

## Cross-lingual priming of cognates and interlingual homographs from L2 to L1

**Eva D. Poort**, Department of Experimental Psychology, University College London, United Kingdom; Neurobiology of Language department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, [eva.poort@mpi.nl](mailto:eva.poort@mpi.nl)

**Jennifer M. Rodd**, Department of Experimental Psychology, University College London, United Kingdom, <https://orcid.org/0000-0002-8608-7244>, [j.rodd@ucl.ac.uk](mailto:j.rodd@ucl.ac.uk)

The aim of the current study was to explore whether lexical processing in a bilingual's *first* language (L1) can be influenced by recent experience in their *second* language (L2). We focussed on word forms that exist in both their languages, and have either the same meaning (cognates) or a different meaning (interlingual homographs). Our previous experiments provided evidence for the reverse form of cross-lingual priming: processing of interlingual homographs in a bilingual's L2 is delayed by recent experience with these words in their L1, while processing of cognates can be speeded up (Poort et al., 2016; Poort & Rodd, 2019b). In the current experiment, Dutch–English bilinguals ( $n = 106$ ) first encountered cognates ( $n = 50$ ), interlingual homographs ( $n = 50$ ) and translation equivalents ( $n = 50$ ) embedded in *English* sentences. After a 15 minute delay they made *Dutch* semantic relatedness judgements to these target words. Significant cross-lingual priming was observed for the interlingual homographs, but not for the cognates. The magnitude of this L2-to-L1 priming effect did not differ from our earlier L1-to-L2 priming effect (Poort & Rodd, 2019b). We also addressed subsidiary questions regarding the (unprimed) processing of cognates and interlingual homographs. Consistent with our previous findings (Poort & Rodd, 2019b), we found a large interlingual homograph inhibition effect in an L1 semantic relatedness task, but no evidence for a cognate facilitation effect in this task. These findings together emphasise the high level of cross-lingual interaction in the bilingual mental lexicon, especially in language-switching contexts.

Data, scripts, materials and pre-registration available via [osf.io/2swyg/](https://osf.io/2swyg/).



## 1. Introduction

Bilinguals regularly encounter word forms that exist in both of their languages. These word forms often have the same meaning, like the Dutch–English cognate *wolf*. Alternatively, they can have two entirely unrelated meanings, like the interlingual homograph *angel*, which means ‘insect’s sting’ in Dutch. A wealth of research indicates that bilinguals process these two types of words differently from translation equivalents, pairs of words like *wortel* and *carrot* that share their meaning but do not have similar orthographic or phonological forms. For example, in lexical decision experiments, participants respond more quickly to cognates than to translation equivalents, but tend to respond more slowly to interlingual homographs (for a review, see e.g. Dijkstra, 2005; Dijkstra & Van Heuven, 2012; Poort & Rodd, 2017a; Dijkstra & Van Heuven, 2018; Poort & Rodd, 2019b). Such findings support the view that words from a multilingual’s different languages are stored in a shared mental lexicon, and can become active in parallel during lexical access.

Additional evidence for cross-lingual interaction in the lexicon comes from a study showing that processing of cognates and interlingual homographs in a bilingual’s second language (L2) is influenced by recent experience with these words in their native language (L1; Poort et al., 2016). In this experiment, Dutch–English bilinguals first read sentences in *Dutch* that each contained either a cognate, an interlingual homograph or the Dutch translation of an English control word. After a delay of approximately 15 minutes, they then made *English* lexical decisions to these words. For the cognates, priming was beneficial: participants responded 28 ms more *quickly* to primed cognates than unprimed cognates. In contrast, for interlingual homographs, where there is a meaning change between the two languages, an interference effect was observed: priming *delayed* participants’ responses by 49 ms relative to unprimed homographs, suggesting that the increased availability of the Dutch meanings interfered with the participants’ ability to process the English reading of these ambiguous word forms. Given that many bilinguals regularly switch between their languages in their daily lives, this cross-lingual priming could have consequences for the ease with which they access and process the meanings of individual words in those languages after a switch.

In two follow-up experiments, we explored whether such cross-lingual priming also occurs for non-identical cognates (e.g. *kat* in Dutch and *cat* in English; Poort & Rodd, 2017b). These experiments, however, did not replicate the original findings: there was no statistical or numerical evidence for cross-lingual priming for the non-identical cognates, nor for the identical cognates or identical interlingual homographs. This inconsistency across experiments may result from the inconsistent need for participants to fully access semantic representations during lexical decision tasks (for details, see Poort & Rodd, 2017b, 2019b). This view was supported by our subsequent replication of Poort et al.’s (2016) original finding of facilitative priming for cognates but disruptive priming for interlingual homographs using a semantic relatedness task, which

does consistently require participants to access the words' meanings (Poort & Rodd, 2019b, Exp. 2). Still, the priming effects in this second follow-up study were smaller than expected based on the earlier findings: only 5 ms for the cognates and 10 ms for the interlingual homographs.

It remains uncertain, therefore, how much interaction occurs on a daily basis between word forms that exist in both of a bilingual's languages, or what factors might modulate the size of this cross-lingual priming effect. Importantly, all the aforementioned experiments measured the extent to which experience in the participants' L1 (Dutch) influenced subsequent processing in their L2 (English). The primary aim of the current experiment is therefore to determine whether the cross-lingual priming effect can also be observed when we reverse the direction of transfer, by priming items in the participants' L2 (English) and then assessing the processing of these primed words in their L1 (Dutch). In addition, by directly comparing the results of the current experiment to our earlier data (Poort & Rodd, 2019b), we can investigate whether this switch in direction of priming will increase or decrease the magnitude of the observed effect.

Intuitively, most bilinguals would assume that experience in their L2 would be unlikely to influence subsequent processing of words in their L1 *more* than their L1 can influence their L2. Many bilinguals feel that their second language is more malleable than their first language, because it is the one they learnt later in life and so will always feel like a work in progress. In contrast, they have a lifetime of experience with their first language, which lends it a sense of relative stability. Current experimental evidence, while not decisive, appears to be in line with this subjective experience. If there exists any asymmetry at all in the extent to which each of a bilingual's languages can exert influence on the other language, L1 indeed seems more likely to influence L2 than the other way around.

Support for this comes from a large body of research on (masked) translation priming that has shown that priming effects are consistently obtained when an L2 target item is primed by its L1 translation (i.e. L1-to-L2 priming; De Groot & Nas, 1991; Williams, 1994; Gollan et al., 1997; Jiang, 1999; Jiang & Forster, 2001; Kim & Davis, 2003; Basnight-Brown & Altarriba, 2007; Voga & Grainger, 2007; Duyck & Warlop, 2009; Dimitropoulou et al., 2011a, 2011b; Lupker et al., 2015; Wen & Van Heuven, 2017). When the direction of priming is reversed and an L1 target is primed by its L2 translation (i.e. L2-to-L1 priming), the priming effect is not as consistently observed and is usually weaker (Gollan et al., 1997; Grainger & Frenck-Mestre, 1998; Jiang, 1999; Jiang & Forster, 2001; Finkbeiner et al., 2004; Duyck & Warlop, 2009; Schoonbaert et al., 2009; Davis et al., 2010; Dimitropoulou et al., 2011a, cf. 2011b; Witzel & Forster, 2012; Nakayama et al., 2013; Wang, 2013; Chen et al., 2014; Wang & Forster, 2014).

The situation is not as simple as this, however, and proficiency appears to modulate priming significantly: the asymmetry in priming appears most robustly in unbalanced bilinguals (Gollan et al., 1997; Jiang, 1999; Finkbeiner et al., 2004; Nakayama et al., 2013), but decreases or is even entirely absent in balanced bilinguals (Perea et al., 2008; Duñabeitia, Dimitropoulou, et

al., 2010; Duñabeitia, Perea, et al., 2010). For example, Davis et al. (2010) found that a group of beginning bilingual participants only showed significant masked priming effects for pairs of Spanish–English cognates (e.g. *rich–rico*) when the primes were in Spanish (L1) and the targets in English (L2). However, for more highly proficient bilinguals, Davis et al. (2010) found no effect of language dominance on the magnitude of this cognate priming effect. Similarly, using a picture naming paradigm with highly fluent Spanish–English bilinguals, Francis et al. (2003, Exp. 1) found that priming from L1-to-L2 was equally effective as priming from L2-to-L1.

Going even further, Lee et al. (2018) observed L2-to-L1 priming in *unbalanced* Korean–English bilinguals, although only when the stimulus onset asynchrony (SOA) was 150 ms. When the SOA was only 60 ms, no L2-to-L1 priming was observed, which they argue is due to the fact that 60 ms is not sufficient time for the L2 translation to prime the L1 target. This could also explain the lack of L2-to-L1 priming effects in unbalanced bilinguals in previous research. In other words, there may not be an asymmetry and, under the right circumstances, L2-to-L1 priming may be observed as consistently and be equally effective as L1-to-L2 priming.

Further support for this idea comes from previous research on monolingual word-meaning priming (Rodd et al., 2013; Rodd et al., 2016; Betts et al., 2018; Gilbert et al., 2018; Gaskell et al., 2019; Rodd, 2020; Gilbert et al., 2021), which as a paradigm is much closer in design to our bilingual studies (Poort et al., 2016; Poort & Rodd, 2017b, 2019b; Exp. 2) than (masked) translation priming. This research has shown that an encounter with one meaning of a homonym (e.g. the ‘tree covering’ meaning of *bark*) can boost the availability of this primed meaning relative to its unprimed meaning (i.e. the ‘dog noise’ meaning of *bark*). Importantly, larger priming effects are observed when the prime encounter is with the strongly *subordinate* meaning (Rodd et al., 2013; Betts, 2018). This effect of meaning dominance most likely arises because, for (relatively rare) subordinate meanings, even a single encounter is highly informative, indicating that this meaning is more likely to occur in the near future than the participant would otherwise have expected. In contrast, a single encounter with the highly frequent dominant meaning, which is more regularly encountered by participants in their everyday lives, is unlikely to produce a meaningful change to subsequent behaviour (Rodd et al., 2013; Betts, 2018). By analogy, we might therefore expect that for bilingual participants, an encounter with the relatively less familiar *L2* interpretation of the word form would produce a more substantial change in how this word form is subsequently processed in L1, compared with the influence of an encounter with its L1 interpretation on L2 processing.

The current experiment addresses these questions about cross-lingual priming from L2 to L1 by replicating our previous experiment (Poort & Rodd, 2019b, Exp. 2), but switching the direction of priming: participants first encountered cognates and interlingual homographs (and translation equivalents) embedded in English (L2) sentences; the impact of this experience on processing of these words in the participants’ L1 was then tested using a Dutch semantic

relatedness task. Setting the current experiment up as a replication of the previous experiment allowed for a direct, statistical comparison of the magnitude of the priming effect between these two experiments. We predicted that this comparison would show that the priming effect is greater when cognates and interlingual homographs are primed in the participants' L2 (English, as in the current experiment) than when they were primed in the participants' L1 (Dutch, as in the previous experiment).

Furthermore, in addition to these primary aims regarding the presence of cross-lingual priming, the data from the unprimed conditions in the current study can also contribute to the ongoing discussion surrounding the cognate facilitation and interlingual homograph inhibition effects. A number of studies have shown that (regardless of any priming), in lexical decision experiments, participants tend to respond more quickly to cognates than single-language control words, but more slowly or equally quickly to interlingual homographs compared with control words (for a review, see e.g. Dijkstra, 2005; Dijkstra & Van Heuven, 2012; Poort & Rodd, 2017a; Dijkstra & Van Heuven, 2018; Poort & Rodd, 2019b). Using a semantic relatedness task, we previously found a different pattern of results, however: we observed a strong interlingual homograph inhibition effect, but found *no* evidence for a cognate facilitation effect (Poort & Rodd, 2019b).

This pattern of results across these two tasks is most consistent with a 'semantic settling' account of lexical access (Rodd et al., 2004; Rodd, 2020) in which familiar word meanings correspond to stable states within a high-dimensional lexical-semantic space. Under this view, word-meaning access is characterised as a settling process during which the system takes time to resolve on a single familiar meaning (Armstrong & Plaut, 2016). We previously hypothesised that task differences in the processing of cognates and interlingual homographs arise because these tasks tap into different stages of this settling process (Poort & Rodd, 2019b). In particular, this hypothesis was informed by drawing analogies with research with monolinguals showing that the processing of different types of semantically ambiguous words also varies across tasks (Rodd et al., 2002, 2004; Beretta et al., 2005; Hino et al., 2006; Klepousniotou & Baum, 2007; Armstrong & Plaut, 2008, 2016).

Specifically, we drew a comparison between cognates and polysemous words, which are ambiguous within a given language (e.g. *run* evokes a different sense in each of the phrases *the athlete/politician/film/river runs*; Rodd, 2018). We noted that, like a polysemous word, the two (L1 and L2) interpretations of a cognate are highly similar, but often not entirely identical, as in the case of the cognate *alarm*, which in both Dutch and English is used to refer to a warning signal and the state of apprehension or fear immediately following the perception of such a signal. In English, however, the word *alarm* is also used to refer to the device that wakes a person up in the morning, which in Dutch is called a *wekker*. Like cognates, polysemous words are known to show a processing benefit (relative to unambiguous words) in monolingual lexical

decision tasks in which participants do not need to settle on one specific sense in order to make a decision, which can be made on the basis of a general assessment of the familiarity of such words. Indeed, a benefit can arise from their relatively rich semantic representations (Rodd et al., 2002). Crucially, such ambiguity between related senses can delay processing in more explicitly semantic tasks that require participants to settle on a specific semantic representation (i.e. to resolve the ambiguity between the different senses; Rodd et al., 2002).

We further suggested that interlingual homographs, in contrast, are more akin to homonyms, which have multiple unrelated meanings within a language (e.g. *bark* meaning ‘the covering of a tree’ or ‘the sound a dog makes’). Monolinguals typically show a strong processing cost for these words in tasks that involve semantic processing (Hino et al., 2006) that is driven by strong competition between their unrelated and mutually inconsistent meanings (Rodd, 2020). This is consistent with the findings of slower responses by bilinguals to interlingual homographs in a semantic relatedness task (Poort & Rodd, 2019b), which is most likely also due to competition between the words’ unrelated, and mutually inconsistent, meanings. When a decision can be made based on a general assessment of the familiarity of the homonym or interlingual homograph (e.g. in lexical decision tasks), both monolinguals and bilinguals do not consistently exhibit such a large processing disadvantage (or at all).

The participants in the current experiment completed the semantic relatedness task in their native language instead of their second language as previously (Poort & Rodd, 2019b). The current experiment, therefore, allowed us to provide convergent evidence on this issue, by testing the prediction from this semantic settling account that our previous findings should also replicate in the participants’ native language. We predicted we would observe slower responses to interlingual homographs compared with translation equivalents, but that responses to cognates should not differ from these unambiguous word forms. We also tested an additional prediction, informed by research using lexical decision and picture naming tasks and eye-tracking during reading of a novel, that suggests that the cognate facilitation effect is greater in L2 than L1 (Kroll et al., 1999; Van Hell & Dijkstra, 2002; Cop et al., 2016). Although we did not expect to find a cognate effect, we predicted that the same logic would apply to interlingual homographs and, therefore, that when participants were tested in their native language (as in the current experiment), the interlingual homograph inhibition effect would be smaller than when participants were tested in their second language (as in the previous experiment).

In summary, the primary aim of this experiment was to obtain additional evidence regarding the high level of interaction between the words in a bilingual’s two languages by testing whether recent experience with ambiguous word forms in their L2 can influence subsequent processing of these items in their L1 and, secondary to that, whether L2-to-L1 priming effects will be greater than L1-to-L2 priming effects. In addition, data from the unprimed conditions of this experiment will speak to the debate concerning whether semantic settling accounts that were derived to



explain processing of ambiguous word forms in monolinguals can also account for how word form ambiguity affects bilingual speakers (Rodd, 2020), by testing whether we can replicate the presence of an interlingual homograph inhibition effect and lack of a cognate facilitation effect in a semantic relatedness judgements task and, secondary to that, whether the size of the interlingual homograph inhibition effect will be smaller in L1 than L2.

## 2. Methods

This experiment was pre-registered using the Center for Open Science's Preregistration Challenge template on the Open Science Framework ([osf.io/wk73f](https://osf.io/wk73f)). All deviations from the pre-registration are noted. The stimuli, data, processing and analysis scripts and analysis output can be found on the Open Science Framework ([osf.io/2swyg/](https://osf.io/2swyg/)).

### 2.1 Participants

#### 2.1.1 Sample size rationale

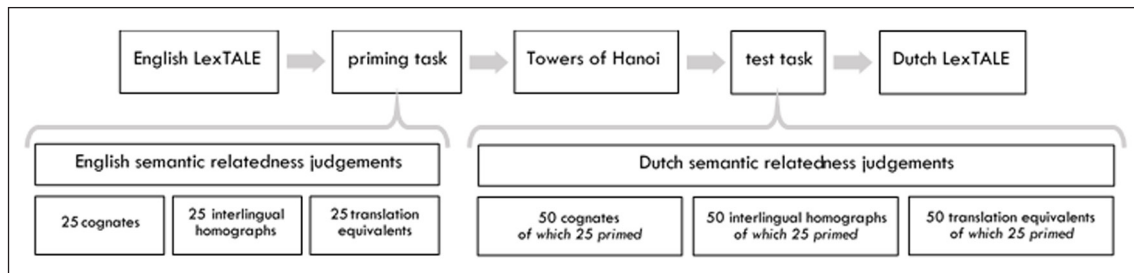
Our sample size was constrained and informed by two factors: money and a smallest-effect-size-of-interest power analysis. The power analyses were conducted in R (version 3.4.3; R Core Team, 2017) using the *simr* package (Green & MacLeod, 2016), based on the data from our previous experiment (Poort & Rodd, 2019b, Exp. 2). We focussed on the critical priming effects in the reaction time analyses (cognates: primed vs. unprimed; interlingual homographs: primed vs. unprimed) and simulated priming effects of approximately 20 ms for the cognates and interlingual homographs. These power curves (see the Supplementary materials, Appendix A) indicated that at least 60 participants were required to achieve 80% power for the simple effects analyses and between 100 and 120 to achieve 80% power for the two  $2 \times 2$  analyses that contrasted the priming effects for the cognates and interlingual homographs to the translation equivalents. As our budget allowed us to recruit 120 participants, we set this as our recruitment target sample size, allowing for a typical exclusion rate of 15% (based on previous experience) to obtain an analysis sample size of at least 100 participants. This would provide 95–100% power for the simple effects analyses and 75–90% for the two  $2 \times 2$  analyses.<sup>1</sup>

#### 2.1.2 Participant characteristics

A total of 120 participants were recruited through Prolific ([www.prolific.co](https://www.prolific.co)). The participants gave informed consent and were paid £7.50 for their participation. The UCL Experimental Psychology Ethics Committee provided approval of our study protocol (Project ID: EP/2017/009).

---

<sup>1</sup> We also set a time-limit of 30 days on recruitment, after which we would terminate recruitment regardless of our achieved sample size, but this turned out to be unnecessary.



**Figure 1:** Broad overview of the experimental procedure (top row) and stimuli included in the priming and testing tasks (bottom row). Note that details regarding subversions, counterbalancing and fillers are not included.

As for our previous experiment (Poort & Rodd, 2019b; Exp. 2), participants had to meet the following eligibility criteria: (1) age 18–50, (2) Dutch/Belgian national currently resident in the Netherlands or Belgium, (3) native speaker of Dutch,<sup>2</sup> (4) fluent speaker of English<sup>3</sup> and (5) not diagnosed with any language disorder(s). Nine participants did not meet all of these eligibility criteria, according to their responses to the demographics and language background questionnaire.<sup>4</sup> Although we did not specify this in our pre-registration, we contacted these participants, as these answers conflicted with what they had indicated on the Prolific website. Eight participants responded that they had made a mistake and did meet the criteria, so they were included in the final sample. The ninth participant was excluded.

Participants were further excluded from the analyses if they did not meet all of the following data quality checks: (1) at least 80% correct on both the priming and testing tasks, (2) a score of at least 50% on both of the two language proficiency measures (the Dutch and English LexTALES; see section 2.2) and (3) an average priming delay of less than 30 minutes. Thirteen participants were excluded based on these criteria. We then compared each participant’s testing task

<sup>2</sup> Native language was defined as: “Your native language is the language that you have learnt from birth from at least one of your parents or guardians. You can have more than one native language, for example if your parents or guardians both speak a different language with you.” Two participants had a second native language that was not English (Arabic and Moroccan).

<sup>3</sup> Language fluency was defined as: “You speak a language fluently if you can have a conversation in that language with ease and if you can write a letter (or other piece of text) in that language without difficulty.” Before starting the experiment, participants had to indicate they agreed with a statement that they met the eligibility criteria, including that they spoke English fluently according to this definition. In the demographics questionnaire they were also explicitly asked if they spoke English (or any other languages) fluently (according to the same definition) or as a native language (according to the definition in the previous footnote). Four participants considered English to be one of their native languages. Eighteen participants spoke at least one other language fluently (French, German, Mandarin, Norwegian, Portuguese, Spanish and/or Swedish).

<sup>4</sup> In our pre-registration we stated that participants who did not meet the criteria would not be allowed to continue with the experiment. We later found out that this is against Prolific policy, so participants who did not meet the eligibility criteria were still able to finish the experiment.



performance on the target items (cognates, interlingual homographs and translation equivalents) to the grand mean of all participants to identify statistical outliers. All participants had performed within three standard deviations of the mean ( $M = 87.7\%$ ,  $SD = 5.2\%$ ), so no further exclusions were made.

The 106 participants included in the analysis (73 males, 32 females, 1 non-binary;  $M_{age} = 24.0$  years,  $SD_{age} = 5.9$  years) started learning English from an average age of 7.9 ( $SD = 2.8$  years) and had an average of 16.0 years of experience with English ( $SD = 6.3$  years). On average, the participants rated their proficiency a 9.2 out of 10 in Dutch ( $SD = 0.7$ ) and an 8.5 in English ( $SD = 0.8$ ). A two-sided paired  $t$ -test showed this difference to be significant [ $t(105) = 8.913$ ,  $p < .001$ ]. These self-ratings were confirmed by their high LexTALE scores in both languages, which a two-sided paired  $t$ -test showed were also, though only slightly, higher in Dutch [Dutch:  $M = 87.5\%$ ,  $SD = 5.9\%$ ; English:  $M = 83.4\%$ ,  $SD = 9.4\%$ ;  $t(105) = 4.456$ ,  $p < .001$ ].

## 2.2 Design and procedure

This experiment employed the same mixed design as before (Poort & Rodd, 2019b, Exp. 2). Priming was a within-participants/within-items factor: for each participant, half of the items were primed while the others were unprimed. There were two versions of the experiment such that participants saw each item only once but across participants items occurred in both the primed and unprimed conditions. Word type was a within-participants/between-items factor: participants saw words from all three word types, but each item of course belonged to only one word type.

The experiment was created and conducted using Gorilla Experiment Builder ([www.gorilla.sc](http://www.gorilla.sc)) (Anwyl-Irvine et al., 2020). Participants first completed a self-report demographics questionnaire in English. The experiment then comprised five separate tasks: (1) the English version of the LexTALE (Lemhöfer & Broersma, 2012), (2) the English semantic relatedness prime task (mean duration in mm:ss: 09:44), (3) a filler task included to create a delay between priming and test, namely the Towers of Hanoi task (instructions in Dutch; maximum duration set to 4 minutes), (4) the Dutch semantic relatedness test task (mean duration: 11:04) and (5) the Dutch version of the LexTALE (Lemhöfer & Broersma, 2012). The experiment ended with a self-report language background questionnaire and debrief statement (in Dutch). Across participants, Dutch semantic relatedness judgements to primed items were made on average 15 minutes and 9 seconds after they were primed in the English semantic relatedness task ( $SD = 02:11$ ;  $range = 12:10-25:24$ ). The tasks are described briefly below. For full details, see Poort and Rodd (2019b). The experiment is also available to preview and clone via Gorilla Open Materials ([app.gorilla.sc/openmaterials/245694](http://app.gorilla.sc/openmaterials/245694)). Finally, **Figure 1** presents an overview of the experimental procedure and breakdown of stimuli in the priming and test tasks.

### **2.2.1 Prime task: English semantic relatedness judgements**

This task served to prime the cognates, interlingual homographs and translation equivalents in sentence contexts. Participants were instructed to read each sentence<sup>5</sup> and, to ensure they comprehended the prime sentences, to indicate with button presses whether a subsequent probe was semantically related to the sentence. The 50 target sentences for each of the three word types were pseudorandomly divided into two sets of 75, matched for all key variables and prime sentence length, for use in the two versions of the experiment. Including the 24 primed filler items (see 2.3), participants read a total of 99 sentences, half followed by related probes and half by unrelated probes. The sentences were divided into four blocks with a 15-second break in between blocks. Participants read the sentences at their own pace, but each sentence was displayed for at least 1,000 ms and at most 4,000 ms. As soon as they finished reading the sentence, they pressed the spacebar to continue to the probe, which remained on screen until the participant responded or until 2,000 ms passed. The inter-trial interval was 500 ms.

### **2.2.2 Towers of Hanoi task**

This task served to introduce a delay between priming and testing, while minimising exposure to additional linguistic material. The Towers of Hanoi is a puzzle in which discs of progressively smaller sizes must be moved from one peg to another in as few moves as possible, never placing a larger disc on top of a smaller one. Participants were given 4 minutes to complete as many puzzles as they could, starting with a puzzle with three discs and three pegs. Every subsequent puzzle had the same number of pegs but one disc more than the previous puzzle. In this web-based implementation, participants moved the discs from peg to peg using the mouse. The instructions were presented in Dutch to minimise any general language switching cost on the Dutch semantic relatedness task.

### **2.2.3 Test task: Dutch semantic relatedness judgements**

During the Dutch semantic relatedness task, the participants saw all 150 related target-probe pairs (“yes”-responses) and 150 unrelated filler-probe pairs (“no”-responses) and were asked to indicate, by means of a button press, as quickly and accurately as possible, whether the word they saw first was related in meaning to the word they saw second. The 300 pairs were divided into six blocks with a 15-second break in between blocks. During each trial, the target or filler item appeared first on screen and remained for 200 ms. The probe appeared after a 50-ms blank screen and remained on screen until the participant responded or until 2,000 ms passed.

---

<sup>5</sup> As is typical in such experiments, participants were not given explicit instruction to read silently, but given their high level of proficiency it is assumed that participants did not read the sentences out loud.

**Table 1:** Means (and standard deviations) for the Dutch and English characteristics, similarity ratings and prime sentence length of the cognates, interlingual homographs and translation equivalents (and fillers). *Frequency* refers to the SUBTLEX frequency of the word in occurrences per million (Brysbaert & New, 2009, for English; Keuleers et al., 2010, for Dutch). *Log10(frequency)* refers to the SUBTLEX log-transformed raw word frequency ( $\log_{10}[\text{raw frequency} + 1]$ ). *OLD20* refers to Yarkoni et al.'s (2008) measure of orthographic complexity of a word, expressed as its mean orthographic Levenshtein distance to its 20 closest neighbours. The similarity ratings were obtained by Poort and Rodd (2019a) and were given on a scale from 1 (not at all similar) to 7 ((almost) identical). *Sentence length* refers to the length of the English prime sentences shown during the semantic relatedness task.

	DUTCH CHARACTERISTICS			ENGLISH CHARACTERISTICS			SIMILARITY RATINGS			SENTENCE LENGTH	
	frequency	log10(frequency)	word length	frequency	log10(frequency)	word length	OLD20	meaning	spelling		pronunciation
identical cognates	39.0 (59.4)	2.92 (0.49)	4.48 (0.97)	44.8 (57.0)	3.10 (0.48)	4.48 (0.97)	1.59 (0.31)	6.85 (0.22)	7.00 (0.00)	5.87 (0.68)	9.64 (1.76)
interlingual homographs	39.2 (82.9)	2.70 (0.68)	4.30 (0.93)	56.2 (108)	2.98 (0.66)	4.30 (0.93)	1.46 (0.32)	1.15 (0.28)	7.00 (0.01)	5.53 (0.79)	9.60 (1.59)
translation equivalents	27.8 (26.3)	2.88 (0.46)	4.68 (0.87)	30.2 (26.0)	3.02 (0.41)	4.46 (0.93)	1.57 (0.31)	6.86 (0.19)	1.12 (0.24)	1.11 (0.21)	9.56 (1.45)
fillers	42.6 (96.5)	2.75 (0.67)	4.57 (1.23)	—	—	—	—	—	—	—	—

**Table 2:** Examples of English prime sentences for a cognate, interlingual homograph and translation equivalent, along with their original Dutch prime sentences. Note that the participants did not see the original Dutch sentences at any point during the experiment.

	Dutch word form	English word form	English prime sentence	original Dutch prime sentence
cognate	wolf	wolf	The dog is a domesticated subspecies of the <b>wolf</b> .	De hond is een gedomesticeerde ondersoort van de <b>wolf</b> .
interlingual homograph	angel	angel	Kate said Gary was a true <b>angel</b> .	Alleen vrouwelijke bijen hebben een <b>angel</b> .
translation equivalent	wortel	carrot	You can always make a donkey happy with a <b>carrot</b> .	Een ezel kun je altijd blij maken met een <b>wortel</b> .

Participants received a warning that they were responding too slowly if they had not responded 1,500 ms after the probe first appeared. This warning remained on screen for 500 ms, during which time the participant could still respond. The inter-trial interval was 1,000 ms.

## 2.3 Materials

The materials we used were adapted from our previous experiment (Poort & Rodd, 2019b, Exp. 2). The target items themselves were not changed, but the filler items, prime sentences, probes for the prime task and probes for the test task were either translated or newly created according to the original criteria.

### 2.3.1 Targets

The target items we selected for our previous experiment (Poort & Rodd, 2019b, Exp. 2) came from our database of cognates, interlingual homographs and translation equivalents (Poort & Rodd, 2019a). In Dutch, the items ranged in frequency from 0.57 to 477.21 occurrences per million (according to the SUBTLEX-NL database; Keuleers et al., 2010), were between 3 and 7 letters long and had OLD20 values (Yarkoni et al., 2008) between 1 and 2.70. In English, the items ranged in frequency from 0.98 to 590.69 occurrences per million (according to the SUBTLEX-US database; Brysbaert & New, 2009), were between 3 and 8 letters long and had OLD20 values between 1 and 2.80.

In our previous experiment, we had matched the three word types only on *English* log-transformed frequency, word length and OLD20. Since we planned to analyse responses to these items in a *Dutch* task, we used independent-samples two-tailed Welch's *t*-tests to also compare

Dutch log-transformed frequency, word length and OLD20 between the three word types. These showed that the word types were less well matched in Dutch: there was a marginally significant difference in log-transformed frequency between the cognates and interlingual homographs ( $p = .062$ ), as well as a significant difference in length between the interlingual homographs and translation equivalents ( $p = .037$ ) and in OLD20 between the cognates and interlingual homographs ( $p = .001$ ) and the cognates and translation equivalents ( $p = .036$ ). As noted in our pre-registration, we therefore ran the  $2 \times 3$  analysis both with and without these three variables included as covariates to determine whether these differences had a considerable impact on our results. **Table 1** lists means and standard deviations per word type for each of the matching criteria (and word frequency in occurrences per million) for both Dutch and English, as well as the meaning, spelling and pronunciation similarity ratings obtained by Poort and Rodd (2019a).

### 2.3.1.1 Prime sentences and probes

For the cognates and translation equivalents, the Dutch priming sentences that we had used previously (Poort & Rodd, 2019b, Exp. 2) were translated to English (see **Table 2** for examples). For the interlingual homographs, new English prime sentences were written according to the same criteria we used before (see Poort & Rodd, 2019b, Exp. 2). These English sentences served to prime the items in the English semantic relatedness judgements task. (Note that the participants did not see the Dutch originals at any point during the experiment.) Independent-samples two-tailed Welch's  $t$ -tests showed no significant differences between word types in terms of prime sentence length (all  $ps > .8$ ). Whenever possible, we also used the English translations of the Dutch probes we had used previously, for use in the English prime task; where necessary, new English probe words were chosen.

### 2.3.1.2 Test probes

For the test task, each target item was paired with a second probe word, but now in Dutch (e.g. *code-geheim*), which was always semantically related to it. Again, for the cognates and translation equivalents, wherever possible we used the Dutch translations of the English probes we used previously (Poort & Rodd, 2019b, Exp. 2). We assigned new probes to the interlingual homographs (e.g. *angel-heaven* became *angel-wesp*, where *wesp* means 'wasp'). The probes were of roughly equal frequency, length and orthographic complexity as the targets themselves. They ranged in frequency from 0.02 to 482.99 occurrences per million, were between 3 and 9 letters long and had OLD20 values between 1 and 3.35. Means and standard deviations of these variables for each set of probes per word type can be found in **Table 3**. The sets of probes for the three word types did not significantly differ from each other in terms of log-transformed frequency, word length or OLD20 (all  $ps > .1$ ).

**Table 3:** Means (and standard deviations) of the Dutch characteristics of the probes for the semantic relatedness task for the cognates, interlingual homographs, English controls and the fillers. *Frequency* refers to the SUBTLEX frequency of the word in occurrences per million (Brysbaert & New, 2009). *Log10(frequency)* refers to the SUBTLEX log-transformed raw word frequency ( $\log_{10}[\text{raw frequency} + 1]$ ). *OLD20* refers to Yarkoni et al.'s (2008) measure of orthographic complexity of a word, expressed as its mean orthographic Levenshtein distance to its 20 closest neighbours.

	DUTCH CHARACTERISTICS OF THE PROBES			
	frequency	log10(frequency)	word length	OLD20
cognates	58.6 (118)	2.79 (0.75)	5.74 (1.47)	1.82 (0.60)
interlingual homographs	34.7 (51.8)	2.71 (0.78)	5.58 (1.65)	1.70 (0.48)
translation equivalents	30.6 (45.2)	2.60 (0.76)	5.36 (1.66)	1.64 (0.55)
fillers	28.5 (76.9)	2.54 (0.67)	5.54 (1.56)	1.81 (0.59)

*Note.* OLD20 information was not available for the probe *examen* for the cognate *test–test* and the probe *zoethout* for the interlingual homograph *drop–drop*.

### 2.3.2 Fillers

The test task also included 150 filler items paired with unrelated probes. Again, whenever possible we used the Dutch translations of the fillers included in our previous experiment (e.g. *pinda–jazz* and *mug–kantlijn*; Poort & Rodd, 2019b, Exp. 2). If translating was not possible, we selected new words using Dutch random word generator websites. The 150 fillers included 15 additional cognates, interlingual homographs and English controls each that were also paired with unrelated probes, and eight of these fillers for each type were primed. These manipulations ensured there was no strong relationship between either word type or priming on the one hand and the response that participants were required to make on the other. Finally, we ensured that no target, filler, probe or their translations occurred twice in the experiment.

## 3. Results

All analyses were carried out in R (version 3.4.4; R Core Team, 2018) using the *lme4* package (version 1.1-17; Bates et al., 2015), following guidelines proposed by Barr et al. (2013) for model fitting (with the same amendments as previously; Poort & Rodd, 2019b) and using Type III Sums of Squares likelihood ratio tests to determine significance. Unless specified, the significance level was set at .05. Reaction times were analysed using the *lmer()* function with the default Nelder-Mead optimiser; accuracy data were analysed using the *glmer()* function with the *bobyqa*



optimiser. The output for all analyses can be found in the `experimentOverview.xlsx` document in our OSF project. This section reports the results of our planned confirmatory analyses. No exploratory analyses are reported.

Three items (the cognates *motto–motto* and *nest–nest* and the translation equivalent *bot–rude*) were excluded from the analyses, as accuracy on the Dutch semantic relatedness task (43.4%, 50.0%, and 53.7%, respectively) was more than three standard deviations below the items' word type mean. After excluding these items, we re-checked the matching of the word types and found that the difference in log-transformed frequency between the cognates and interlingual homographs was now significant ( $p = .047$ ). This did not change our analysis, as we had already planned to include log-transformed frequency in the  $2 \times 3$  analysis (see below).

Finally, note that the figures display the (harmonic) participant means, while the effects (and marginal means, *MM*) reported in the text were derived from the estimates of the fixed effects provided by the model summary (calculations are provided in the Excel document `experimentOverview.xlsx` in our OSF project). In figures that show harmonic participant means, we first calculated condition means for each participant (the dots) of the inverse-transformed reaction times and then transformed these back to the original scale before calculating an overall average per condition (the horizontal bars). For example, if a condition mean of inverse-transformed RTs for a particular participant was 1.63, their harmonic condition mean was  $1000/1.63 = 613$  ms and this is plotted as a dot in the figures. This can result in small discrepancies between the numbers reported in the text and the figures.

### 3.1 Prime task: English semantic relatedness task

High accuracy ( $M = 91.8\%$ ,  $SD = 3.9\%$ ,  $range = 80.8\%–100\%$ ) confirmed participants had processed the sentence meanings correctly. To determine whether any of the observed effects of priming in the testing task could have been due to differences between the word types or priming versions at the time of priming, a  $3 \times 2$  analysis was conducted on the accuracy data with the fixed factors word type (3 within-participants/between-items levels: cognate, interlingual homograph, translation equivalent) and priming version (2 between-participants/within-items levels: version 1, version 2). The maximal model converged for this analysis and included a random intercept by participants and items as well as a by-participants random slope for word type and a by-items random slope for version. Correlations between the random slopes and intercepts were also included. This analysis revealed that the main effect of word type was significant [ $\chi^2(2) = 10.45$ ,  $p = .005$ ], but the main effect of version was not [ $\chi^2(1) = 0.049$ ,  $p = .825$ ,  $\Delta = 0.2\%$ ], nor was the interaction between word type and version [ $\chi^2(1) = 3.784$ ,  $p = .151$ ].

As per our pre-registration, because the main effect of word type was significant, we conducted three pairwise comparisons to directly compare the cognates, interlingual homographs and translation equivalents to each other. The maximal model also converged for these three analyses and included a random intercept by participants and items and a by-participants random slope for word type. The  $p$ -values for the pairwise comparisons were compared against a Bonferroni-corrected  $\alpha$  of .0167. Probes were more accurately judged to be (un)related to sentences containing a cognate ( $MM = 98.8\%$ ) than a translation equivalent [ $MM = 95.6\%$ ;  $\chi^2(1) = 10.78, p = .001, \Delta = 3.1\%$ ] or an interlingual homograph [ $MM = 96.6\%$ ;  $\chi^2(1) = 6.394, p = .011, \Delta = 2.2\%$ ]. There was no significant difference between sentences containing an interlingual homograph ( $MM = 96.6\%$ ) and a translation equivalent [ $MM = 95.9\%$ ;  $\chi^2(1) = 0.378, p = .539, \Delta = 0.8\%$ ].

### 3.2 Towers of Hanoi task

All participants completed at least one puzzle, confirming task engagement. On average, participants completed 2.4 puzzles ( $mode = 2, range = 1-4$ ).

## 3.3 Test task: Dutch semantic relatedness task

### 3.3.1 Analysis procedure

The same analysis procedure was employed for the reaction times and accuracy data. In all cases, positive effects of priming indicate a facilitative effect of priming (i.e. faster reaction times and higher accuracy for primed items), while negative effects indicate a disruptive effect of priming (i.e. slower reaction times and lower accuracy for primed items). Positive (negative) effects of word type indicate an advantage (disadvantage) for the first-named word type over the second-named word type. The significance level ( $\alpha$ -level) for all analyses was set at .05 unless otherwise noted.

Two fixed factors were included in the main  $2 \times 3$  analysis: priming (2 within-participant/within-items levels: unprimed, primed) and word type (3 within-participant/between-items levels: cognate, interlingual homograph, translation equivalent). The maximal random effects structure of this model included a correlated random intercept and random slope for word type, priming and their interaction by participants and a correlated random intercept and random slope for priming by items. This maximal model did not converge for the reaction time analysis, nor did any other model with a random slope, so we used a random intercepts-only model. For the accuracy analysis, the maximal model also did not converge. As per our pre-registration, we removed the correlations between the random effects, but the model still did not converge, so next we removed the random slope for priming by items, which was the random effect with the smallest variance in the maximal model. The random effects in this model were again allowed to correlate and this model converged. Furthermore, as per our pre-registration, the

$2 \times 3$  analysis was conducted both with and without the covariates item length, log-transformed frequency and OLD20, as there were significant differences between the three word types on these variables. These covariates were centred to have a mean of zero prior to including them in the model. Random slopes were not included for the covariates. The results of these analyses are only reported when the significance level of the effects of priming or word type (or their interaction) are different compared to the analysis without covariates. For transparency, the output is included in the `experimentOverview.xlsx` document on the OSF.

In addition, three  $2 \times 2$  analyses were conducted comparing the effect of priming for the cognates and interlingual homographs, the cognates and translation equivalents and the interlingual homographs and translation equivalents. The maximal random effects structure for these models included a correlated random intercept and random slope for word type, priming and their interaction by participants and a correlated random intercept and random slope for priming by items. For the reaction times, the maximal model again did not converge, so as per our pre-registration we removed the correlations between the random effects, after which the model did converge. For the accuracy analyses the maximal model converged. All  $p$ -values from these three analyses were compared against a Bonferroni-corrected  $\alpha$  of .0167.

To examine the effect of priming for each of the three word types separately, three simple effects analyses were conducted. The maximal model converged for both the reaction times and accuracy analysis and included a correlated random intercept and random slope for priming by both participants and items. The  $p$ -values for these three analyses were compared against a Bonferroni-corrected  $\alpha$  of .0167.

Finally, three pairwise comparisons were conducted on the unprimed data only, comparing the cognates, interlingual homographs and translation equivalents to each other. Only the unprimed trials were analysed as priming was expected to essentially increase both the cognate facilitation effect and the interlingual homograph inhibition effect. The maximal model converged for both the reaction times and accuracy analysis and included a correlated random intercept and random slope for word type by participants and a random intercept by items. The  $p$ -values for these three analyses were compared against a Bonferroni-corrected  $\alpha$  of .0167.

**Table 4** presents an overview of which effects were significant in which of the analyses we describe below. Naturally, given our hypotheses, our pre-registration also included analyses that compare the effects of priming and word type between the current experiment and our previous experiment (Poort & Rodd, 2019b, Exp. 2). These analyses revealed no significant differences between the two experiments. For brevity, we report these analyses in Appendix B of the Supplementary materials instead of in the main text. A summary of these results is presented in **Table 5**.

**Table 4:** Overview of the significant effects per analysis and dependent variable (reaction times and accuracy). Significant effects are significant at the  $\alpha$ -level specified for that analysis in the Analysis procedure. Effect sizes (unstandardised) are provided only for significant effects and only if the factor had no more than two levels. As before, positive (negative) effects of priming indicate a facilitative (disruptive) effect of priming (i.e. faster (slower) reaction times and higher accuracy for primed items); positive (negative) effects of word type indicate an advantage (disadvantage) for the first-named word type over the second-named word type. ‘%pts’ means percentage points.

	REACTION TIMES			ACCURACY		
	priming	word type	interaction	priming	word type	interaction
$2 \times 3$	sig. (-5 ms)	sig.	n.s.	n.s.	sig.	n.s.
$2 \times 2s$						
cognates vs. interlingual homographs	n.s.	sig. (59 ms)	n.s.	n.s.	sig. (8.7 %pts)	n.s.
cognates vs. translation equivalents	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
interlingual homographs vs. translation equivalents	sig. (-7 ms)	sig. (-65 ms)	n.s.	n.s.	sig. (9.1 %pts)	n.s.
<b>simple effects</b>						
cognates	n.s.	—	—	n.s.	—	—
interlingual homographs	sig. (-16 ms)	—	—	n.s.	—	—
translation equivalents	n.s.	—	—	n.s.	—	—
<b>pairwise comparisons</b>						
cognates vs. interlingual homographs	—	sig. (50 ms)	—	—	sig. (8.4 %pts)	—
cognates vs. translation equivalents	—	n.s.	—	—	n.s.	—
interlingual homographs vs. translation equivalents	—	sig. (-56 ms)	—	—	sig. (9.7 %pts)	—

**Table 5:** Overview of the significant effects per analysis and dependent variable (reaction times and accuracy) for the between-experiment analyses (see Supplementary materials, Appendix B). Significant effects are significant at the  $\alpha$ -level specified for that analysis in the Analysis procedure outlined in the Supplementary materials, Appendix B. Effect sizes (unstandardised) are provided only for significant effects. Positive (negative) effects of priming indicate a facilitative (disruptive) effect of priming (i.e. faster (slower) reaction times and higher accuracy for primed items); positive (negative) effects of word type indicate an advantage (disadvantage) for the first-named word type over the second-named word type. Effects of experiment were never significant. ‘%pts’ means percentage points.

	REACTION TIMES			ACCURACY		
	priming/ word type	experiment	interaction	priming/ word type	experiment	interaction
<b>priming effects</b>						
cognates	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
interlingual homographs	sig. (-14 ms)	n.s.	n.s.	n.s.	n.s.	n.s.
<b>word type effects</b>						
cognates vs. interlingual homographs	sig. (46 ms)	n.s.	n.s.	sig. (7.5 %pts)	n.s.	n.s.
cognates vs. translation equivalents	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
interlingual homographs vs. translation equivalents	sig. (-50 ms)	n.s.	n.s.	sig. (-8.5 %pts)	n.s.	n.s.

### 3.3.2 Reaction times

Reaction times (RTs) faster than 300 ms or slower than 1500 ms were discarded (1.1% of the data), as were RTs for incorrect trials and trials that participants had not responded to (11.0% of the remaining data). The RTs were inverse-transformed (inverse-transformed RT =  $1000/\text{raw RT}$ ) to remedy violations of the assumption of normality and to be consistent with our previous approach (Poort & Rodd, 2019b, Exp. 2). After inverse-transforming the RTs, any inverse-transformed RTs were removed that were more than three standard deviations above or below a participant’s mean inverse-transformed RT (0.1% of the remaining data).

**Figure 2** shows the reaction times per word type and priming condition. The violin is a symmetrical density plot showing the distribution of reaction times. Each dot in the figure represents a participant’s condition mean. The grey lines connect primed and unprimed means

from the same participant for each word type. The horizontal bar and the number above it provide the mean across all participants for that condition. The figure suggests that priming was not effective for the cognates or translation equivalents, but hints at a disruptive effect of priming for the interlingual homographs. The figure further shows that there was considerable between-participant variation in reaction times and in the size of the priming effect for the three word types.

### 3.3.2.1 2×3

The main effect of priming was significant [ $\chi^2(1) = 4.339, p = .037$ ], with participants responding on average 5 ms more slowly to primed items than to unprimed items. The main effect of word type was significant [ $\chi^2(2) = 37.01, p < .001$ ], indicating that there was a difference in reaction times between the three word types. The interaction between word type and priming was significant [ $\chi^2(2) = 7.057, p = .029$ ].

### 3.3.2.2 2×2s

The 2×2 analysis that included the cognates and interlingual homographs showed no significant main effect of priming at the Bonferroni-corrected  $\alpha$ -level of .0167 [ $\chi^2(1) = 5.642, p = .018, \Delta = -7$  ms]. The main effect of word type was significant [ $\chi^2(1) = 20.74, p < .001$ ], with participants responding on average 59 ms more quickly to the cognates than to the interlingual homographs. The interaction between word type and priming was also not significant [ $\chi^2(1) = 4.929, p = .026; \alpha = .0167$ ].

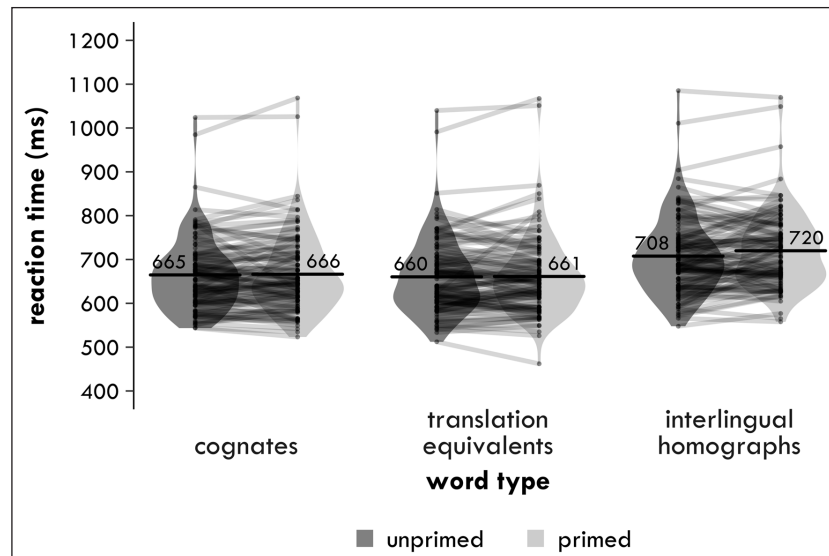
The 2×2 analysis that included the cognates and translation equivalents showed no significant effects of priming [ $\chi^2(1) = 0.017, p = .897, \Delta = -0.3$  ms], word type [ $\chi^2(1) = 0.460, p = .498, \Delta = -6$  ms] or the interaction between priming and word type [ $\chi^2(1) = 0.001, p = .998$ ].

In the 2×2 analysis that included the interlingual homographs and translation equivalents, the main effect of priming was significant [ $\chi^2(1) = 5.782, p = .016$ ], with participants responding on average 7 ms more slowly to primed items than to unprimed items. The main effect of word type was also significant [ $\chi^2(1) = 27.02, p < .001$ ], with participants responding on average 65 ms more slowly to the interlingual homographs than to the translation equivalents. The interaction between word type and priming was not significant [ $\chi^2(1) = 5.143, p = .023$ ].

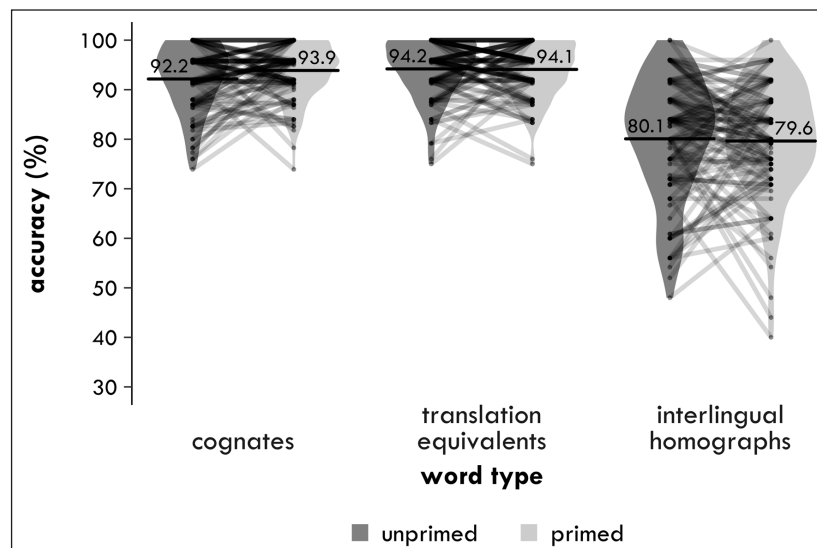
### 3.3.2.3 Simple effects of priming

The simple effect of priming was not significant for the cognates [ $\chi^2(1) = 0.064, p = .800, \Delta = -1$  ms] or for the translation equivalents [ $\chi^2(1) = 0.009, p = .926, \Delta = -0.3$  ms]. However, the effect of priming was significant for the interlingual homographs [ $\chi^2(1) = 10.21, p = .001$ ], with participants responding on average 16 ms more slowly to primed interlingual homographs than to unprimed interlingual homographs.





**Figure 2:** Harmonic participant means of the current experiment's inverse-transformed Dutch semantic relatedness reaction times (in milliseconds) by word type (cognates, translation equivalents, interlingual homographs; x-axis) and priming (unprimed, dark grey; primed, light grey). Each point represents a condition mean for a participant with lines connecting means from the same participant. Each bar provides the mean across all participants in that condition. The violin is a symmetrical density plot.



**Figure 3:** Participant means of the current experiment's Dutch semantic relatedness accuracy (in percentages correct) by word type (cognates, translation equivalents, interlingual homographs; x-axis) and priming (unprimed, dark grey; primed, light grey). Each point represents a condition mean for a participant with lines connecting means from the same participant. Each bar provides the mean across all participants in that condition. The violin is a symmetrical density plot.

### 3.3.2.4 Pairwise comparisons of unprimed trials

The pairwise comparisons on the unprimed trials revealed a significant difference between the cognates and the interlingual homographs [ $\chi^2(1) = 16.93, p < .001$ ], with participants responding on average 50 ms more quickly to the cognates. There was also a significant difference between the interlingual homographs and the English controls [ $\chi^2(1) = 21.58, p < .001$ ], with participants responding on average 56 ms more slowly to the interlingual homographs. There was no significant difference between the cognates and the English controls [ $\chi^2(1) = 0.416, p = .519, \Delta = -5$  ms].

### 3.3.3 Accuracy

In line with the trimming procedure for the reaction times, any trials with RTs faster than 300 ms or slower than 1500 ms were removed.

**Figure 3** shows the percentages correct per word type and priming condition. As in **Figure 2**, the violin is a symmetrical density plot showing the distribution of percentages correct. Each dot in the figure represents a participant's condition mean. The grey lines connect primed and unprimed means from the same participant for each word type. The horizontal bar and the number above it provide the mean across all participants for that condition. The figure suggests that priming was not effective for the translation equivalents, but hints at a small facilitative effect of priming for the cognates and a small disruptive effect for the interlingual homographs. The figure further shows that there was a ceiling effect, especially for the cognates and translation equivalents and that, again, there was considerable between-participant variation in accuracy and in the size of the priming effect for the three word types. Note that the analyses below modelled the accuracy of individual responses (i.e. 1 if correct or 0 if incorrect) using logistic mixed effects models, not the participant means displayed in the figure.

#### 3.3.3.1 2×3

The main effect of priming was not significant [ $\chi^2(1) = 0.091, p = .764, \Delta = 0.1$  percentage points]. The main effect of word type was significant [ $\chi^2(2) = 34.07, p < .001$ ], indicating that there was a difference in accuracy between the three word types. The interaction between word type and priming was not significant [ $\chi^2(2) = 3.712, p = .156$ ].

#### 3.3.3.2 2×2s

In the 2×2 analysis that included the cognates and interlingual homographs, the main effect of priming was not significant [ $\chi^2(1) = 1.802, p = .180, \Delta = 1.0$  percentage points]. The main effect of word type was significant [ $\chi^2(1) = 19.53, p < .001$ ], with participants' accuracy being on average 8.7 percentage points higher for the cognates than the interlingual homographs. The interaction between word type and priming was also not significant [ $\chi^2(1) = 3.403, p = .065$ ].

In the  $2 \times 2$  analysis that included the cognates and translation equivalents, the main effect of priming was again not significant [ $\chi^2(1) = 0.076, p = .783, \Delta = 0.2$  percentage points], nor was the main effect of word type [ $\chi^2(1) = 0.407, p = .523, \Delta = -0.5$  percentage points] or the interaction between word type and priming [ $\chi^2(1) = 1.572, p = .210$ ].

In the  $2 \times 2$  analysis that included the interlingual homographs and translation equivalents, the main effect of priming was also not significant [ $\chi^2(1) = 0.369, p = .543, \Delta = -0.4$  percentage points]. The main effect of word type was significant [ $\chi^2(1) = 25.21, p < .001$ ], with participants' accuracy being on average 9.1 percentage points lower for the interlingual homographs than the English controls. The interaction between word type and priming was not significant [ $\chi^2(1) = 0.011, p = .916$ ].

### 3.3.3.3 Simple effects of priming

None of the simple effects of priming were significant [for the cognates:  $\chi^2(1) = 2.030, p = .154, \Delta = 1.1$  percentage points; for the interlingual homographs:  $\chi^2(1) = 0.002, p = .970, \Delta = -0.04$  percentage points; for the translation equivalents:  $\chi^2(1) = 1.185, p = .276, \Delta = -0.8$  percentage points].

### 3.3.3.4 Pairwise comparisons of unprimed trials

The pairwise comparisons on the unprimed trials revealed a significant difference between the cognates and the interlingual homographs [ $\chi^2(1) = 15.74, p < .001$ ], with participants' accuracy being on average 8.4 percentage points higher for the cognates. There was also a significant difference between the interlingual homographs and the English controls [ $\chi^2(1) = 26.80, p < .001$ ], with participants' accuracy being on average 9.7 percentage points lower for the interlingual homographs. There was no significant difference between the cognates and the English controls [ $\chi^2(1) = 2.000, p = .157, \Delta = -1.4$  percentage points].

## 4. Discussion

This experiment addressed several questions regarding how bilinguals process cognates and interlingual homographs in their L1 and L2. Chief among these was whether processing of these words in their L1 would be affected by having encountered them approximately 15 minutes earlier in their L2. We predicted that this cross-lingual priming manipulation would result in faster and more accurate semantic relatedness judgements for cognates but slower and less accurate ones for interlingual homographs. In contrast, priming was not predicted to affect processing of the translation equivalents (also known as noncognates). We found that priming indeed did not affect the translation equivalents, but neither did it result in a significant benefit for the cognates. For the interlingual homographs, our hypothesis was confirmed in part: participants responded 16 ms more slowly to pairs including a primed interlingual homograph (compared with pairs

including an unprimed homograph), but this significant priming effect did not carry over into the accuracy data nor was it significantly different from the null effects for the translation equivalents and the cognates.<sup>6</sup>

That we observed L2-to-L1 priming might be surprising, given that for many bilinguals their native language feels more stable and less susceptible to short-term changes than their second language. This finding supports the notion that the bilingual mental lexicon is indeed highly interactive, and that cross-lingual priming is bidirectional. On a daily basis, many bilinguals will regularly switch from speaking in one language to the other. This can result in language interference effects in many different experimental paradigms in both language production and comprehension (for a review, see Declerck & Philipp, 2015). In particular, as we have shown previously, it may lead to situations of cross-lingual priming that can have consequences for the ease with which bilinguals can access and process the meanings of individual words in those languages after a switch. We have now shown that this is not only the case when bilinguals switch from their native language to their second language, but also when they switch from their second to their native language. In addition, the fact that cross-lingual priming has been observed again using a semantic relatedness task as the test task adds support to the idea that such priming effects are more evident in semantic relatedness tasks than lexical decision tasks. Evidence for priming using lexical decision tasks has been inconsistent, perhaps due to the inconsistent need to fully access semantic representations during this task (Poort & Rodd, 2019b). This experiment highlights again the need for researchers in the field of bilingualism to move beyond using lexical decision tasks to study lexical access.

Based on findings from the monolingual word-meaning priming literature (Rodd et al., 2013; Betts, 2018), we had predicted that the priming effects for the cognates and interlingual homographs in this experiment would be larger than the priming effects observed in our previous experiment (Poort & Rodd, 2019b), in which these words were primed in the participants' L1 instead. Although numerically larger, the priming effect for the interlingual homographs that we observed here (16 ms) was not statistically larger than the effect previously found (10 ms). The priming effect for the cognates in the current experiment was numerically smaller and in the opposite direction (-1 ms) compared to the previous experiment (5 ms), but this difference was also not significant. This means that, in contrast to our hypothesis, we did not find any evidence for an asymmetry of cross-lingual priming.

---

<sup>6</sup> Keen readers will remember that there was a significant difference in priming task accuracy between the word types. Accuracy in the priming task was highest for the cognates, then the interlingual homographs and then the translation equivalents. We therefore think it unlikely that the priming effect for the interlingual homographs was caused by better processing at the time of priming and, analogously, that the lack of priming for the cognates was caused by worse processing.

While this finding is, therefore, not entirely in line with the observation that, in monolingual participants, the subordinate meaning of a homonym is just as likely or even more likely to interfere with processing of the dominant meaning as the dominant meaning is to interfere with processing of the subordinate meaning (Rodd et al., 2013; Betts, 2018), it does fit in with (but unfortunately does not clarify) previous research on (masked) translation and repetition priming that produced mixed findings regarding the strength of L1-to-L2 and L2-to-L1 priming (Gollan et al., 1997; Jiang, 1999; Jiang & Forster, 2001; Francis et al., 2003; Finkbeiner et al., 2004; Perea et al., 2008; Duyck & Warlop, 2009; Schoonbaert et al., 2009; Davis et al., 2010; Duñabeitia, Dimitropoulou, et al., 2010; Duñabeitia, Perea, et al., 2010; Dimitropoulou et al., 2011a, 2011b; Witzel & Forster, 2012; Nakayama et al., 2013; Wang, 2013; Chen et al., 2014; Wang & Forster, 2014; Lee et al., 2018). As discussed in Section 1, this may be due to the fact that our participants were relatively balanced bilinguals, but were still more proficient in their L1 than L2.

Since this research suggests that priming is more effective from L1 to L2 in bilinguals who are more proficient in their L1 than their L2 (Gollan et al., 1997; Jiang, 1999; Finkbeiner et al., 2004; Nakayama et al., 2013), while there is no evidence for a directional asymmetry in more balanced bilinguals (Perea et al., 2008; Davis et al., 2010; Duñabeitia, Dimitropoulou, et al., 2010; Duñabeitia, Perea, et al., 2010), it could be that priming will be greater from L2 to L1 in (simultaneous) bilinguals who are equally proficient in both languages. Given this, and the fact that this absence of evidence should also not be taken as evidence for the absence of such an asymmetry, we suggest that future experiments should investigate our hypothesis further by studying bilinguals who are at least equally proficient in their L2 as their L1.

Although this finding of cross-lingual priming adds to a growing literature concerning interaction between the lexical representations of words in a bilingual's different languages, the precise mechanism by which this priming occurs remains somewhat unclear. Early reports of word-meaning priming in monolingual speakers were taken as evidence that an encounter with one meaning of an ambiguous word (e.g. the 'tree covering' meaning of *bark*) resulted in a direct strengthening of this particular form-to-meaning mapping in the lexicon, making this meaning more available in the near future, relative to its alternative 'dog noise' meaning (Rodd et al., 2013). However, more recently it has been suggested that word-meaning priming does not reflect a direct change in the mental lexicon itself. In Gaskell et al.'s (2019) contextual binding account, they suggest that when homonyms (and other words) are encountered in sentence contexts, a temporary representation is formed in which the word is bound to the context in which it was experienced, in order to facilitate comprehension. They propose that this contextually bound representation of a word provides an additional source of information alongside long-term lexical knowledge that could influence the subsequent interpretation of that same word in the minutes (or hours) that follow exposure (Gaskell et al., 2019). Additional work is needed to discriminate

between the predictions of these different mechanistic accounts of how recent lexical experience can influence how ambiguous word forms are processed for both within-language and cross-lingual priming.

Our final question of interest concerned the unprimed trials only. Previously, we found that the cognate facilitation effect disappears when a semantic relatedness task is used rather than the more usual lexical decision task and, simultaneously, that the interlingual homograph inhibition effect is enhanced (Poort & Rodd, 2019b, Exp. 1). In line with our current hypothesis, our data also showed that there was a strong interlingual homograph inhibition effect (of 56 ms and 9.7 percentage points) but no cognate facilitation effect. We based the explanation of our findings on research in the monolingual domain on semantic ambiguity resolution (Rodd et al., 2002, 2004; Armstrong & Plaut, 2008, 2016; Rodd, 2018). We posited that the representation of a cognate in the mental lexicon consists of two or more related senses (since many cognates do not have identical meanings, but often differ in their nuances). Like the senses of a polysemous word, they facilitate processing when only a superficial impression of the word's meaning or word-likeness suffices (e.g. in lexical decision), but may compete to some extent for selection when a specific sense needs to be retrieved (e.g. in semantic relatedness). The representation of an interlingual homograph, in contrast, is thought to be linked to two distinct meanings, as for a homonym. These meanings may or may not compete when the situation does not require settling on a particular meaning (e.g. in lexical decision), but certainly will when one of these two meanings needs to be selected (e.g. in semantic relatedness). In other words, our replication of our previous finding (Poort & Rodd, 2019b, Exp. 1) provides further confirmation for the theory we outlined then.

We also predicted that, because the semantic relatedness task was performed in the participants' first language (Dutch), the interlingual homograph inhibition effect would be smaller in this experiment than in our previous experiments (Poort & Rodd, 2019b), where the effect was tested in the bilinguals' second language (English). This was not the case, and surprisingly, the interlingual homograph inhibition effect was numerically *larger* in the current experiment than the previous by 19 ms and 3.2 percentage points (56 ms and 9.7 percentage points compared with 37 ms and 6.5 percentage points). In line with this but also against our prediction, the difference between the cognates and interlingual homographs was not smaller in this experiment than in our previous experiment (50 ms and 8.4 percentage points versus 41 ms and 6.3 percentage points previously). However, these (between-participants) differences were not significant and so should not be taken as strong evidence for an asymmetry in the opposite direction than predicted.

These findings diverge from previous literature (Kroll et al., 1999; Van Hell & Dijkstra, 2002; Cop et al., 2016) showing that cognate effects tend to be larger in L2 than L1. Van Hell



and Dijkstra (2002), for example, found that the cognate facilitation effect in their lexical decision experiment was larger when participants were tested in their second language than when they were tested in their native language. This finding likely matches most researchers' (and lay bilinguals') intuition, as the representation of a cognate is thought to be less stable in a second language. This means there would be more scope for the L1 representation to aid processing in the L2. If we extend this reasoning to interlingual homographs, we would expect that the L1 representation would be more likely to interfere with processing in the L2 than vice versa. Certainly anecdotally, interlingual homographs tend to be especially tricky in one's second language, and we imagine this is even more true in a semantic relatedness task than a lexical decision task. It is, therefore, unclear to us why we failed to observe a similar effect for interlingual homographs. We suggest that future research using a larger sample should explore in more detail how the characteristics of the participants may influence the size of the cognate facilitation effect and interlingual homograph inhibition effect. For example, in this case as well, language proficiency and differences in recent and longer-term experiences (i.e. language dominance) might be modulating the relative magnitude of these effects in the bilingual's different languages. Furthermore, a similar study that employed a within-participants design instead of the between-participants design that we used would be able to establish with greater certainty whether priming from L1 to L2 or L2 to L1 is stronger.

In summary, this experiment replicated the finding that a bilingual's prior experience with an interlingual homograph can influence its subsequent processing in their other language (Poort et al., 2016; Poort & Rodd, 2019b). For the first time, we have also shown that this interference effect is not restricted to the processing of words in the participants' (potentially more malleable) L2: in the current experiment participants' ability to process these words in their (more stable) L1 was hindered by their prior encounter with the word form in their L2. In addition, we showed that the interlingual homograph inhibition effect is present not only when participants make semantic relatedness judgements in their L2, but also in their L1. These findings add to a growing literature that emphasises the high level of interaction between words in a bilingual's different languages within their mental lexicon (for a review, see e.g. Dijkstra, 2005; Dijkstra & Van Heuven, 2012; Poort & Rodd, 2017a; Dijkstra & Van Heuven, 2018; Poort & Rodd, 2019b).

---

## Data Accessibility Statement

All data and supplementary materials are available on the Open Science Framework via <https://osf.io/b9az4/> (doi: [10.17605/OSF.IO/2SWYG](https://doi.org/10.17605/OSF.IO/2SWYG)). The experiment itself can be previewed and cloned via Gorilla Open Materials ([app.gorilla.sc/openmaterials/245694](https://app.gorilla.sc/openmaterials/245694)).

## Ethics and Consent

The UCL Experimental Psychology Ethics Committee provided approval of the study protocols (Project ID: EP/2017/009).

## Acknowledgements

We would like to thank the Experimental Psychology Society (EPS) for the Small Grant (award number 178661) awarded to Jennifer M. Rodd that funded this research. We would also like to thank the two anonymous reviewers for their valuable comments, which helped us improve our paper.

## Competing Interests

The authors have no competing interests to declare.

## Author contributions

- Eva D. Poort: conceptualisation, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, validation, visualisation, writing – original draft, writing – review & editing.
- Jennifer M. Rodd: conceptualisation, funding acquisition, methodology, project administration, resources, supervision, writing – original draft, writing – review & editing.

---

## References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*, 388–407. DOI: <https://doi.org/10.3758/s13428-019-01237-x>

Armstrong, B. C., & Plaut, D. C. (2008). *Settling dynamics in distributed networks explain task differences in semantic ambiguity effects: Computational and behavioral evidence*. 30th Annual Meeting of the Cognitive Science Society, Washington, DC, USA. <https://escholarship.org/uc/item/6rv2c5hh>. DOI: <https://doi.org/10.1037/e527312012-207>

- Armstrong, B. C., & Plaut, D. C. (2016). Disparate semantic ambiguity effects from semantic processing dynamics rather than qualitative task differences. *Language, Cognition and Neuroscience*, 31(7), 940–966. DOI: <https://doi.org/10.1080/23273798.2016.1171366>
- 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. DOI: <https://doi.org/10.1016/j.jml.2012.11.001>
- Basnight-Brown, D. M., & Altarriba, J. (2007). Differences in semantic and translation priming across languages: The role of language direction and language dominance. *Memory & Cognition*, 35(5), 953–965. DOI: <https://doi.org/10.3758/BF03193468>
- Bates, D., Maechler, M., Bolker, B. M., & Walker, S. (2015). Fitting linear mixed effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>
- Beretta, A., Fiorentino, R., & Poeppel, D. (2005). The effects of homonymy and polysemy on lexical access: An MEG study. *Cognitive Brain Research*, 24(1), 57–65. DOI: <https://doi.org/10.1016/j.cogbrainres.2004.12.006>
- Betts, H. N. (2018). *Retuning lexical-semantic representations on the basis of recent experience* [Ph.D. dissertation, University College London, London, United Kingdom]. <https://discovery.ucl.ac.uk/id/eprint/10049908>
- Betts, H. N., Gilbert, R. A., Cai, Z. G., Okedara, Z. B., & Rodd, J. M. (2018). Retuning of lexical-semantic representations: Repetition and spacing effects in word-meaning priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(7), 1130–1150. DOI: <https://doi.org/10.1037/xlm0000507>
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41(4), 977–990. DOI: <https://doi.org/10.3758/BRM.41.4.977>
- Chen, B., Zhou, H., Gao, Y., & Dunlap, S. (2014). Cross-language translation priming asymmetry with Chinese-English bilinguals: A test of the Sense Model. *Journal of Psycholinguistic Research*, 43(3), 225–240. DOI: <https://doi.org/10.1007/s10936-013-9249-3>
- Cop, U., Dirix, N., Van Assche, E., Drieghe, D., & Duyck, W. (2016). Reading a book in one or two languages? An eye movement study of cognate facilitation in L1 and L2 reading. *Bilingualism: Language and Cognition*, 20(4), 1–23. DOI: <https://doi.org/10.1017/S1366728916000213>
- Davis, C., Sánchez-Casas, R., García-Albea, J. E., Guasch, M., Molera, M., & Ferré, P. (2010). Masked translation priming: Varying language experience and word type with Spanish–English bilinguals. *Bilingualism: Language and Cognition*, 13, 137–155. DOI: <https://doi.org/10.1017/S1366728909990393>
- De Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, 30(1), 90–123. DOI: [https://doi.org/10.1016/0749-596X\(91\)90012-9](https://doi.org/10.1016/0749-596X(91)90012-9)

- Declerck, M., & Philipp, A. M. (2015). A review of control processes and their locus in language switching. *Psychonomic Bulletin & Review*, 22(6), 1630–1645. DOI: <https://doi.org/10.3758/s13423-015-0836-1>
- Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). Oxford University Press.
- Dijkstra, T., & Van Heuven, W. J. B. (2012). Word recognition in the bilingual brain. In M. Faust (Ed.), *The handbook of the neuropsychology of language* (pp. 451–471). Blackwell Publishing. DOI: <https://doi.org/10.1002/9781118432501.ch22>
- Dijkstra, T., & Van Heuven, W. J. B. (2018). Visual word recognition in multilinguals. In S.-A. Rueschemeyer & M. G. Gaskell (Eds.), *The Oxford handbook of psycholinguistics* (2 ed., pp. 118–143). Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780198786825.013.6>
- Dimitropoulou, M., Duñabeitia, J. A., & Carreiras, M. (2011a). Masked translation priming effects with low proficient bilinguals. *Memory & Cognition*, 39(2), 260–275. DOI: <https://doi.org/10.3758/s13421-010-0004-9>
- Dimitropoulou, M., Duñabeitia, J. A., & Carreiras, M. (2011b). Two words, one meaning: Evidence of automatic co-activation of translation equivalents. *Frontiers in Psychology*, 2, 188. DOI: <https://doi.org/10.3389/fpsyg.2011.00188>
- Duñabeitia, J. A., Dimitropoulou, M., Uribe-Etxebarria, O., Laka, I., & Carreiras, M. (2010). Electrophysiological correlates of the masked translation priming effect with highly proficient simultaneous bilinguals. *Brain Research*, 1359, 142–154. DOI: <https://doi.org/10.1016/j.brainres.2010.08.066>
- Duñabeitia, J. A., Perea, M., & Carreiras, M. (2010). Masked translation priming effects with highly proficient simultaneous bilinguals. *Experimental Psychology*, 57(2), 98–107. DOI: <https://doi.org/10.1027/1618-3169/a000013>
- Duyck, W., & Warlop, N. (2009). Translation priming between the native language and a second language: New evidence from Dutch–French bilinguals. *Experimental Psychology*, 56(3), 173–179. DOI: <https://doi.org/10.1027/1618-3169.56.3.173>
- Finkbeiner, M., Forster, K., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language*, 51(1), 1–22. DOI: <https://doi.org/10.1016/j.jml.2004.01.004>
- Francis, W. S., Augustini, B. K., & Sáenz, S. P. (2003). Repetition priming in picture naming and translation depends on shared processes and their difficulty: Evidence From Spanish-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(6), 1283–1297. DOI: <https://doi.org/10.1037/0278-7393.29.6.1283>
- Gaskell, M. G., Cairney, S. A., & Rodd, J. M. (2019). Contextual priming of word meanings is stabilized over sleep. *Cognition*, 182, 109–126. DOI: <https://doi.org/10.1016/j.cognition.2018.09.007>
- Gilbert, R. A., Davis, M. H., Gaskell, M. G., & Rodd, J. M. (2018). Listeners and readers generalize their experience with word meanings across modalities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(10), 1533–1561. DOI: <https://doi.org/10.1037/xlm0000532>

- Gilbert, R. A., Davis, M. H., Gaskell, M. G., & Rodd, J. M. (2021). The relationship between sentence comprehension and lexical-semantic retuning. *Journal of Memory and Language*, 116, 104188. DOI: <https://doi.org/10.1016/j.jml.2020.104188>
- Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning Memory and Cognition*, 23, 1122–1139. DOI: <https://doi.org/10.1037/0278-7393.23.5.1122>
- Grainger, J., & Frenck-Mestre, C. (1998). Masked priming by translation equivalents in proficient bilinguals. *Language and Cognitive Processes*, 13(6), 601–623. DOI: <https://doi.org/10.1080/016909698386393>
- Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493–498. DOI: <https://doi.org/10.1111/2041-210X.12504>
- Hino, Y., Pexman, P. M., & Lupker, S. J. (2006). Ambiguity and relatedness effects in semantic tasks: Are they due to semantic coding? *Journal of Memory and Language*, 55(2), 247–273. DOI: <https://doi.org/10.1016/j.jml.2006.04.001>
- Jiang, N. (1999). Testing processing explanations for the asymmetry in masked cross-language priming. *Bilingualism: Language and Cognition*, 2(1), 59–75. DOI: <https://doi.org/10.1017/S1366728999000152>
- Jiang, N., & Forster, K. I. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory and Language*, 44(1), 32–51. DOI: <https://doi.org/10.1006/jmla.2000.2737>
- Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, 42(3), 643–650. DOI: <https://doi.org/10.3758/BRM.42.3.643>
- Kim, J., & Davis, C. (2003). Task effects in masked cross-script translation and phonological priming. *Journal of Memory and Language*, 49, 484–499. DOI: [https://doi.org/10.1016/S0749-596X\(03\)00093-7](https://doi.org/10.1016/S0749-596X(03)00093-7)
- Klepousniotou, E., & Baum, S. R. (2007). Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics*, 20(1), 1–24. DOI: <https://doi.org/10.1016/j.jneuroling.2006.02.001>
- Kroll, J. F., Dijkstra, A., Janssen, N., & Schriefers, H. J. (1999). Cross-language lexical activity during production: Evidence from cued picture naming. *11th Congress of the European Society for Cognitive Psychology*. Gent, Belgium.
- Lee, Y., Jang, E., & Choi, W. (2018). L2-L1 translation priming effects in a lexical decision task: Evidence from low proficient Korean-English bilinguals [original research]. *Frontiers in Psychology*, 9. DOI: <https://doi.org/10.3389/fpsyg.2018.00267>
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, 44(2), 325–343. DOI: <https://doi.org/10.3758/s13428-011-0146-0>

- Lupker, S. J., Perea, M., & Nakayama, M. (2015). Non-cognate translation priming effects in the same–different task: Evidence for the impact of “higher level” information. *Language, Cognition and Neuroscience*, 30(7), 781–795. DOI: <https://doi.org/10.1080/23273798.2015.1015430>
- Nakayama, M., Sears, C. R., Hino, Y., & Lupker, S. J. (2013). Masked translation priming with Japanese–English bilinguals: Interactions between cognate status, target frequency and L2 proficiency. *Journal of Cognitive Psychology*, 25(8), 949–981. DOI: <https://doi.org/10.1080/20445911.2013.839560>
- Perea, M., Duñabeitia, J. A., & Carreiras, M. (2008). Masked associative/semantic priming effects across languages with highly proficient bilinguals. *Journal of Memory and Language*, 58(4), 916–930. DOI: <https://doi.org/10.1016/j.jml.2008.01.003>
- Poort, E. D., & Rodd, J. M. (2017a). The cognate facilitation effect in bilingual lexical decision is influenced by stimulus list composition. *Acta Psychologica*, 180, 52–63. DOI: <https://doi.org/10.1016/j.actpsy.2017.08.008>
- Poort, E. D., & Rodd, J. M. (2017b). Studies of cross-lingual long-term priming. *PsyArXiv*. DOI: <https://doi.org/10.31234/osf.io/ert8k>
- Poort, E. D., & Rodd, J. M. (2019a). A database of Dutch–English cognates, interlingual homographs and translation equivalents. *Journal of Cognition*, 2(1), 1–15. DOI: <https://doi.org/10.5334/joc.67>
- Poort, E. D., & Rodd, J. M. (2019b). Towards a distributed connectionist account of cognates and interlingual homographs: Evidence from semantic relatedness tasks. *PeerJ*, 7, e672. DOI: <https://doi.org/10.7717/peerj.6725>
- Poort, E. D., Warren, J. E., & Rodd, J. M. (2016). Recent experience with cognates and interlingual homographs in one language affects subsequent processing in another language. *Bilingualism: Language and Cognition*, 19(1), 206–212. DOI: <https://doi.org/10.1017/S1366728915000395>
- R Core Team. (2017). *R: A language and environment for statistical computing*. In (Version 3.4.3) R Foundation for Statistical Computing.
- R Core Team. (2018). *R: A language and environment for statistical computing*. In (Version 3.4.4) R Foundation for Statistical Computing.
- Rodd, J. M. (2018). Lexical ambiguity. In S.-A. Rueschemeyer & M. G. Gaskell (Eds.), *The Oxford Handbook of Psycholinguistics* (2nd ed., pp. 96–117). Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780198786825.013.5>
- Rodd, J. M. (2020). Settling into semantic space: An ambiguity-focused account of word-meaning access. *Perspectives on Psychological Science*, 15(2), 411–427. DOI: <https://doi.org/10.1177/1745691619885860>
- Rodd, J. M., Cai, Z. G., Betts, H. N., Hanby, B., Hutchinson, C., & Adler, A. (2016). The impact of recent and long-term experience on access to word meanings: Evidence from large-scale internet-based experiments. *Journal of Memory and Language*, 87, 16–37. DOI: <https://doi.org/10.1016/j.jml.2015.10.006>
- Rodd, J. M., Cutrin, B. L., Kirsch, H., Millar, A., & Davis, M. H. (2013). Long-term priming of the meanings of ambiguous words. *Journal of Memory and Language*, 68(2), 180–198. DOI: <https://doi.org/10.1016/j.jml.2012.08.002>



- Rodd, J. M., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, 46(2), 245–266. DOI: <https://doi.org/10.1006/jmla.2001.2810>
- Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2004). Modelling the effects of semantic ambiguity in word recognition. *Cognitive Science*, 28(1), 89–104. DOI: <https://doi.org/10.1016/j.cogsci.2003.08.002>
- Schoonbaert, S., Duyck, W., Brysbaert, M., & Hartsuiker, R. J. (2009). Semantic and translation priming from a first language to a second and back: Making sense of the findings. *Memory & Cognition*, 37(5), 569–586. DOI: <https://doi.org/10.3758/MC.37.5.569>
- Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, 9(4), 780–789. DOI: <https://doi.org/10.3758/BF03196335>
- Voga, M., & Grainger, J. (2007). Cognate status and cross-script translation priming. *Memory & Cognition*, 35, 938–952. DOI: <https://doi.org/10.3758/BF03193467>
- Wang, X. (2013). Language dominance in translation priming: Evidence from balanced and unbalanced Chinese–English bilinguals. *The Quarterly Journal of Experimental Psychology*, 66(4), 727–743. DOI: <https://doi.org/10.1080/17470218.2012.716072>
- Wang, X., & Forster, K. (2014). Is translation priming asymmetry due to partial awareness of the prime? *Bilingualism: Language and Cognition*, 18(4), 657–669. DOI: <https://doi.org/10.1017/S1366728914000650>
- Wen, Y., & Van Heuven, W. J. B. (2017). Non-cognate translation priming in masked priming lexical decision experiments: A meta-analysis. *Psychonomic Bulletin & Review*, 24(3), 879–886. DOI: <https://doi.org/10.3758/s13423-016-1151-1>
- Williams, J. N. (1994). The relationship between word meanings in the first and second language: Evidence for a common, but restricted, semantic code. *European Journal of Cognitive Psychology*, 6(2), 195–220. DOI: <https://doi.org/10.1080/09541449408520143>
- Witzel, N. O., & Forster, K. I. (2012). How L2 words are stored: The episodic L2 hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(6), 1608–1621. DOI: <https://doi.org/10.1037/a0028072>
- Yarkoni, T., Balota, D., & Yap, M. J. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, 15(5), 971–979. DOI: <https://doi.org/10.3758/PBR.15.5.971>

