priming effect across *Foot* and/or *Loan Types*. For that, we first fitted a model including *Prime* (control/experimental), *Loan Type* (NONEX, DISYLL, TRISYLL) and *Foot* (1FOOT, 2FEET) as fixed factors. There was no significant three-way interaction between the factors but the interaction between *Prime* and *Loan Type* reached significance ($\chi^2 = 11.99$, $p = .002$). The best fitting model is summarised in Table 8 which shows that the significant *Prime* * *Loan Type* interaction stems from NONEX items which differ both from the TRISYLL and DISYLL items. Thus, the priming effect in the NONEX conditions is significantly larger than for the other two loan types. There was no difference between DISYLL and TRISYLL loans.

Table 8. Summary of the best fitting model. Square brackets indicate the factor level comparisons. The DISYLL - TRISYLL comparison for the main effects as well as for the interaction with the factor *Prime* was run in a separate model where DISYLL was set to the baseline for the factor *Loan Type*.

	RTsLog		
Predictors	Estimates	CI	p
(Intercept)	6.49	6.45 – 6.53	<0.001***
Prime	-0.07	-0.08 – -0.06	<0.001***
Loan Type [NONEX vs. DISYLL]	-0.05	-0.08 – -0.02	<0.001**
Loan Type [NONEX vs. TRISYLL]	-0.02	-0.05 – 0.01	0.119
Loan Type [DISYLL vs. TRISYLL]	0.03	0.00 – 0.06	0.024*
Foot Type	-0.02	-0.04 – 0.00	0.092

Prime * Type [NONEX vs. DISYLL]	0.02	0.00 – 0.04	0.026*
Prime * Type [NONEX vs. TRISYLL]	0.03	0.01 – 0.05	0.001**
Prime * Type [DISYLL vs. TRISYLL]	0.01	-0.01-0.03	0.205

Syntax: RTsLog ~ prime * type + foot + (1+type+prime+foot|subject) + (1|item), optimizer: nlminbwrap

EEG Results

Figure 3 shows topographic plots of the priming effects (control items minus experimental items) for all six conditions whilst the brainwaves in Figure 4 shows both ERPs for the control and experimental items over selected electrodes for each of the conditions. Main effects and interactions are summarised in Table 9 and the Appendix whilst Table 10 illustrates significant post-hoc comparisons in the individual ROIs.

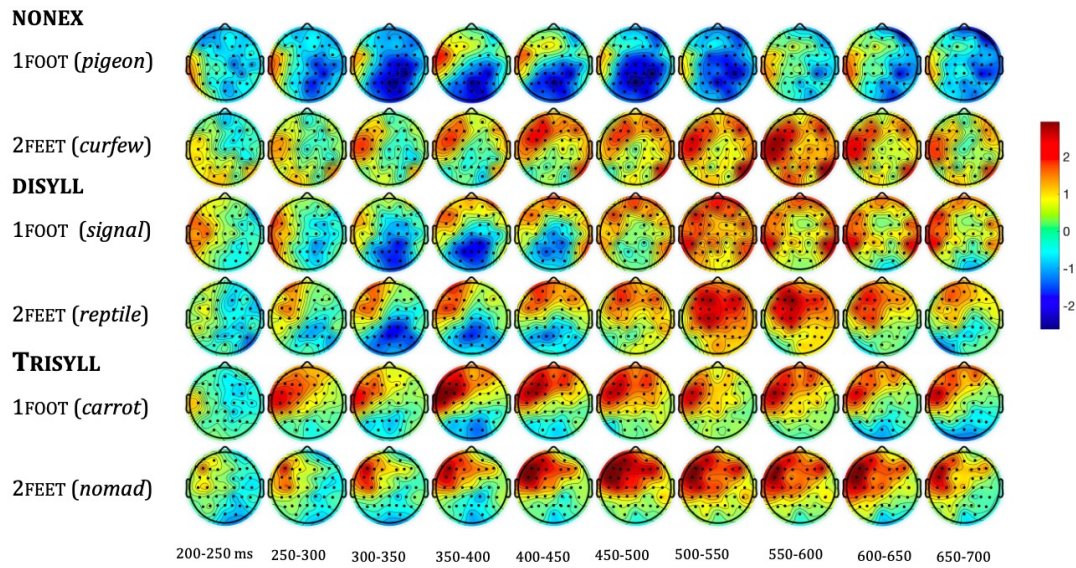
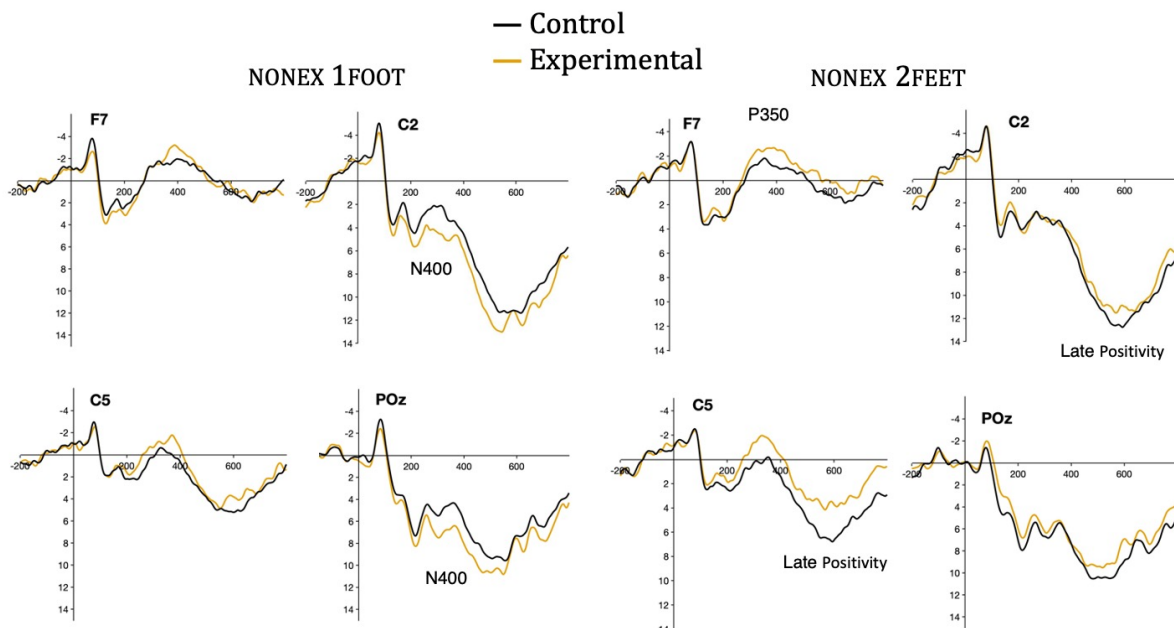


Figure 3. Topographic plots showing the difference between the control minus the experimental items for each of the three Loan Types and for the 1FOOT and 2FEET distinction between 200-700 ms post stimulus onset.



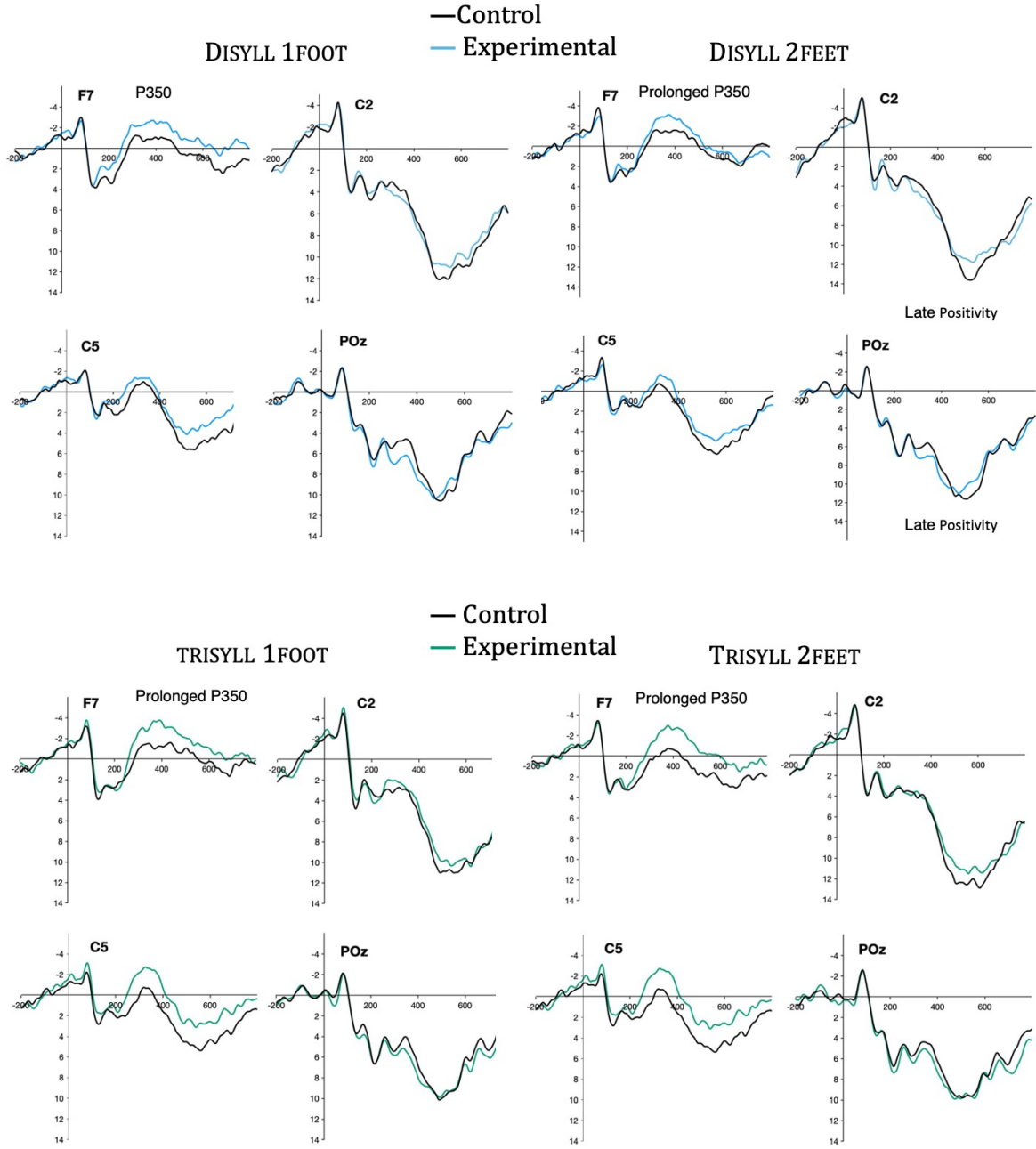


Figure 4. Brainwaves time-locked to the target-word onset illustrating the different priming effects (i.e. experimental vs. control condition) in each of the three *Loan Types* and for the *1FOOT* and *2FEET* conditions.

Time-window 300-400ms (P350)

In the 300-400ms time-window the ANOVA with the highest order significant interaction was *ROI * Loan Type * Foot* ($F(4.93, 108.43) = 5.60, p = .0001, \eta^2_G = 0.03$). Lower order significant interactions are listed in Table 9. We then ran separate ANOVAs for the three *Loan Types* (NONEX, DISYLL, TRISYLL) which showed a *ROI * Prime* interaction for all three *Loan Types* (DISYLL: $F(2.73, 60.14) = 18.10, p < .0001, \eta^2_G = 0.013$; TRISYLL: $F(2.18, 47.95) = 22.10, p < .0001, \eta^2_G = 0.013$; NONEX: $F(2.15, 47.37) = 8.47, p = .0005, \eta^2_G = 0.007$). For DISYLL and TRISYLL items, we further found a significant effect of *Foot* including as well as a *Prime * Foot* interaction. All significant interactions can be found in the Appendix.

Post-hoc pair-wise t-tests for each of the six conditions across the five ROI showed a significant positivity in the left anterior region for 2FOOT DISYLL ($t = 3.04$, $p = .006$), 2FOOT NONEX ($t = 2.65$, $p = .015$) and both TRIYLL conditions (1FOOT TRIYLL: $t = 3.72$, $p = .001$; 2FEET TRIYLL: $t = 3.16$, $p = .005$). Significant negativities were found for 2FOOT NONEX in all but the left anterior region with the strongest over right posterior electrodes ($t = -3.48$, $p = .002$). We also found a significant negativity in the 2FOOT DISYLL condition ($t = -2.18$, $p = .040$) which, however, was not confirmed by the mass permutation test. All significant t-tests are summarised in Table 10.

Time-window 300-500ms (N400)

In the N400 time-window, the $ROI * Prime * Loan Type * Foot$ ANOVA performed on the mean activity between 300 and 500ms revealed a significant main effect of *Loan Type* ($F(1.61, 35.12) = 8.70$, $p = .002$, $\eta^2_G = 0.006$) and *Foot* ($F(1,22) = 8.39$, $p = .008$, $\eta^2_G = 0.003$). A significant four-way interaction was found for $ROI * Prime * Loan Type * Foot$ ($F(4.05, 89.17) = 3.11$, $p = .018$, $\eta^2_G = 0.001$). Further significant interactions are indicated in Table 9. ANOVAs for the individual *Loan Types*, revealed a three-way interaction for $ROI * Foot * Prime$ (NONEX: $F(2.56, 56.4) = 6.50$, $p = .001$, $\eta^2_G = 0.004$). Significant two-way interactions for the different *Loan Types* are summarised in the Appendix. Furthermore, the TRIYLL and DISYLL conditions revealed a significant main effect of *Foot* (DISYLL: $F(1,22) = 7.24$, $p = .013$, $\eta^2_G = 0.009$; TRIYLL: $F(1,22) = 6.32$, $p = .02$, $\eta^2_G = 0.008$).

Post-hoc t-tests comparing experimental and control items in the five ROIs across the six conditions, showed a significant positivity in the left anterior region for all shared loanword conditions but for DISYLL 1FOOT (DISYLL 2FOOT: $t = 3.36$, $p = .003$; TRIYLL 2FOOT: $t = 3.82$, $p < .0001$; TRIYLL 1FOOT: $t = 3.66$, $p = .001$). In the 1FOOT NONEX condition a negativity was found which reached significance in all ROIs but the left anterior region with the effect being strongest over centro-parietal ($t = -3.21$, $p = .004$) and right posterior sites ($t = -3.62$, $p = .002$). All significant t-tests are marked in Table 10.

Both the topographical and time distribution is comparable to the N400 found in previous studies. As for the positivities, the prolonged effect (extending from the previous 300-400 time-window), is comparable to the effect found by Kobor and colleagues (2018).

Time-window 500-650ms

Finally, in the 500-650ms time-window, we found a significant main effect of *Foot* ($F(1, 22) = 21.36$, $p = .0001$, $\eta^2_G = 0.01$) and *Prime* ($F(1,22) = 17.88$, $p = .0003$, $\eta^2_G = 0.01$). Significant two-way interactions are summarised in Table 9. Individual ANOVAs for the three *Loan Types* showed a significant main effect for the factors *Prime* (DISYLL: $F(1,22) = 9.35$, $p = .006$, $\eta^2 = 0.02$; TRIYLL: $F(1,22) = 6.86$, $p = .016$, $\eta^2 = 0.016$) and *Foot* (DISYLL: $F(1,22) = 12.80$, $p = .002$, $\eta^2 = 0.018$; TRIYLL: $F(1,22) = 25.95$, $p < .0001$, $\eta^2 = 0.036$).

Post-hoc t-tests showed significant differences (i.e. positivities) between experimental and control conditions in the left anterior region for all conditions but the 1FOOT NONEX (NONEX 2FEET: $t = 2.31$, $p = .030$; DISYLL 1FOOT: $t = 2.09$, $p = .048$; DISYLL 2FEET: $t = 3.44$, $p = .002$; TRIYLL 1FOOT: $t = 3.80$, $p = .0009$; TRIYLL 2FEET: $t = 3.64$, $p = .001$).

In addition, the DISYLL 2FEET condition elicited a significant positivity in the right anterior ($t = 2.29$, $p = .032$) and centro-parietal region ($t = 2.68$, $p = .034$). Finally, the NONEX 2FEET condition also elicited a significant positivity in the right anterior and posterior regions (apart from the centro-parietal region) with p-values ranging from $p = .033$ -.047. All significant t-tests are summarised in Table 10.

Table 9. Results from the ROI * Prime * Loan Type * Foot ANOVA for each of the three time-windows.

* <.05, **<.01, *** <.001

<i>Effect</i>	<i>300 - 400ms</i>	<i>300 - 500ms</i>	<i>500 - 650ms</i>
Prime			***
Loan Type	**	**	
Foot	*	**	***
ROI * Prime	***	***	***
ROI * Loan Type	*	**	
Prime * Loan Type			
ROI * Foot			
Prime * Foot			
Loan Type * Foot			*
ROI * Prime * Loan Type		**	
ROI * Prime * Foot			
ROI * Loan Type * Foot	***	***	
Prime * Loan Type * Foot			
ROI * Prime * Loan Type * Foot		*	

Table 10. Post-hoc paired-wise t-tests for each of the five ROIs for the factor Loan Type (i.e., Non-existent in German, Disyllabic in German, Trisyllabic in German) for each of the three time-windows.

<.05, **<.01, *** <.001; (+) and (-) indicate whether the difference between the control and experimental condition is positive or negative.

	<i>300-400 ms</i>			<i>300-500 ms</i>			<i>500-650 ms</i>		
<i>Effect</i>	<i>NONEX</i>	<i>DISYLL</i>	<i>TRISYLL</i>	<i>NONEX</i>	<i>DISYLL</i>	<i>TRISYLL</i>	<i>NONEX</i>	<i>DISYLL</i>	<i>TRISYLL</i>
1FOOT									
Left anterior			** (+)			** (+)		* (+)	** (+)
Right anterior	* (-)			* (-)					
Centro-parietal	* (-)			** (-)					
Left posterior	* (-)			* (-)					
Right posterior	** (-)			** (-)					
2FEET									
Left anterior	* (+)	** (+)	** (+)	* (+)	** (+)	*** (+)	* (+)	** (+)	*** (+)
Right anterior							* (+)	* (+)	
Centro-parietal								* (+)	
Left posterior							* (+)		
Right posterior		* (-)					* (+)		

Cluster mass permutation test

For each of the six conditions, we ran a separate cluster mass permutation test comparing experimental vs. control items. For each of the six tests, we report only significant clusters ($p < .05$) and their corresponding t-scores. In all, but the DISYLL 1FOOT condition (all clusters $p > .1$), the permutation resulted in a significant cluster. Results from the cluster mass permutation tests are illustrated in Figure 5. For the NONEX conditions, we found one significant negative cluster for the 1FOOT items ($M = -1906$, $p = .006$) in the observed data spanning from around 250-550ms with a centro-parietal distribution. The 2FEET condition, however, elicited a significant positive cluster ($M = 1190$; $p = .024$) strongest over left anterior sites and starting at around 250ms and lasting throughout the tested time-window (i.e. 700ms). For the DISYLL 2FEET condition, we also found one positive cluster ($M = 1480$, $p = .036$) but with a different topographical distribution.

The onset of the cluster is at around 300ms but distributed across left frontal electrodes whilst later time-points are also significant across more posterior sites. The tests from the TRISYLL conditions show very similar positive clusters, i.e. a prolonged positivity mainly over anterior sites (i.e., left anterior electrodes) starting at 250ms post-stimulus onset and lasting until 700ms (1FOOT: $M = 2019$, $p = .002$; 2FEET: $M = 1775$, $p = .016$).

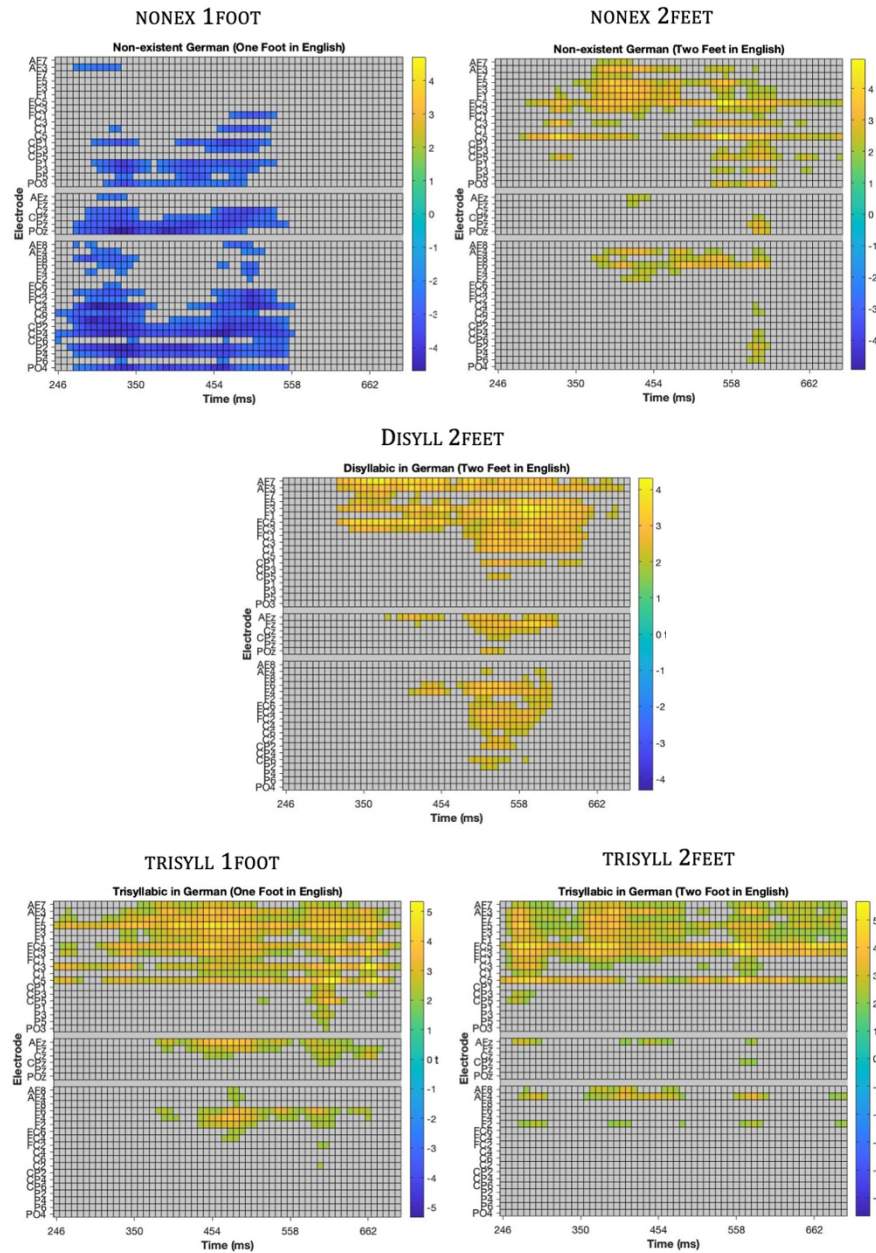


Figure 5. Time-bins and electrodes forming a significant cluster found by running a cluster mass permutation test in the 250-700ms time-window across 44 electrodes.

Discussion

In the present study, we tested the effect of L1 metrical structure on L2 word processing for native German speakers who are highly proficient in English. This also allowed us to gain further insights into the role of metrical feet in word processing and more generally the influence of the L1 phonological system onto L2 word processing. The stimuli were all disyllabic English words with initial stress which had been borrowed from Romance languages at different times. For our experiments, the items were divided into three categories, each with a different relationship to their German counterparts: non-existent in German (NONEX), disyllabic with final stress (DISYLL),

and trisyllabic with non-initial stress (TRISYLL). All items in these three categories consisted either of one foot (1FOOT) or two feet (2FEET) in English. The results can be summarised in four main points. First, the behavioural results showed a significant priming effect across all conditions, i.e., faster RTs for matching primes compared to control primes. However, largest priming effects were found for items which do not have a counterpart in German.

Second, we found that the metrical structure of German modulated different ERP components when comparing the experimental items with their controls. Only the NONEX 1FOOT condition in German showed a negativity in the typical N400 time-window (and for both the ANOVA and the mass permutation test) whilst all other conditions elicited positivity effects. This suggests different integration processes for words without a German counterpart ('non-cognates' in the L2 literature) but only if there is enough phonological overlap. That is, in terms of foot structure the 2FEET condition (e.g., *curfew*) elicited a left anterior positivity (P350) as well as a late positivity with a more posterior distribution.

Third, the conditions with disyllabic counterparts in German comprising of 2FEET showed a left anterior positivity in an early time-window (300-400) as well as a more widely distributed positivity across posterior sites in a later time-window. Cluster mass permutation tests showed that this positivity comprises of only one positive cluster in the 2FEET condition with an onset at around 300ms. Given the posterior distribution, we interpret this effect as a post-lexical effect similar to a P600 (Brouwer, Fitz, & Hoeks, 2012) reflecting processes linked to a re-analysis of the target word triggered by the German counterpart.

Lastly, the conditions with trisyllabic counterparts in German showed a prolonged positivity across anterior frontal electrodes starting at the onset of the testing window (250ms) and spanning across the testing window. Compared to the DISYLL 1FOOT condition, the effect did not spread across posterior sites in a later time-window indicating that, although prolonged, the elicited effect can be linked to the P350 (cf. Friedrich, Felder, Lahiri, & Eulitz, 2013). In the following, we will discuss the implications of these results for the processing of shared loanwords and L2 phonological processing more generally.

Loan (cognate) status

Both our behavioural and ERP results show that the processing of loanwords in an L2 is dependent on the existence of a counterpart in their L1. We attribute the greater priming effect for NONEX loans in our behavioural data to differences in lexical competition: if no similar words exist in the L1, there is no competition and thus greater ease of lexical/phonological integration for the L2 speaker and the matching prime the greatest amount of priming in accessing the target word. This is seemingly in contrast with previous studies where 'cognates' facilitated word recognition (see Lijewska, 2020 for an overview). However, when looking at the RTs in the control items (see Table 7) in the two non-shared loanword conditions, we see that those are slower compared to the shared loanword conditions (apart from the TRISYLL 2FEET condition) indicating a visual 'cognate' facilitation effect without the target being primed. Since the listeners can use this to its full advantage as there is no possible inhibitory information or competition which would need to be resolved before a response can be made. Previous studies that did investigate priming effects in relation to 'cognate' status used a masked translation priming paradigm where the prime and target are not presented in the same language and the prime was a whole word (e.g., Davis et al., 2009; Dunabeitia, Perea, & Carreiras, 2010). Thus, the methodological differences between the current study and previous studies do not allow a direct comparison since we assume different underlying processes for fragment primes vs. whole word primes in terms of cohort activation as well as a mixed language task vs. a single language task. In particular, for the fragment priming study, we assume that auditorily presented experimental prime activates the cohort, i.e., all words with the same initial syllable. Consequently, in the experimental conditions structural overlap between prime fragment and target creates

competition in certain conditions which is evident in the modulation of different ERP components.

Further, we found that both the phonological similarity between the non-shared loanwords (1FOOT and 2FEET) and the typical phonological structure in German disyllabic words influence word processing. Disyllabic trochaic words ending in a light syllable are the norm in German words of Germanic origin (e.g., *Schulter* ['ʃʊl.tɐ] 'shoulder', *Hafen* ['ha:fɐ] 'haven, port') which corresponds to all 1FOOT words in English that we used in this experiment. However, for words that consist of two feet, the tendency in German would be to stress the second foot, which is the second syllable rather than the first, e.g., *Reptil* [rep.'ti:l] 'reptile'. Thus, we assume that the 2FEET items in English which do not have phonologically related counterparts in German are processed similarly to the items which do have a German counterpart. We will elaborate on these processing differences when discussing the P350 and the P600 effect.

These findings have important implications for L2 'cognate' processing. Studies so far have made a two-way distinction — either words were classified as 'cognates' with varying degrees of semantic, phonological and orthographical overlap or they functioned as control items in a 'non-cognate' condition. Our results, however, have shown that even within the category of 'non-cognates' (or non-loans) phonological overlap in terms of metrical structure matters considerably. Future research should aim to disentangle whether the impact of phonological overlap also translates to the level of segmental phonology.

Our results also extend Muntendam et al.'s (2022) findings who found that stress placement influences 'cognate' auditory word recognition, especially in terms of how competition unfolds for words shared across languages. Cognate facilitation in their study was the strongest if the initial syllable bore stress in both languages presumably leading to the strongest co-activation whilst differences in stress placement prevented a facilitation effect of cognates. Based on our behavioural priming effects, which we found across all conditions, we assume that regardless of the varying stress patterns (and consequently foot structure), also the loanword in the participants L1 (German) was activated. However, to differing degrees similar as in Muntendam et al.'s study.

Our ERP results further indicate that it is not only stress placement that influences the processing of shared words but also the differences in foot structure are taken into account during processing. Future studies may also want to include a condition with the same foot structure across languages to see whether this would eliminate any interference effects.

N400 vs. P350

The modulation of different ERP components across conditions suggests that different integration processes are at play depending on the participants' L1 phonology. What underlying processes do the P350 and the phonological N400 (or central negativity) index? According to previous studies, the P350 reflects a modality-independent activation of word form representations (Friedrich et al., 2009; Kobor et al., 2018) which exclusively relies on orthographic and/or phonological matching. The phonological N400 within a fragment priming context, however, can be related to phonological expectancies derived from the cohort activated by the prime. Thus, the difference between the two components is that the P350 is not sensitive to the co-activated neighbours whilst the N400 takes the cohort into account. Evidence for this hypothesis comes from Friedrich et al. (2009) who found that the fragment priming paradigm elicited a P350 effect regardless of the length of the prime. Only the N400 was sensitive to the length of the fragment. That is, longer fragments showed a larger reduction of the N400 compared to shorter fragments which can be explained by the larger activated cohort for shorter fragments which results in more competition. The P350 on the other hand was not modulated by the length of the priming fragment suggesting its independence from the activated cohort.

Referring back to the present study, we assume that in the NONEX condition, the L1 cohort would not be activated and thus, the number of competing items is minimal. For the *pigeon* class (1FOOT), we see a reduced N400 which we assume is based on a lack of cohort competition plus a trochee

leading to the strongest lexical integration. For *curfew* words (2FEET), although there is no L1 cohort competition, there is a structural mismatch: Participants' L1 phonology (i.e. German) would prefer the second syllable to be stressed because it is heavy, as opposed to the first syllable as it is most common in English. Based on this structural mismatch, we see a P350 indicating phonological mapping between prime and target. For both TRISYLL conditions, the competition from the L1 cohort prevents an N400 effect but a P350 which is independent of the cohort was still present.

For the 2SYLL 2FEET condition and the NONEX 2FEET condition, it is unclear whether the effect that started over left anterior electrodes was exclusively a late positivity or a P350 plus a late positivity. The brainwaves in Figure 5 suggest that there might be a difference between the prolonged P350 effect strongest over left anterior sites and the late positivities strongest over the posterior midline and the posterior right sites. Given the topographical distribution, the latency and duration, the late central-posterior positivity strongest over the midline or the right hemisphere (for 2SYLL 2FEET and NONEX 2FEET), suggests the modulation of a different ERP component, i.e. the P600.

Late positivity

The modulation of the ERPs in the conditions with DISYLL (2FEET) vs. TRISYLL German counterparts suggests different underlying processes when encountering such words. Based on the different topographical distribution, we attribute the late positivity effect for the DISYLL conditions to a P600 component reflecting post-lexical reanalysis based on the German counterpart which is, in terms of foot structure, most similar to the English word presented to the participants. A similar late positivity was found in Peeters, Dijkstra, and Grainger (2013) when comparing 'cognate' vs. 'non-cognate' processing in a lexical decision task with 'identical cognates' (i.e., words with the exact same orthography in L1 and L2). The authors suggested that this P600 was elicited by a conflict created by a possible French (L1) reading of the target word which was presented in English (L2). We assume that a similar conflict based on phonological similarity has been triggered by the words with disyllabic counterparts in German. The cluster mass permutation test only showed a significant positive cluster for the items with 2FEET in English (e.g., *reptile*) but not for the 1FOOT items. This also fits with the similarity hierarchy where the metrically most similar words in German elicit the greatest effort difficulties. Thus, a word with two feet, such as *reptile*, is metrically more similar to the German counterparts than the one-foot items (e.g., *signal*). This would then also explain the NONEX 2FEET conditions that also elicited a posterior extension of the earlier positivity in the 500-650 time-window and is also evident in the mass permutation test. As discussed earlier, such two feet words would usually receive final stress in English and thus potentially triggered a re-analysis after lexical processing. Presumably, the TRISYLL items, due to the different number of syllables, were phonologically too distant to the English counterparts so that no reanalysis was triggered after lexical integration of the target word.

In line with most recent L2 processing literature, our results also support an integrated bilingual lexicon (e.g., Dijkstra & van Heuven, 2001) as both languages affect the reactions to the target despite the task being completely in English. Much of the research into phonological effects in L2 processing, especially when dealing with items shared across languages, has focused either on the degree of overlap in terms of phonemes or, if on the suprasegmental level, on word stress only. There have to date not been any studies, to the best of our knowledge, which investigate the metrical foot as a potential processing unit and its importance in bilingual processing.

To conclude, our study showed that the German L1 metrical phonology guides L2 processing of shared loanwords even in highly proficient L2 speakers. Moreover, we could demonstrate that metrical phonology even plays a role in the processing of loanwords that only exist in English (L2), i.e., the more similar the foot structure to the L1's typical metrical pattern, the easier the processing. For shared loanwords, our data indicates that the more similar a word is in terms of

its metrical phonology across L1 and L2, the more effortful the processing of a word indicating effects of the German phonological system.

References

- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv preprint arXiv:1506.04967.
- 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. doi:<https://doi.org/10.1006/brln.2001.2532>
- Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about Semantic Illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, 1446, 127-143. doi:<https://doi.org/10.1016/j.brainres.2012.01.055>
- Brown, C., & Hagoort, P. (1993). The Processing Nature of the N400: Evidence from Masked Priming. *Journal of Cognitive Neuroscience*, 5(1), 34-44. doi:10.1162/jocn.1993.5.1.34
- 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. doi:10.1109/42.750253
- Chwilla, D. J., & Kolk, H. H. J. (2005). Accessing world knowledge: Evidence from N400 and reaction time priming. *Cognitive Brain Research*, 25(3), 589-606. doi:<https://doi.org/10.1016/j.cogbrainres.2005.08.011>
- Comesana, 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. doi:10.1037/xlm0000065
- Comesana, M., Sanchez-Casas, R., Soares, A. P., Pinheiro, A. P., Rauber, A., Frade, S., & Fraga, I. (2012). The interplay of phonology and orthography in visual cognate word recognition: an ERP study. *Neurosci Letters*, 529(1), 75-79. doi:10.1016/j.neulet.2012.09.010
- , 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. and Jesse, A. (2021). Word Stress in Speech Perception. In *The Handbook of Speech Perception* (eds J.S. Pardo, L.C. Nygaard, R.E. Remez and D.B. Pisoni). <https://doi.org/10.1002/9781119184096.ch9>
- Davis, C., Sanchez-Casas, R., Garcia-Albea, J. E., Guasch, M., Molero, M., & Ferre, P. (2009). Masked translation priming: Varying language experience and word type with Spanish-English bilinguals. *Bilingualism: Language and Cognition*, 13(2), 137-155. doi:10.1017/s1366728909990393
- De Rosario-Martinez, H., Fox, J., Team, R. C., & De Rosario-Martinez, M. H. (2015). Package 'phia'. CRAN Repos. Retrieved, 1, 2015.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9-21. doi:<https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: The Neglected Role of Phonology. *Journal of Memory and Language*, 41, 496-518.
- Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5(3), 175-197.

- 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. doi:10.1016/j.jml.2009.12.003
- Domahs, U., Wiese, R., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). The processing of German word stress: evidence for the prosodic hierarchy. *Phonology*, 25(1), 1-36. <https://doi.org/10.1017/S0952675708001383>
- 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>
- Downey, S. S., Hallmark, B., Cox, M. P., Norquest, P., & Lansing, J. S. (2008). Computational Feature-Sensitive Reconstruction of Language Relationships: Developing the ALINE Distance for Comparative Historical Linguistic Reconstruction. *Journal of Quantitative Linguistics*, 15(4), 340-369. doi:10.1080/09296170802326681
- Dunabeitia, J. A., Perea, M., & Carreiras, M. (2010). Masked translation priming effects with highly proficient simultaneous bilinguals. *Experimental Psychogyl*, 57(2), 98-107. doi:10.1027/1618-3169/a000013
- 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. *Scientific Reports*, 11(1), 12812. doi:10.1038/s41598-021-92259-z
- Friedrich, C. K., Felder, V., Lahiri, A., & Eulitz, C. (2013). Activation of words with phonological overlap. *Front Psychol*, 4, 556. doi:10.3389/fpsyg.2013.00556
- 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 Research*, 20(2), 300-308. doi: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. doi: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. doi: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. doi:<https://doi.org/10.1016/j.biopsycho.2008.04.012>
- Grainger, J., & Jacobs, A.M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, 22, 696-713. <https://doi.org/10.1037/0033-295X.103.3.518>
- Grosjean, F., & Gee, J. P. (1987). Prosodic structure and spoken wor recognition. *Cognition*, 25 (1-2), 135-155. [https://doi.org/10.1016/0010-0277\(87\)90007-2](https://doi.org/10.1016/0010-0277(87)90007-2)
- 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. doi:<https://doi.org/10.1111/j.1469-8986.2011.01273.x>
- Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. University of Chicago Press.
- Henrich, K., Alter, K., Wiese, R., & Domahs, U. (2014). The relevance of rhythmical alternation in language processing: An ERP study on English compounds. *Brain and language*, 136, 19-30.
- Hoshino, N., & Thierry, G. (2011). Language selection in bilingual word production: Electrophysiological evidence for cross-language competition. *Brain research*, 1371, 100-109. <https://doi.org/10.1016/j.brainres.2010.11.053>
- Jared, D., & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words?. *Journal of Memory and Language*, 44(1), 2-31. <https://doi.org/10.1006/jmla.2000.2747>
- Kennard, Holly & Aditi Lahiri (2020). Nonesuch phonemes in loanwords. *Linguistics*, 58, 83-108. doi: 10.1515/ling-2019-0033.

- 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. *Int J Psychophysiol*, 128, 100-110. doi:10.1016/j.ijpsycho.2018.04.006
- Kondrak, G. (2003). Phonetic Alignment and Similarity. *Computers and the Humanities*, 37(3), 273-291. doi:10.1023/A:1025071200644
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of statistical software*, 82, 1-26.
- Lahiri, A. (2001). Metrical patterns. In E. König & M. Haspelmath (eds.) *Language Typology and Language Universals*, 1347-1367. Berlin: Mouton de Gruyter.
DOI: [10.1515/9783110171549.2](https://doi.org/10.1515/9783110171549.2).
- Lemhofer, K., & Broersma, M. (2012). Introducing LexTALE: a quick and valid Lexical Test for Advanced Learners of English. *Behav Res Methods*, 44(2), 325-343. doi:10.3758/s13428-011-0146-0
- Levenshtein, V. I. 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10, 707-710.
- Lijewska, A. (2020). Cognate Processing Effects in Bilingual Lexical Access. In R. Heredia & A. Cieślicka (Eds.), *Bilingual Lexical Ambiguity Resolution* (pp. 71-95). Cambridge: Cambridge University Press. doi:10.1017/9781316535967.005
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 8: 213. doi:10.3389/fnhum.2014.00213
- Magne, C., Astésano, C., Aramaki, M., Ystad, S., Kronland-Martin, 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>
- Marian, V., Blumenfeld, H., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, 50, 940-967. doi:<https://doi.org/10.1044/1092-4388>
- 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 G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148-172). The MIT Press.
- 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. doi: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. doi:10.1017/s1366728922000037
- Norris, D., & Cutler, A. (1988). The relative accessibility of phonemes and syllables. *Perception & Psychophysics*, 43(6), 541-550.
- 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. *Handbook of neuroimaging data analysis*, 175-197.
- 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. doi:10.1016/j.jml.2012.12.003
- Praamstra, P., Meyer, A. S., & Levelt, W. J. M. (1994). Neurophysiological Manifestations of Phonological Processing: Latency Variation of a Negative ERP Component Timelocked to Phonological Mismatch. *Journal of Cognitive Neuroscience*, 6(3), 204-219. doi:10.1162/jocn.1994.6.3.204

- Reinisch, E., Jesse, A., & McQueen, J. M. (2010). Early use of phonetic information in spoken word recognition: lexical stress drives eye movements immediately. *Q J Exp Psychol*, 63(4), 772-783. doi:10.1080/17470210903104412
- Scharinger, M., & Felder, V. (2011). ERP signatures of cross-modal semantic fragment priming: Early context effects in speech perception. *International Journal of Psychophysiology*, 80(1), 19-27. doi:https://doi.org/10.1016/j.ijpsycho.2011.01.006
- Schutz, K., Schendzielarz, I., Zwitserlood, P., & Vorberg, D. (2007). Nice wor_ if you can get the wor_: subliminal semantic and form priming in fragment completion. *Conscious Cogn*, 16(2), 520-532. doi:10.1016/j.concog.2006.09.001
- 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. doi:10.1006/jmla.2000.2783
- 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. doi: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. doi:10.1080/17470218.2013.850521
- Wiese, R. (2000) *The Phonology of German*, Oxford University Press.
- Willems, R. M., der Haegen, L. V., Fisher, S. E., & Francks, C. (2014). On the other hand: including left-handers in cognitive neuroscience and neurogenetics. *Nature Reviews Neuroscience*, 15(3), 193-201. doi:10.1038/nrn3679
- Zonneveld, W., Trommelen, M., Jessen, M., Bruce, G. & Árnason, K. (1999) "8. Wordstress in West-Germanic and North-Germanic languages". In *4 Word Prosodic Systems in the Languages of Europe* (ed. Harry van der Hulst). Berlin, New York: De Gruyter Mouton, (pp. 477-604). https://doi.org/10.1515/9783110197082.2.477
- Zwitserlood, P. (1996). Form Priming. *Language and Cognitive Processes*, 11(6), 589-596. doi:10.1080/016909696386980

APPENDIX A

Interactions and main effects from ANOVAs for the three tested time-windows and split according to the Loan Type (i.e., Non-existent in German, Disyllabic in German, Trisyllabic in German).

* <.05, **<.01, *** <.001

	<i>300-400 ms</i>		
<i>Effect</i>	<i>NONEX</i>	<i>DISYLL</i>	<i>TRISYLL</i>
Prime			
Foot		*	*
ROI * Prime	**	***	***
ROI * Foot		**	*
Prime * Foot	*		
ROI * Prime * Foot			

	<i>300-500 ms</i>		
<i>Effect</i>	<i>NONEX</i>	<i>DISYLL</i>	<i>TRISYLL</i>
Prime			
Foot		*	*
ROI * Prime	*	***	***
ROI * Foot		*	*
Prime * Foot	*		
ROI * Prime * Foot	**		

	<i>500-650 ms</i>		
<i>Effect</i>	<i>NONEX</i>	<i>DISYLL</i>	<i>TRISYLL</i>
Prime		**	*
Foot		**	***
ROI * Prime	**		**
ROI * Foot			
Prime * Foot			
ROI * Prime * Foot			