Short title: Cognate effect in spoken and written word production

Comparing the cognate effect in spoken and written second language word production

Merel Muylle[1], Eva Van Assche[2] & Robert J. Hartsuiker[1]

[1]Department of Experimental Psychology, Ghent University, Ghent, Belgium
[2] Thomas More University of Applied Sciences, Antwerp, Belgium

Address for correspondence:

Merel Muylle
Department of Experimental Psychology
Ghent University
Henri Dunantlaan 2
B-9000 Gent (Belgium)
E-mail: merel_muylle@hotmail.com

Abstract

Cognates – words that share form and meaning between languages – are processed faster than control words. However, is this facilitation effect merely lexical in nature or does it cascade to phonological/orthographic (i.e., sub-lexical) processes? This study compared cognate effects in spoken and typewritten production, which share lexical, but not sub-lexical processes. Dutch-English bilinguals produced English names for pictures representing Dutch-English cognates and control words in either the spoken or typewritten modality. Onset latencies were shorter and accuracy was higher for cognates vs. control words and this effect was similar in both modalities. Compared to controls, total latencies in the written modality were similar for cognates with much cross-linguistic overlap, but longer for ones with less overlap. Additionally, error analysis showed that cognates were more affected by L1 interference than controls. These results suggest two different cognate effects: one at the lexical and one at the sub-lexical level.

Keywords: bilingualism, word production, cognate effect, writing

Converging evidence suggests that bilinguals activate both their mother tongue (L1) and their second language (L2) simultaneously when processing linguistic information (e.g., Dijkstra & van Heuven, 2002; Van Hell & Dijkstra, 2002). Also in a unilingual situation, when only one of both languages is relevant, the other language is active to some extent, and even influences processing in the relevant language. There is already much information available about bidirectional influences of L1 and L2 in the bilingual's reading, spoken word recognition and speech production (e.g., Duyck & De Houwer, 2008; Green, 1998; Kroll & Stewart, 1994; Lagrou, Hartsuiker & Duyck, 2011; van Heuven, Dijkstra & Grainger, 1998). In contrast, little is known about the cognitive processes underlying writing, especially in bilinguals. To illustrate this, according to the table of contents of *The Oxford Handbook of Language Production* (Goldrick, Ferreira & Miozzo, 2014), 19 chapters are dedicated to speech production and only 3 to writing. In their review on architectures, representations and processes of language production, Alario, Costa, Ferreira and Pickering (2006) also pointed to the need for more research in the written domain. They argued that models of language production should entail both spoken and written production.

Despite the many similarities between spoken and written production, there are also some important differences between both modalities. For instance, writing requires phoneme-to-grapheme conversion on top of the lexical and phonological selection that takes place in speech production. Moreover, writing can be considered to be less automatic than speaking, because it is a relatively recent skill in humans. Most models of spoken production assume that there are three stages: a) selection of concepts, b) selection of lexical forms and c) selection of phonological forms (e.g., Caramazza, 1997; Dell, 1986; Levelt, Roelofs & Meyer, 1999). It is often assumed that the central (i.e., conceptual and lexical) processes are shared between spoken and written modalities, whereas the peripheral, post-lexical processes, such as phonological and orthographic selection, and motor processes are different (e.g., Bonin, Chalard, Méot & Fayol, 2002; Bonin, Méot, Lagarrigue & Roux, 2015; Hillis, Rapp, Romani & Caramazza, 1990; Perret & Laganaro, 2012, 2013). Evidence for shared central processes stems from studies that found similar results in spoken and written production for the semantic interference effect (Bonin & Fayol, 2000) and word frequency effect (Bonin & Fayol, 2002), which are considered to have a central locus (Dell, 1986; Caramazza, 1997, but see Almeida, Knobel, Finkbeiner & Caramazza, 2007; Jescheniak & Levelt, 1994, for an alternative view). In addition, a neuro-imaging study comparing spoken and written picture naming found that lexical processes represented in the left fusiform gyrus (BA 37), were modality-independent (DeLeon et al., 2007).

The complicated relation between speaking and writing suggests that caution is needed when assuming that certain phenomena that occur in speech, such as parallel activation of both L1 and L2, also generalize to writing. One example of a study that addressed the role of L1 in L2 writing is Meade, Midgley, Dijkstra and Holcomb (2018), who investigated the role of orthographic similarity with L1 in the acquisition and typewritten production of new pseudo-words in a non-existing language. This study found that pseudo-words that were very similar to L1 words were typed with higher accuracy than less similar pseudo-words. However, it is not clear whether these similarity effects are also present in highly proficient L2 speakers. In general, very little is known about the interplay between L1 and L2 during writing. The current study aims to gain more insight in whether and how bilinguals activate their L1 during L2 written production and how this relates to spoken production.

Several studies use cognates as a tool to investigate the interaction between languages in the bilingual mind. Cognates are words that share a similar phonological form and meaning between languages, for instance 'shell' in English and 'schelp' in Dutch. Van Hell and Dijkstra (2002) found that bilinguals had faster reaction times for cognates in comparison to control words during a lexical decision task. They argued that bilinguals are faster to recognize cognates, because these are represented in the lexicons of both languages. For instance, when processing the word ‘shell’, English-Dutch bilinguals receive activation from both the English word form ‘shell’ and the Dutch word form ‘schelp’, which enhances recognition. In contrast, English monolinguals only receive activation from the English word form and hence do not process cognates faster than non-cognates. Van Hell and Dijkstra (2002) called this bilingual advantage the *cognate facilitation effect*. A similar advantage was also found in other tasks, such as translation tasks (e.g., de Groot & Nas, 1991), auditory lexical decision tasks (Woutersen, de Bot & Weltens, 1995), picture naming tasks (e.g., Costa, Caramazza, & Sebastian-Galles, 2000; Hoshino & Kroll, 2008; see below for a more in depth discussion of this topic) and sentence and text reading comprehension tasks (e.g., Duyck, Van Assche, Drieghe & Hartsuiker, 2007; Van Assche, Drieghe, Duyck, Welvaert & Hartsuiker, 2011; Van Assche, Duyck, Hartsuiker & Diependaele, 2009). Moreover, these reading comprehension studies showed that even a highly constraining semantic context, denoted by the surrounding sentences, cannot extinguish cross-lingual interactions (i.e., cognates were read faster than non-cognates in both low- and high-constraining sentences), although a recent meta-analysis by Lauro and Schwartz (2017) has shown that such interactions are smaller for high- vs. low-constraining sentences. The language of the observed sentence can be seen as the context to select the target language, but this seems to be insufficient to deactivate the other language entirely. Nevertheless, the amount of cross-lingual activation may depend on whether L1 or L2 is the target language (Palma et al., 2019) and on executive control skills in the bilingual speaker (Pivneva et al., 2014). Cognate effects are also present in L1 in a unilingual context (Van Hell & Dijkstra, 2002) and can even be found in trilinguals, where the effects add up over the three languages (Lemhöfer, Dijkstra & Michel, 2004).

Most of the tasks discussed above focus on recognition, whereas the focus of the current study will be on production. To investigate the cognate facilitation effect in spoken word production, the most frequently used paradigm is the picture naming task, in which participants say aloud the name of the picture on the screen. In a Spanish experiment with this task, Costa and colleagues (2000) found shorter naming latencies for cognates in comparison to control words in Catalan-Spanish bilinguals, but not in Spanish monolinguals, showing evidence for activation of non-selected lexical items. Moreover, the difference between naming latencies of cognates vs. non-cognates was larger in the non-dominant language. In addition, proficiency seems to modulate the magnitude of the cognate effect (Christoffels, de Groot & Kroll, 2006; Van Hell & Dijkstra, 2002). The cognate facilitation effect during picture naming is also reflected in ERP components (Christoffels, Firk & Schiller, 2007; Strijkers, Costa & Thierry, 2010). In sum, converging evidence supports the hypothesis that bilinguals activate lexical representations of both languages when producing in one of them.

The cognate facilitation effect in spoken picture naming has been taken as evidence for theories that propose language-nonselective access of lexical information during production (e.g., Costa et al., 2000; Starreveld et al., 2014). A different view was proposed by Costa, Pannunzi, Deco and Pickering (2017), who argued that the cognate facilitation effect could also be a result of the learning context in which bilinguals acquire new L2 vocabulary; because there are many resemblances between the translation equivalents in both languages, cognates might be picked up faster and earlier during learning than non-cognates, resulting in a strong memory trace for – and readily accessible representations of – cognates. Their claims were supported by a computational model that was able to generate phenomena that were typically considered as evidence for parallel activation, after turning off L1 representations of the model. More recently, the findings of this model were challenged by Oppenheim, Wu and

Thierry (2018), who pointed out a number of issues with the model code. They showed that by implementing some adaptations to the model (which they argued to be necessary), it could no longer generate the phenomena without maintaining the connections with L1 representations. In the current study, we will, amongst others, investigate whether there are L1 intrusions (i.e., errors) during the typewritten and spoken production of cognates vs. control words, given that Costa et al.'s (2017) learning theory would not predict such intrusions (as L1 is not activated during L2 production according to this theory). A language-nonselective account on the other hand, is compatible with such intrusions, although the activation of the non-target language may be task dependent (see Martin & Nozari, 2020). As such, the absence of L1 intrusions does not necessarily provide evidence against the non-selective account (we return to this issue in the discussion). Before we present more precise predictions about cognate effects in spoken and written picture naming, we first need to be more explicit about the locus of cognate effects in word production and about the role of phonological and orthographical information in written and spoken production. We turn to these issues in the sections below.

The locus of the cognate effect in word production

Despite the large number of studies investigating the cognate facilitation effect, researchers have not yet reached consensus about where this effect originates in the production process. In 2000, Costa and colleagues interpreted cognate facilitation as evidence for a cascading model of word production: when a bilingual names a picture, lexical representations of both languages will become active and will each activate representations at the phoneme level. In the case of a cognate, this process will be facilitated because of the overlap in phonemes. According to this view, the cognate effect is situated in the mapping between central and peripheral levels, as the activation from both languages cascades from the lexical to the phonological level. Moreover, an ERP study by Christoffels et al. (2007) showed a difference between cognates and non-cognates around 300 ms, indicating that the cognate facilitation effect indeed has a phonological origin given estimates of the time course of language production processes (Indefrey & Levelt, 2004).

In contrast, Strijkers et al. (2010) found an earlier time-window (i.e., 150-200 ms) for this effect and argued that the mechanism responsible seems to be lexical access rather than phonological retrieval. Because of the tight link between L1 and L2 phonology in cognates, translation equivalents become more often co-activated than in the case of non-cognates (because the lexical item activates its phonology, which in turn activates words with similar phonology, including the translation equivalent), resulting in a strong connection between both lexical items that is no longer mediated by phonological access. In 2017, Costa and colleagues proposed that such a connection is formed during early stages of L2 acquisition, because cognates are picked up faster than non-cognates (see also de Groot & Keijzer, 2000; Lotto & de Groot, 1998). As such, the cognate effect may have become strictly lexical in nature for more proficient bilinguals. More recently, Muscalu and Smiley (2019) found in a written translation task that the cognate facilitation effect is limited to onset latencies, whereas full word typing duration is subject to cognate *interference* effects. They concluded that cognate facilitation is a lexical process, whereas cognate interference is a sub-lexical process.

In conclusion, it remains unclear whether the cognate facilitation effect is a purely lexical or – at least partly – phonological phenomenon. A possible way to gain more insight into this issue might be to compare cognate effects in spoken and written production. In the next section, we therefore discuss the representations involved in both of these modalities.

Spoken vs. written production

Although there is a clear distinction between spoken and written production at the level of motor processes (e.g., mouth and tongue muscles in speech vs. hand and finger movements in writing) and in their building blocks (phonemes vs. graphemes), the role of phonology and orthography in both modalities is less intuitive. On the one hand, it seems that orthography does not necessarily play a role in speech, given that people who cannot read and write are nevertheless able to speak. In addition, Ferrand, Grainger and Segui (1994) showed in a masked priming experiment that pre-activation of phonological, but not orthographic information facilitated spoken picture naming. Note that Damian and Bowers (2003) did find an effect of orthography in priming of spoken production, whereas other studies failed to replicate this finding (e.g., Alario, Perre, Castel & Ziegler, 2007; Bi, Wei, Janssen & Han, 2009; Damian & Bowers, 2009; Roelofs, 2006; Zhang & Damian, 2012). Furthermore, Hoshino and Kroll (2008) found a similar cognate facilitation effect in a spoken picture naming task between languages with the same (i.e., English-Spanish) vs. a different script (i.e., English-Japanese), which indicates that written word forms are not (strongly) activated in speech. In sum, it seems that orthographic activation during speech is either very weak or non-existent.

On the other hand, there has been much debate about whether phonology plays a role in written language. Do people also activate phonological representations during reading and writing? In other words, does writing merely entail an additional stage (i.e., phoneme-grapheme conversion) on top of speech production processes, or is the access of graphemic information independent of phonological information? According to the *orthographic autonomy hypothesis* (Miceli, Benvegnu, Capasso & Caramazza, 1997; Rapp, Benzing & Caramazza, 1997), orthographic information can be accessed without mediation of the phonological system, whereas the *obligatory phonological mediation hypothesis* (Afonso & Álvarez, 2011; Bonin, Peereman & Fayol, 2001; Geschwind, 1974; Luria, 1970; Qu, Damian & Li, 2016) states that the phonological system is necessarily involved when processing orthographic information. Studies on this matter indicate that it is *possible* to by-pass phonology during writing, as shown in brain-damaged patients, who were able to write down picture names without being able to name them orally (e.g., Rapp & Caramazza, 1997; Shelton & Weinrich, 1997) and in languages with low phonology-orthography overlap, such as Mandarin Chinese (Zhang & Wang, 2015, 2016). However, phonology seems to be activated for most of the time in healthy individuals during written production in alphabetic languages (Bonin et al., 2001). The specific contribution of phonology may however depend on the type of task: in a copy task phonology may be less involved compared to a spelling-to-dictation task. For written picture naming, both an indirect (i.e., phonologically mediated) and a direct route (i.e., direct link between lexicon and orthographic representations) may be involved (Bonin et al., 2015, 2001; Damian, Dorjee & Stadthagen-Gonzalez, 2011; Damian & Qu, 2013; Qu, Damian, Zhang & Zhu, 2011).

In sum, the abovementioned studies indicate that written picture naming benefits from both phonological and orthographic activation, whereas spoken picture naming depends primarily on phonology. When taking this to a bilingual situation, the activation of words that have overlap in orthography and phonology between languages, such as cognates, cascades into both orthography and phonology in written production, but mainly into phonology in spoken production. Thus, if the locus of the cognate facilitation effect would be in the cascading between lexicon and phonological/orthographic form, a larger cognate facilitation effect should be observed in written as opposed to spoken production. In contrast, if the cognate facilitation effect is purely lexical, no differences should be observed between both modalities. To tests these predictions, we carried out the study below.

The present study

The current study compared cognate effects in written and spoken word production by means of a picture naming paradigm. Dutch-English bilinguals produced the names of pictures in English by either naming them orally or typing the names on a computer keyboard as fast and accurately as possible. For spoken naming, onset reaction times (RTs) were operationalized as voice onset time, and for writing, this was first keystroke latency. In addition, we measured the total duration to produce the word (i.e., from onset to pressing <enter>). The critical pictures represented a cognate word (i.e., words that have very large form and meaning overlap between L1 and L2), like "bed" [bɛd] (Dutch: "bed" [bɛt]), while control pictures represented non-cognate words (that had no form overlap across languages), for instance "dog" [dɒg] (Dutch: "hond" [hɔnt]). Because the cognate facilitation effect has been observed in spoken production, it can be hypothesized that this is also the case for written production. Moreover, if the cognate facilitation effect is purely lexically driven, participants should have faster onset RTs and higher accuracy for cognates compared to control words to a similar extent in spoken and written picture naming, because central processes are shared between both modalities. In that case, there may also be a cognate interference effect during the full word production (i.e., longer total duration for cognates compared to control words, cf. Muscalu & Smiley, 2019) and this effect may differ across modalities, because this should reflect peripheral processes. In contrast, if the cognate facilitation effect is (partly) peripherally located, larger facilitation should be observed in written production, given that the lexical representation cascades to both phonological and orthographic representations during writing, resulting in stronger activation of the translation equivalent compared to spoken production, that relies mainly on phonology. Furthermore, if cognate effects result from the activation of L1 during L2 processing, influences of L1 phonology, orthography, or even the lexicon should be observed into L2 production. In other words, L1 intrusions into L2 production would indicate that the L1 and L2 are co-activated, in contrast to Costa et al.'s (2017) learning hypothesis. In order to investigate whether there are L1 influences in the productions (e.g., writing 'shelp' instead of 'shell' [Dutch: 'schelp']), we conducted an error analysis on the responses.

Method

Participants

In this study 80 students (21 males and 59 females; age: *M* = 21.5, *SD* = 4.79) received either credits or payment for participation. Half of them were assigned to the spoken condition and the other half to the written condition. All participants were native Dutch speakers and had learned English as a second language before the age of 14. They had normal or corrected-to-normal vision. In order to ensure sufficient vocabulary knowledge to name most of the pictures, only highly proficient participants were selected based on their scores for the English (L2) LexTALE test (Lemhöfer & Broersma, 2012). Students that were interested to participate were sent an e-mail in which they were asked to complete the LexTALE tests online at www.lextale.com. The cut-off score was set at 70%. In addition, the participants completed the Dutch version of the LexTALE test (L1).

The average LexTALE scores were 89.05% (SD = 6.66%) for Dutch and 82.42% (SD 7.60%) for English. All scores ranged between 67.5% and 100% for the Dutch version, and between 70% and 100% for the English version. Furthermore, the majority of the participants had better scores for L1 than L2, showing the expected bilingual pattern (62 out of 80). This was also reflected in the self-ratings of proficiency (on a 1-7 Likert scale; Dutch: *M* = 6.15, *SD* = 0.60; English: *M* = 5.42, *SD* = 0.69). The self-ratings for French show that French was the L3 (*M* = 3.83, *SD* = 0.91) in terms of dominance for most of the participants (none of them rated French proficiency higher than English proficiency). Twenty-five of the participants in the written condition indicated that they had touch-typing skills. The others used on average 5.7 (*SD* = 2.1) fingers to type. The mean self-rated typing proficiency (on a 1-7 Likert scale, with 1 = 'not proficient at all' and 7 = 'very proficient') was 5.85 (*SD* = 2.91).

In order to check for any differences between participants in the spoken and written conditions, the groups were compared on LexTALE scores, self-rating measures and several further individual characteristics (Table 1). One-way ANOVAs showed no significant differences between the groups for any of the characteristics, except for the self-rated reading proficiency in English ($F(1, 78) = 4.76$, $p < .05$) and Dutch ($F(1, 78) = 4.95$, $p < .05$), and the self-rated speaking proficiency in Dutch ($F(1, 78) = 4.66$, $p < .05$), which were somewhat higher in the written group.

<Insert Table 1 about here>

Stimuli

Because most Flemish students also speak French, and earlier work has demonstrated effects of cognates with L3 (Van Hell & Dijkstra, 2002), we selected only Dutch-English cognates with a different form in French (e.g., 'computer'(NL)-'computer'(E)-'ordinateur'(F)). By doing so, we avoided the possibility that French word forms would influence the processing of the cognates. We selected 48 cognates. For each of them, we selected a control, non-cognate word that was matched in length, frequency, bigram frequency, rank of first letter (i.e., ranking based on the frequency of a letter as first letter in a given language) and neighbourhood for both L1 and L2, by means of WordGen (Duyck, Desmet, Verbeke & Brysbaert, 2004). Finally, 42 filler items were added. In total, 138 black and white drawings, which corresponded to the target and filler words, were taken from Severens, Lommel, Ratinckx and Hartsuiker (2005). Ten items served as practice stimuli. See Appendix A for the list of words and their properties.

Procedure

The programming and data collection of the experiment were done in E-Prime (spoken) and PsychoPy (written; Peirce et al., 2019).[1] First, participants signed an informed consent form after which they were seated in front of a computer screen on which the instructions of the experiment were presented. Trials started with a fixation cross presented during 500 ms, followed by a picture in the centre of the screen. Half of the participants were instructed to type below the name of the picture in English as soon as possible using an AZERTY keyboard[2]. If the target word was unknown, they could skip the trial and go on with the next. There were no time limits on responding. Correction was allowed within the trial by pressing <backspace> after which a {*} appeared and then the participant could type the whole word again. All typing was recorded, including mistakes that the participant corrected. When the word was written, the participant could press <enter> to view the next picture. The other half of the participants were instructed to name the pictures orally. They were also instructed to press <enter> to proceed to the next picture.

There was no familiarization phase of the words and pictures, because we were also interested in errors, which might expose L1 influences. The first 10 trials were practice trials, which were not included in analysis. If something remained unclear, the participant could ask questions before going on with the experiment. The experiment itself consisted of 128 trials and the order of picture presentation was randomized. Each picture was presented only once to the same participant. After every 32 trials, there was a break and after the last trial the student was asked to fill in a questionnaire, containing self-report measures of proficiency in Dutch, English and French. The students were asked to estimate their reading, listening, speaking and writing skills in these three languages on a Likert scale (1-7). Furthermore, the questionnaire surveyed the age of acquisition of and frequency of exposure to English and French. Finally, participants in the written condition estimated their typewriting skills on a Likert scale.

Coding of responses

Responses on the picture naming task were scored as correct when the exact target word was produced. All other responses, including the use of backspaces and (near) synonyms were scored as errors. Errors were divided into eight categories: a) typing errors (only for the typing condition), either as a result of pressing an adjacent button on the keyboard, e.g., “beatr” instead of “bear”, or errors resulting in transpositions of letters, e.g., “bera” instead of “bear”, which we assumed to be motor errors, b) disfluencies (only for the spoken condition)[3], e.g., “che… cherry”, c) L1 translation-related errors , e.g., “flyer” instead of “kite” (Dutch: “vlieger”, a noun derived from the verb “vliegen” [to fly]) or “racket” instead of “rocket” (Dutch: “raket”), d) orthographic errors, errors that reflect a wrong orthographic representation of the word form (only for the typing condition), e.g., “swann” instead of “swan”, e) phonological errors (only for the spoken condition), e.g., “hence” instead of “fence”, f) semantic errors, e.g., “glass” instead of “window”, g) synonyms[4], e.g., “ship” instead of “boat”, and h) no response, which means that participants continued to the next trial without writing/saying a word. Items in this final category were discarded from analysis. Responses could also be assigned to different error categories at the same time, e.g., when someone said “racket” instead of “rocket”, this was coded as a phonological error, but also as an L1 translation-related error (Dutch: “raket”).

Results

All data and analysis scripts are available on the Open Science Framework (link: https://osf.io/pkvnt/?view_only=3f9447fb9edc4e119f9873ca86962196).

Cognate effect

The number of valid responses (i.e., everything except no response) was 93.77% for the spoken condition and 96.19% for the written condition; the accuracy of the valid responses was 85.67% in the spoken and 77.17% in the written condition. We first excluded outlier reaction times (over 3*SD above the mean, i.e., onset RT: > 2894 ms for the spoken group and > 6437 ms for the written group; total duration: > 4374 ms for the spoken group and > 6409 ms for the written group). After this, one outlier word was excluded from analysis, because the accuracy was too low (i.e., ‘picture’: 8.8% correct). The cognate effects for onset RTs, total duration and accuracy across modalities can be found in Figure 1.

<Insert Figure 1 around here>

Onset reaction times. The mean onset RT for correct responses was 1165 ms (*SD* = 547) for cognates and 1305 ms (*SD* = 545) for controls in the spoken group and 1427 ms (*SD* = 730) for cognates and 1605 ms (*SD* = 731) for controls in the written group. There was a significant positive correlation between mean spoken and written reaction times aggregated per item (Pearson's $r = .85$, df= 93, p < .001).

The cognate effect and its interaction with spoken and written condition was tested for onset RTs using linear mixed effects models by means of the lme4 package (Bates, Mächler, Bolker & Walker, 2015) in R (R Core Team, 2016). For the random part of the model, the maximal random effects structure (Barr, Levy, Scheepers & Tily, 2013) was included and in case of singularity or other non-convergence, it was reduced until convergence by first removing correlations between random slopes and intercepts and next removing random slopes with coefficients very close to 0 (see Matuschek, Kliegl, Vasishth, Baayen & Bates, 2017). This resulted in a random intercept for *word* and *subject* and a random slope for *modality* over words. The fixed part consisted of the *modality* (spoken vs. written) * *cognate status* (cognate vs. control) interaction. There was a significant main effect of *modality* (*M* spoken = 1231 ms, *M* written = 1510 ms, $\chi^2(1) = 33.89$, $p < .001$): spoken responses were faster than written responses. In addition, there was a significant main effect of *cognate status* (*M* cognate = 1292 ms, *M* control = 1451 ms, $\chi^2(1) = 6.99$, $p < .01$), in other words, there was a cognate facilitation effect. Post-hoc pairwise comparisons using the phia package (De Rosario-Martinez, 2013) showed that this effect was significant in both the written ($\chi^2(1) = 6.72$, $p < .05$) and spoken modality ($\chi^2(1) = 6.33$, $p < .05$). However, there was no interaction between *modality* and *cognate status* ($\chi^2(1) = 0.74$, $p = .39$), which indicates that the cognate facilitation effect is very similar in spoken and written production in terms of onset latencies. The absence of the interaction was further verified using Bayesian hypothesis testing. Concretely, we compared the full model (H1, without the random slope of *modality*[5]) with a model without the *modality* * *cognate status* interaction (H0) by means of the brms package (Bürkner, 2017). The H_{01} Bayes Factor was 32.3 (average after 10 iterations, with values ranging between 24.5 and 53.3), thus showing strong evidence for the null hypothesis of similar cognate facilitation in the spoken vs. written modality.

Total duration. In this section, we report analyses with regard to the total duration of producing the word (i.e., from onset to pressing <enter>). On average, the total duration was 803 ms (*SD* = 443) for cognates and 819 ms (*SD* = 468) for controls in the spoken group and 1108 ms (*SD* = 549) for cognates and 1066 ms (*SD* = 501) for controls in the written group.

In order to test whether there was a cognate interference effect in the production of the entire word (as predicted by Muscalu & Smiley, 2019), we built linear mixed effects models with *total duration* as outcome variable and *modality * cognate status* as fixed effects. The random effects structure was determined in the same way as described for the onset RT model and consisted of a random intercept for *word* and *subject*, and a random slope for *cognate status* over words. There was again a main effect of modality ($\chi^2(1) = 17.34, p < .001$): it took longer to type words ($M = 1088$ ms) than to say them ($M = 811$ ms). However, there was no effect of *cognate status* ($\chi^2(1) = 0.30, p = .58$) and no interaction between *modality* and *cognate status* ($\chi^2(1) =.94, p = .33$). Here, the H_{01} Bayes factor comparing the model with (H_1) and without the modality*cognate status interaction (H_0) was 1.2 (average after 10 iterations, with values ranging between 0.8 and 1.5). Hence, there is only anecdotal evidence in favour of H_0. Taken together, there is no clear evidence for the cognate interference effect, and this is the case for both the spoken and written group.

Accuracy. The mean accuracy was 0.86 (*SD* = 0.35) for cognates and 0.79 (*SD* = 0.41) for controls in the spoken condition; in the written condition it was 0.79 (*SD* = 0.40) for cognates and 0.73 (*SD* = 0.45) for controls. There was a significant positive correlation between accuracy in the spoken and in the written condition for each item (Pearson's $r = .83$, $df = 93$, $p < .001$).

A generalized linear mixed effects model with the logit link-function was fitted for *accuracy* (because it is a binomial outcome variable), using the same fixed effects structure and the same method to determine the random effects structure as in the RT models. Also here, the random effects consisted of a random intercept for *word* and *subject*, and a random slope for *modality* over *words*. The output of this model revealed a significant *modality* * *cognate status* interaction ($\chi^2(1) = 4.04$, $p < .01$), indicating a difference in the cognate effect between both modalities. Pairwise contrasts showed that the cognate effect was significant in both the spoken ($\chi^2(1) = 8.13$, $p < .01$) and written condition ($\chi^2(1) = 6.60$, $p < .05$), although it was slightly smaller in the written condition. In addition, accuracy was lower in general in the written vs. spoken condition ($\chi^2(1) = 21.57$, $p < .001$).

Cross-lingual similarity. In order to explore whether phonological and orthographic similarity across languages depend upon the degree of phonological and orthographic similarity, we computed for each English-Dutch cognate pair the Levenshtein distance (LD) of the orthographic form and the phonological distance (see Downey, Hallmark, Cox, Norquest & Lansing, 2008) of the phonological form (i.e., IPA codes, using the alineR package; Downey, Sun & Norquest, 2017). Linear mixed effects models with *phonological distance* as fixed effect and a random intercept for *wor*d and *subject* showed that there was no effect of phonological distance on onset RT ($t(47.13) = -0.64$, $p = .52$) or total duration ($t(43.53) = -0.76$, $p = .45$) when these were tested as outcome variable in the spoken group. Similar analyses for the written group with *LD* as fixed effect, showed no effect on the onset RT ($t(47.98) = -0.37$, $p = .72$), but a significant effect on the total duration ($t(47.13) = 3.88$, $p < .001$) in the sense that typing duration was shorter for cognates with more orthographic overlap than those with less overlap. Figure 2 plots the relationship between LD and the difference in total typing duration for cognates vs. their matched controls. In order to assess whether there was cognate interference or facilitation depending on LD, we compared for each LD value (0 to 4) whether there was a difference in duration for cognates belonging to that level and their matched controls. For this analysis, we used an ANOVA with *duration* as outcome variable and the interaction between *cognate status* (cognate vs. control) and *LD* (ordered factor with 5 levels, each representing one of the 5 observed LD values, i.e., 0, 1, 2 , 3 and 4) as predictor. Because this interaction was significant ($F(4, 2752) = 3.15$, $p = .01$), we used post-hoc pairwise contrasts with Holm correction to find out for which LD values there was a significant difference between cognates and controls (i.e., if cognates < controls, there was facilitation, but if cognates > controls, there was interference). There was significant cognate interference for LD 4 ($F(1, 2752) = 7.76$, $p < .05$), and marginally significant cognate interference for LD 3 ($F(1, 2752) = 5.59$, $p = .07$), but no difference between cognates and control words with smaller LDs (LD 2: $F(1, 2752) = 2.66$, $p = .31$; LD 1: $F(1, 2752) = 0.03$, $p = .85$; LD 0: $F(1, 2752) = 0.88$, $p = .69$). These analyses indicate that participants experienced interference for cognates with less orthographic overlap between languages, but no interference (nor facilitation) for cognates with more overlap (compared to matched controls).

<Insert Figure 2 around here>

Typewriting skills. Additional (generalized) linear mixed effects models with a random slope for *word* and one for *subject* (determined using the procedure as described above) tested whether there was a difference in the written cognate effect for participants with ($N = 25$) and without touch typing skill ($N = 15$) (i.e., *cognate status * typing skill*). This was not the case for the onset RT ($\chi^2(1) = 0.85$, $p = .36$) and accuracy ($\chi^2(1) = 0.56$, $p = .45$), but there was an interaction for the total duration ($\chi^2(1) = 6.04$, $p < .05$), in the sense that participants with touch typing skill showed a small, but non-significant cognate interference effect (*M* cognates = 1060 ms; *M* controls = 980 ms; $\chi^2(1) = 1.65$, $p = .40$), whereas participants without such skill did not show any difference between cognates and control words (*M* cognates = 1187 ms; *M* controls = 1199 ms; $\chi^2(1) = 0.01$, $p = .93$). Descriptively, participants with touch typing skill had somewhat faster onset RTs in general ($\chi^2(1) = 3.97$, $p < .105$), but were not more accurate ($\chi^2(1) = 1.25$, $p = .26$) than participants without this skill.

Error Analysis

In a final analysis, we conducted a detailed analysis of the errors. There were 1636 erroneous responses on a total of 7680 trials. Of these responses, 95 belonged to more than one error category (i.e., mixed errors). The proportion of each error type for each condition can be found in Table 2. In addition, some errors (14 in total) even bore witness of L3 (French) influences. For instance, one of the participants wrote "gant" instead of "glove", which is "gant" in French.

Similar to previous analyses, we started with generalized linear mixed effects models containing the *cognate status* * *modality* * *error type* interaction on errors (binomial) with a random intercept for *word* and *subject*, and a random slope for *modality* over words and *cognate status* over subjects. However, the output of this model was singular because all random effects were estimated as 0. Therefore, we built a reduced generalized linear model with only the fixed effects. The output of this model showed a significant *cognate status* * *error type* ($\chi^2(4) = 13.42$, $p < .01$) and a *modality* * *error type* interaction ($\chi^2(4) = 112.36$, $p < .001$), but no three-way interaction ($\chi^2(4) = 7.52$, $p = .11$). Pairwise contrasts revealed that there were significantly more L1 translation-related errors ($\chi^2(1) = 8.06$, $p < .05$) and phonological/orthographic errors ($\chi^2(1) = 7.60$, $p < .05$), but fewer synonyms in cognates compared to controls ($\chi^2(1) = 4.96$, $p = .077$), although the latter is only marginally significant. In addition, the spoken group produced more synonyms ($\chi^2(1) = 34.49$, $p < .001$) and semantic errors ($\chi^2(1) = 43.61$, $p < .001$) than the written group, but fewer sub-lexical errors (phonological vs. orthographic errors: $\chi^2(1) = 27.27$, $p < .001$). However, there was no difference between both modalities in terms of L1 translation-related errors ($\chi^2(1) = 1.16$, $p = .28$).

<Insert Table 2 around here>

Discussion

The aim of the study was to compare the cognate effects in spoken and (type)written word production by means of an L2 picture naming task. Cognate effects were investigated in terms of onset latencies, total duration and accuracy of responses. In addition, we performed an error analysis to further investigate L1 influences during L2 production. Overall, there was a clear cognate facilitation effect at the onset of production for both modalities: the onset response latencies were shorter in cognates compared to control words. The difference between cognates and controls was about 140 ms for spoken and 180 ms for written production. Moreover, statistical analyses failed to find a difference in the cognate effect between both modalities and the absence of such a difference was also confirmed by Bayesian analyses. In addition, we found no differences in total duration between the cognates and control words for either modality, but there was a relatively strong positive correlation (Spearman's *rho* = .51) between orthographic similarity in cognates and total duration in written production. As can be seen in Figure 2, cognates with very strong orthographic overlap were typed as fast as control words, whereas cognates with less overlap were typed more slowly than controls (i.e., cognate interference effect). Furthermore, there was also cognate facilitation in terms of accuracy, in the sense that participants made fewer errors when producing cognates than control words. Finally, error analyses showed that cognates were more susceptible to L1 translation-related errors in comparison to controls.

To our knowledge, the current study is the first to report cognate facilitation effects in typewritten picture naming. Interestingly, the facilitation at onset of production is very similar across modalities, which suggests that the cognate facilitation effect is purely lexical in nature and hence centrally situated. The current findings indicate that bilinguals largely rely on similar processes during word retrieval in L2 spoken and written production.

What about cognate interference effects in whole word production? In the current study, there was no difference in the duration of producing cognates vs. control words in either modality. As such, we found no evidence for the post-onset cognate interference effect that was proposed by Muscalu and Smiley (2019). However, when we took cross-lingual orthographic similarity of the cognates into account (i.e., Levenshtein distance), there was cognate interference for cognates with less cross-lingual overlap, but not for those with much overlap (in comparison with matched controls). Such interference was only observed in the written group. Because the effect was different for very similar and more dissimilar cognates, the effect of cognate status in the total duration for the written group may no longer be visible in the main analysis. The effect of orthographic similarity on total duration in the typing of cognates suggests that when there is limited orthographic overlap between cognates, there is cognate interference at the sub-lexical level (see also Martin & Nozari, 2020). One important difference between our study and the one conducted by Muscalu and Smiley (2019) is the type of task that was used (i.e., picture naming vs. translation, respectively). Indeed, in a translation task, the L1 word form is explicitly given, whereas in a picture naming task, the L1 word form is only indirectly activated. As such, bilinguals may experience stronger interference from the L1 orthographic form when writing down L2 words during a translation task.

In terms of accuracy, there was also cognate facilitation in the sense that participants were more accurate in the production of cognates compared to control words. Other studies investigating the cognate effect in spoken picture naming yielded mixed results in the sense that they either found facilitation in terms of accuracy (Costa et al., 2000, Experiment 2; Hoshino & Kroll, 2008), found no facilitation (Costa et al., 2000, Experiment 1; Starreveld, de Groot, Rossmark & Van Hell, 2014), or did not compare the errors between cognates and control words because error incidence was very low (Christoffels et al., 2006; 2007).

However, most of the mentioned studies did not elaborate on this matter. One reason for the limited attention toward accuracy could be that most of the studies included a familiarization phase of the picture names in order to reduce the chance of errors. The presence or absence of such a phase may also explain why some studies found cognate facilitation effects in accuracy, whereas others did not. Still, spontaneous errors in L2 picture naming can provide important information about the realization of L2 words in bilinguals, and especially how their L1 affects this process.

In addition, the detailed analysis of errors in the current study revealed that beside the larger number of synonym errors for control words, cognates were more often influenced by their L1 equivalents and more prone to orthographic/phonological errors (e.g., ‘shelp’ instead of ‘shell’ because of the Dutch ‘schelp’, an error that occurred in both spoken and written modality). Hence, it seems that the processing facilitation of cognates comes at the expense of an increased vulnerability to (or interference from) erroneous L1 influences. These influences can be considered as additional evidence for cognate interference effects on the sub-lexical level. Indeed, once the bilingual speaker starts typing the word, it can be assumed that the stage of lexical selection is complete and that interference following the onset reflects sub-lexical processes (cf. Muscalu & Smiley, 2019). Concretely, the co-activation of both L1 and L2 features of a target word (i.e., lexical forms and their associated phonology and orthography) causes competition between its L1 and L2 representations. When L1 and L2 representations are identical, there will be no competition between them, but when they are different, there will be interference. This interference is not only reflected in slower RTs, but also in a higher proportion of L1-related erroneous responses in cognates vs. controls. According to interactive activation accounts of language production (e.g., Dell, 1986), such L1-related errors may be the outcome of two types of processes. A first type arises when the cascading activation from the lexical layer into the phonological and orthographic layers leads to stronger activation for the L1 vs. the L2 representations, and feedback to the lexical layer alters the lexical selection (i.e., L1 word form instead of L2 word form). In this case, a bilingual may produce the word of the non-target language (e.g., "schelp" instead of "shell"). A second type arises when there is stronger activation for L1 vs. L2 phonological and orthographic representations, but feedback to the lexical layer does not alter the lexical selection. In that case, the competition between L1 and L2 phonemes and graphemes influences the subsequent selection of motor programs, resulting in code mixing errors (e.g., "shelp" instead of "shell"). The first type of errors may be classified as lexical and the second as sub-lexical.

The L1 influences on the sub-lexical level support the idea that both languages are activated in parallel, but contradict the hypothesis that the cognate effect is merely a learning artefact that does not entail the activation of L1 features (Costa et al., 2017). Another finding that challenges the assumption of the learning account is that while processing cognates reduces switching costs in language-switching tasks, this effect disappears or even reverses when the same cognate stimulus is presented repeatedly throughout the experiment (Li & Gollan, 2018). Such context-dependent effects are hard to explain in terms of the learning account and support the idea of co-activation of both languages. Still, it has to be noted that the extent to which there is co-activation may depend on the context and the task. For instance, bilinguals may have stronger activation of the task-irrelevant language in a language switching task, compared to a monolingual task (see Kroll, Bobb, & Wodniecka, 2006, for a review). As a consequence, one would expect more cross-lingual activation in the former type of task compared to the latter. However, even in a monolingual sentence context, cognate interference effects can be observed when the processing demands are increased (Martin & Nozari, 2020).

The larger number of L1 translation-related errors in the current study is in line with the triggering hypothesis, which states that encountering a cognate word triggers activation of the non-target language in bilinguals (Broersma & De Bot, 2006; Clyne, 1967). Note that the non-target language may be activated for control words as well (see below), but for cognates the activation may be much higher, resulting in more interference. Nevertheless, the responses were more accurate in general for cognates.

L1 influences regarding errors were not limited to cognates, but appeared also in control words, for instance, "flyer" instead of *kite* (Dutch: *vlieger*, derived from the verb *vliegen* – to fly), "mailbus" instead of *mailbox* (Dutch: *brievenbus*), "flathermouse" instead of *bat* (Dutch: *vleermuis*), or "fabric" instead of *factory* (Dutch: *fabriek*). These examples indicate that the L1 word form is often used as retrieval cue for the L2 word, a strategy that is often used in L2 learners (de Groot & Keijzer, 2000). For cognates, this might be a successful strategy, but less so for words that have a different word form in the two languages. Note that an exploratory analysis of the onset RTs in L1 translation-related errors (cognates: $N = 68$; controls: $N = 53$) showed that these RTs were significantly lower in cognates compared to controls ($t(22.17)=3.75$, $p < .01$), which indicates that the higher number of L1 translation-related errors in cognates vs. controls is not the result of top-down guessing strategies that are specific to cognates (e.g., the speaker cannot recall the exact word, but knows that it was similar to the L1 form). Indeed, the fact that the L1 translation-related errors are produced faster in cognates vs. controls shows that the production is more automatic in the former type of words and this favours the parallel activation hypothesis.

In general, the spoken group committed more lexical errors (i.e., semantic errors and synonyms), but fewer sub-lexical errors (phonological/orthographic errors) in comparison with the written group. This indicates that writing is more vulnerable to interference of competing graphemes/phonemes compared to speech (see also Berg, 2002). Interestingly, the vulnerability to L1 interference (i.e., L1 translation-related errors) is similar in both modalities, which suggests that the activation of L1 is comparable during L2 speech and writing. Furthermore, the accuracy was lower and picture naming response latencies tended to be slower in typewriting. The lower accuracy for typewriting is not surprising, given that the incidence of typing errors on the word level is much higher in general compared to speech errors (see Berg, 2002). Slower reaction times were also found in Bonin and Fayol's (2002) and Baus et al.'s (2013) study, who argued that this slowing is unlikely due to differences in the access of conceptual information between both modalities, but rather to the lower degree of automaticity of the typewriting process in comparison with speech. Our findings are in line with this idea, given that despite the slower reaction times in writing, cognate facilitation effects are similar across modalities. Indeed, words that are more easily accessed (such as cognates) yield a similar facilitation in terms of lexical access in both speech and writing.

A limitation of the current study might be that the two groups of bilinguals show a difference in some of the self-rating measures, in the sense that participants in the written group rated themselves somewhat higher for Dutch reading and speaking skills, and for English reading skills. However, there was no significant difference between the two groups for the objective measures of language skills (i.e., LexTALE scores). The higher ratings in the written group may be a result of the fact that this group was somewhat older in general (although there was no significant difference in age between groups) and perhaps more confident of their language skills. Nevertheless, the written group was slower and made more errors than the spoken group, so if participants in the former group would be more proficient in English (and Dutch), this cannot explain the observed differences across modalities. Another limitation is that our findings about writing are based on typewriting, but do not necessarily generalize to handwriting, which requires more in-depth processing of the graphemic forms (James & Engelhardt, 2012). Hence, it might be interesting to compare the cognate effect in hand- and typewriting in future research. Another possible follow-up study could look into the individual contribution of orthographic and phonological representations in spoken and written production by comparing cognates that have both initial phoneme and grapheme overlap in L1 and L2 (e.g., heart-hart) with cognates that start with the same phoneme, but a different grapheme (e.g., cat-kat).

Conclusions

In sum, the current study shows that the cognate facilitation effect at the onset of bilingual word production is largely modality-independent, in line with a central locus for cognate facilitation in bilinguals (e.g., Costa et al., 2000; 2017). In addition to this lexically situated cognate effect, there are also sub-lexical cognate effects (at least in writing): cognates with large orthographic overlap across languages show no effect, whereas cognates with less overlap show interference effects. Finally, the analyses of the errors indicate that bilinguals activate both their L1 and L2 during L2 speech and writing, which contradicts Costa et al.'s (2017) interpretation of the cognate facilitation effect as a learning artefact.

References

Afonso, O., & Álvarez, C. J. (2011). Phonological effects in handwriting production: Evidence from the implicit priming paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1474–1483. https://doi.org/10.1037/a0024515

Alario, F.-X., Costa, A., Ferreira, V. S., & Pickering, M. J. (2006). Architectures, representations and processes of language production. *Language and Cognitive Processes*, *21*(7–8), 777–789. https://doi.org/10.1080/016909600824112

Alario, F.-X., Perre, L., Castel, C., & Ziegler, J. C. (2007). The role of orthography in speech production revisited. *Cognition*. https://doi.org/10.1016/j.cognition.2006.02.002

Almeida, J., Knobel, M., Finkbeiner, M., & Caramazza, A. (2007). The locus of the frequency effect in picture naming: When recognizing is not enough. *Psychonomic Bulletin & Review*, *14*(6), 1177–1182. https://doi.org/10.3758/BF03193109

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

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Berg, T. (2002). Slips of the typewriter key. *Applied Psycholinguistics*, *23*(2), 185–207. https://doi.org/10.1017/S0142716402002023

Bi, Y., Wei, T., Janssen, N., & Han, Z. (2009). The contribution of orthography to spoken word production: Evidence from Mandarin Chinese. *Psychonomic Bulletin & Review*, *16*(3), 555–560. https://doi.org/10.3758/PBR.16.3.555

Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, *93*(1), 89–114. https://doi.org/10.1348/000712602162463

Bonin, P., & Fayol, M. (2000). Writing words from pictures: What representations are activated, and when? *Memory & Cognition*, *28*(4), 677–689. https://doi.org/10.3758/BF03201257

Bonin, P., & Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. *European Journal of Cognitive Psychology*, *14*(3), 289–313. https://doi.org/10.1080/09541440143000078

Bonin, P., Méot, A., Lagarrigue, A., & Roux, S. (2015). Written object naming, spelling to dictation, and immediate copying: Different tasks, different pathways? *Quarterly Journal of Experimental Psychology*, *68*(7), 1268–1294. https://doi.org/10.1080/17470218.2014.978877

Bonin, P., Peereman, R., & Fayol, M. (2001). Do phonological codes constrain the selection of orthographic codes in written picture naming? *Journal of Memory and Language*, *45*(4), 688–720. https://doi.org/10.1006/jmla.2000.2786

Broersma, M., & De Bot, K. (2006). Triggered codeswitching: A corpus-based evaluation of the original triggering hypothesis and a new alternative. *Bilingualism: Language and Cognition*, *9*(1), 1–13. https://doi.org/10.1017/S1366728905002348

Bürkner, P.-C. (2017). brms : An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Caramazza, A. (1997). How Many Levels of Processing Are There in Lexical Access? *Cognitive Neuropsychology*, *14*(1), 177–208. https://doi.org/10.1080/026432997381664

Christoffels, I. K., de Groot, A. M. B., & Kroll, J. F. (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency☆. *Journal of Memory and Language*, *54*(3), 324–345. https://doi.org/10.1016/j.jml.2005.12.004

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, *1147*, 192–208. https://doi.org/10.1016/j.brainres.2007.01.137

Clyne, M. G. (1967). *Transference and triggering: Observations on the language assimilation of postwar German-speaking migrants in Australia.* Martinus Nijhoff.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The Cognate Facilitation Effect: Implications for Models of Lexical Access. *Journal of Experimental Psychology: Learning Memory and Cognition*. https://doi.org/10.1037/0278-7393.26.5.1283

Costa, A., Pannunzi, M., Deco, G., & Pickering, M. J. (2017). Do Bilinguals Automatically Activate Their Native Language When They Are Not Using It? *Cognitive Science*. https://doi.org/10.1111/cogs.12434

Damian, M. F., & Bowers, J. S. (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory and Language*, *49*(1), 119–132. https://doi.org/10.1016/S0749-596X(03)00008-1

Damian, M. F., & Bowers, J. S. (2009). Assessing the role of orthography in speech perception and production: Evidence from picture-word interference tasks. *European Journal of Cognitive Psychology*. https://doi.org/10.1080/09541440801896007

Damian, M. F., Dorjee, D., & Stadthagen-Gonzalez, H. (2011). Long-term repetition priming in spoken and written word production: Evidence for a contribution of phonology to handwriting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 813–826. https://doi.org/10.1037/a0023260

Damian, M. F., & Qu, Q. (2013). Is handwriting constrained by phonology? Evidence from Stroop tasks with written responses and Chinese characters. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00765

de Groot, A. M. B., & Keijzer, R. (2000). What Is Hard to Learn Is Easy to Forget: The Roles of Word Concreteness, Cognate Status, and Word Frequency in Foreign-Language Vocabulary Learning and Forgetting. *Language Learning*, *50*(1), 1–56. https://doi.org/10.1111/0023-8333.00110

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. https://doi.org/10.1016/0749-596X(91)90012-9

De Rosario-Martinez, H. (2013). phia: Post-Hoc Interaction Analysis. In *Available: http://CRAN.R-project.org/package5phia*.

DeLeon, J., Gottesman, R. F., Kleinman, J. T., Newhart, M., Davis, C., Heidler-Gary, J., Lee, A., & Hillis, A. E. (2007). Neural regions essential for distinct cognitive processes underlying picture naming. *Brain*, *130*(5), 1408–1422. https://doi.org/10.1093/brain/awm011

Dell, G. S. (1986). A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review*. https://doi.org/10.1037/0033-295X.93.3.283

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Downey, S. S., Hallmark, B., Cox, M. P., Norquest, P., & Lansing, J. S. (2008). Computational Feature-Sensitive Reconstruction of Language Relationships: Developing the ALINE Distance for Comparative Historical Linguistic Reconstruction. *Journal of Quantitative Linguistics*, *15*(4), 340–369. https://doi.org/10.1080/09296170802326681

Downey, S. S., Sun, G., & Norquest, P. (2017). alineR: an R Package for Optimizing Feature-Weighted Alignments and Linguistic Distances. *The R Journal*, *9*(1), 138. https://doi.org/10.32614/RJ-2017-005

Duyck, W., & De Houwer, J. (2008). Semantic access in second-language visual word processing: Evidence from the semantic Simon paradigm. *Psychonomic Bulletin & Review*, *15*(5), 961–966. https://doi.org/10.3758/PBR.15.5.961

Duyck, W., Desmet, T., Verbeke, L. P. C., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 488–499. https://doi.org/10.3758/BF03195595

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 663–679. https://doi.org/10.1037/0278-7393.33.4.663

Ferrand, L., Grainger, J., & Segui, J. (1994). A study of masked form priming in picture and word naming. *Memory & Cognition*. https://doi.org/10.3758/BF03200868

Geschwind, N. (1974). Problems in the anatomical understanding of the aphasias. In A. L. Benton (Ed.), *Contributions to Clinical Neuropsychology* (pp. 431–451). Aldine. https://doi.org/10.1007/978-94-010-2093-0_20

Goldrick, M., Ferreira, V. S., & Miozzo, M. (2014). The Oxford Handbook of Language Production. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Eds.), *The Oxford Handbook of Language Production*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199735471.001.0001

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67–81. https://doi.org/10.1017/S1366728998000133

Hillis, A. E., Rapp, B., Romani, C., & Caramazza, A. (1990). Selective impairment of semantics in lexical processing. *Cognitive Neuropsychology*, *7*(3), 191–243. https://doi.org/10.1080/02643299008253442

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition*, *106*(1), 501–511. https://doi.org/10.1016/j.cognition.2007.02.001

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*. https://doi.org/10.1016/j.cognition.2002.06.001

James, K. H., & Engelhardt, L. (2012). The effects of handwriting experience on functional brain development in pre-literate children. *Trends in Neuroscience and Education*. https://doi.org/10.1016/j.tine.2012.08.001

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 824–843. https://doi.org/10.1037/0278-7393.20.4.824

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism*. https://doi.org/10.1017/S1366728906002483

Kroll, J. F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language*, *33*(2), 149–174. https://doi.org/10.1006/jmla.1994.1008

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2011). Knowledge of a second language influences auditory word recognition in the native language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 952–965. https://doi.org/10.1037/a0023217

Lauro, J., & Schwartz, A. I. (2017). Bilingual non-selective lexical access in sentence contexts: A meta-analytic review. *Journal of Memory and Language*, *92*, 217–233.

https://doi.org/10.1016/j.jml.2016.06.010

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. https://doi.org/10.3758/s13428-011-0146-0

Lemhöfer, K., Dijkstra, T., & Michel, M. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes*, *19*(5), 585–611. https://doi.org/10.1080/01690960444000007

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–75. https://doi.org/10.1017/S0140525X99001776

Li, C., & Gollan, T. H. (2018). Cognates facilitate switches and then confusion: Contrasting effects of cascade versus feedback on language selection. *Journal of Experimental Psychology: Learning Memory and Cognition*. https://doi.org/10.1037/xlm0000497

Lotto, L., & de Groot, A. M. B. (1998). Effects of Learning Method and Word Type on Acquiring Vocabulary in an Unfamiliar Language. *Language Learning*, *48*(1), 31–69. https://doi.org/10.1111/1467-9922.00032

Luria, A. R. (1970). *Traumatic Aphasia*. DE GRUYTER MOUTON. https://doi.org/10.1515/9783110816297

Martin, C. D., & Nozari, N. (2020). Language control in bilingual production: Insights from error rate and error type in sentence production. *Bilingualism: Language and Cognition*, 1–15. https://doi.org/10.1017/S1366728920000590

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

Meade, G., Midgley, K. J., Dijkstra, T., & Holcomb, P. J. (2018). Cross-language

Neighborhood Effects in Learners Indicative of an Integrated Lexicon. *Journal of Cognitive Neuroscience*, *30*(1), 70–85. https://doi.org/10.1162/jocn_a_01184

Miceli, G., Benvegnu, B., Capasso, R., & Caramazza, A. (1997). The independence of phonological and orthographic lexical forms: Evidence from aphasia. *Cognitive Neuropsychology*, *14*(1), 35–69. https://doi.org/10.1080/026432997381619

Muscalu, L. M., & Smiley, P. A. (2019). The illusory benefit of cognates: Lexical facilitation followed by sublexical interference in a word typing task. *Bilingualism: Language and Cognition*, *22*(04), 848–865. https://doi.org/10.1017/S1366728918000792

Oppenheim, G., Wu, Y. J., & Thierry, G. (2018). Found in Translation: Late Bilinguals Do Automatically Activate Their Native Language When They Are Not Using It. *Cognitive Science*, *42*(5), 1700–1713. https://doi.org/10.1111/cogs.12618

Palma, P., Whitford, V., & Titone, D. (2019). Cross-Language Activation and Executive Control Modulate Within-Language Ambiguity Resolution: Evidence From Eye Movements. *Journal of Experimental Psychology: Learning Memory and Cognition*. https://doi.org/10.1037/xlm0000742

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*. https://doi.org/10.3758/s13428-018-01193-y

Perret, C., & Laganaro, M. (2012). Comparison of electrophysiological correlates of writing and speaking: A topographic ERP analysis. *Brain Topography*. https://doi.org/10.1007/s10548-011-0200-3

Perret, C., & Laganaro, M. (2013). Why are written picture naming latencies (not) longer than spoken naming? *Reading and Writing*, *26*(2), 225–239. https://doi.org/10.1007/s11145-012-9365-8

Pivneva, I., Mercier, J., & Titone, D. (2014). Executive control modulates cross-language lexical activation during L2 reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning Memory and Cognition*. https://doi.org/10.1037/a0035583

Qu, Q., Damian, M. F., & Li, X. (2016). Phonology contributes to writing: Evidence from a masked priming task. *Language, Cognition and Neuroscience*, *31*(2), 251–264. https://doi.org/10.1080/23273798.2015.1091086

Qu, Q., Damian, M. F., Zhang, Q., & Zhu, X. (2011). Phonology Contributes to Writing. *Psychological Science*, *22*(9), 1107–1112. https://doi.org/10.1177/0956797611417001

R Core Team. (2016). R: A language and environment for statistical computing. In *R Foundation for Statistical Computing*.

Rapp, B., Benzing, L., & Caramazza, A. (1997). The autonomy of lexical orthography. *Cognitive Neuropsychology*. https://doi.org/10.1080/026432997381628

Rapp, B., & Caramazza, A. (1997). From Graphemes to Abstract Letter Shapes: Levels of Representation in Written Spelling. *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/10.1037/0096-1523.23.4.1130

Roelofs, A. (2006). The influence of spelling on phonological encoding in word reading, object naming, and word generation. *Psychonomic Bulletin & Review*, *13*(1), 33–37. https://doi.org/10.3758/BF03193809

Severens, E., Lommel, S. Van, Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, *119*(2), 159–187. https://doi.org/10.1016/j.actpsy.2005.01.002

Shelton, J. R., & Weinrich, M. (1997). Further Evidence of a Dissociation between Output Phonological and Orthographic Lexicons: A Case Study. *Cognitive Neuropsychology*, *14*(1), 105–129. https://doi.org/10.1080/026432997381637

Starreveld, P. A., de Groot, A. M. B., Rossmark, B. M. M., & Van Hell, J. G. (2014). Parallel language activation during word processing in bilinguals: Evidence from word production in sentence context. *Bilingualism: Language and Cognition*, *17*(2), 258–276. https://doi.org/10.1017/S1366728913000308

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhp153

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language*, *64*(1), 88–107. https://doi.org/10.1016/j.jml.2010.08.006

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does Bilingualism Change Native-Language Reading? *Psychological Science*, *20*(8), 923–927. https://doi.org/10.1111/j.1467-9280.2009.02389.x

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin and Review*. https://doi.org/10.3758/BF03196335

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, *39*(3), 458–483. https://doi.org/10.1006/jmla.1998.2584

Woutersen, M., De Bot, K., & Weltens, B. (1995). The bilingual lexicon: Modality effects in processing. *Journal of Psycholinguistic Research*, *24*(4), 289–298. https://doi.org/10.1007/BF02145058

Zhang, Q., & Damian, M. F. (2012). Effects of Orthography on Speech Production in Chinese. *Journal of Psycholinguistic Research*, *41*(4), 267–283. https://doi.org/10.1007/s10936-011-9193-z

Zhang, Q., & Wang, C. (2015). Phonology is not accessed earlier than orthography in Chinese written production: evidence for the orthography autonomy hypothesis. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00448

Zhang, Q., & Wang, C. (2016). The temporal courses of phonological and orthographic encoding in handwritten production in Chinese: An ERP study. *Frontiers in Human Neuroscience*. https://doi.org/10.3389/fnhum.2016.00417

Footnotes

[1] Originally, both experiments were conducted with E-Prime, but later on we discovered an issue in the written data. Therefore, we tested a new group of participants in the written condition using PsychoPy software and replaced the old data with this new data.
[2] For those participants who were more familiar with a QWERTY keyboard, a QWERTY version of the experiment was provided.
[3] Errors that were self-interrupted, such as "h… house", were coded as disfluencies, and thus incorrect, to avoid distortion of reaction times.
[4] Synonyms were scored as an error even though they were a correct description of the picture, because the target words were matched on several features, whereas the synonyms were not.
[5] Random slopes were left out because brms (or any other existing package for Bayesian linear effects models) does not deal well with them.

Table 1. *Group comparisons of participant characteristics*.

characteristic	Spoken *M (SD)*	written *M (SD)*
Age	20.5 (3.36)	22.6 (5.74)
sex (N of females)	32	27
LexTALE Dutch	88.78 (5.79)	89.34 (7.54)
LexTALE English	81.63 (6.97)	83.22 (8.19)
self-ratings Dutch		
- speaking*	6.1 (0.67)	6.4 (0.68)
- reading*	6.3 (0.74)	6.6 (0.55)
- writing	6.1 (0.78)	6.3 (0.55)
- proficiency	6.0 (0.62)	6.3 (0.55)
self-ratings English		
- age of acquisition	10.5 (2.44)	10.3 (2.77)
- speaking	5.3 (0.88)	5.6 (0.82)
- reading*	5.7 (0.76)	6.0 (0.56)
- writing	4.9 (0.93)	5.2 (0.79)
- proficiency	5.3 (0.73)	5.5 (0.65)
self-ratings French		
- age of acquisition	10.0 (1.38)	9.9 (1.28)
- speaking	3.8 (1.04)	3.6 (1.04)
- reading	4.2 (1.19)	4.6 (0.87)
- writing	3.7 (0.99)	3.5 (0.96)
- proficiency	3.9 (0.91)	3.8 (0.93)

Note. *$p < .05$

Table 2. *Proportion of errors for each condition**.

	Spoken		written	
error type	cognate	Control	cognate	control
disfluency	0.04	0.03	-	-
typing error	-	-	0.26	0.24
L1 translation-related error	0.14	0.10	0.18	0.10
phonological/orthographic error	0.05	0.01	0.15	0.12
semantic error	0.42	0.45	0.22	0.29
synonym	0.35	0.41	0.19	0.25

*Some errors could belong to more than one category.

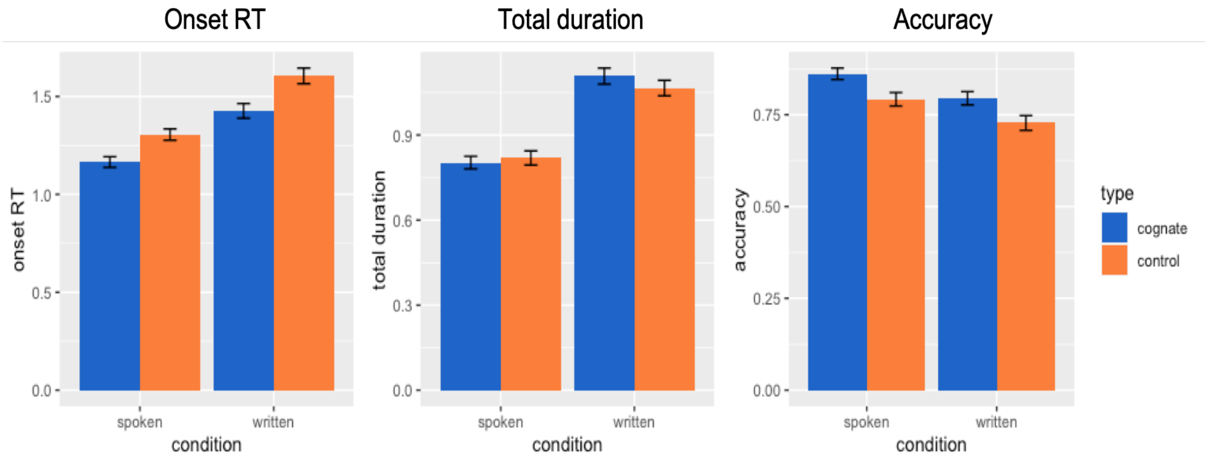

Figure 1. *Cognate effects in onset RTs, total duration and accuracy for the spoken and written modality.*

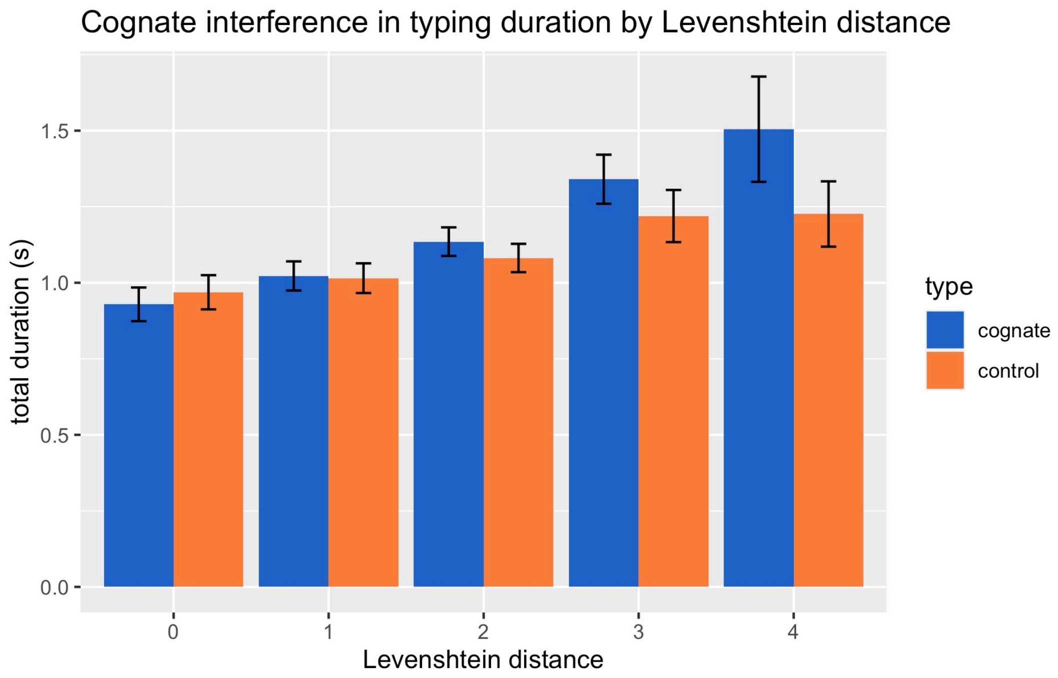

Figure 2. *Cognate interference effects in typing duration as function of Levenshtein distance between cognates and their translations (e.g., "ship"-"schip" has Levenshtein distance 1, whereas "cookie"-"koek" has Levenshtein distance 4).*

Appendix

Appendix A. *List of words and their characteristics.*

Cognates					Matched controls				
English	Dutch	log freq/ million	neighbours (N)	rank 1st letter	English	Dutch	log freq/ million	neighbours (N)	rank 1st letter
apple	appel	1.51	2	6	*fence*	hek	1.52	1	10
bear	beer	2.11	19	4	*tear*	traan	2.11	16	5
bed	bed	2.46	14	4	*boy*	jongen	2.57	17	4
bell	bel	1.63	13	4	*tail*	staart	1.63	13	5
boat	boot	1.90	9	4	*coat*	jas	1.86	10	2
book	boek	2.63	12	4	*fire*	vuur	2.34	13	10
bread	brood	1.89	6	4	*nurse*	verpleegster	1.76	3	19
bridge	brug	1.84	2	4	*stairs*	trap	1.67	0	1
clock	klok	1.62	8	2	*witch*	heks	1.53	7	15
cook	kok	1.97	12	2	*lock*	slot	1.90	14	12
cookie	koek	0.70	3	2	*wallet*	portefeuille	0.90	3	15
fish	vis	2.31	4	10	*tree*	boom	2.31	5	5
fist	vuist	1.45	7	10	*whip*	zweep	1.45	9	15
flag	vlag	1.46	9	10	*belt*	riem	1.48	11	4
foot	voet	2.54	12	10	*wall*	muur	2.35	13	15
glass	glas	2.18	3	11	*dress*	kleed	2.25	5	7
hair	haar	2.32	5	13	*city*	stad	2.46	2	2
hammer	hamer	1.38	2	13	*bucket*	emmer	1.32	1	4
hand	hand	2.92	9	13	*road*	straat	2.40	9	8
hat	hoed	1.85	20	13	*bag*	zak	1.93	22	4
heart	hart	2.23	0	13	*chair*	stoel	2.17	2	2
helmet	helm	1.11	1	13	*knight*	ridder	1.11	0	22
hook	haak	1.76	11	13	*roof*	dak	1.78	8	8
house	huis	2.85	6	13	*watch*	uurwerk	2.48	8	15
ladder	ladder	1.20	2	12	*turkey*	kalkoen	1.18	0	5
moon	maan	1.79	11	9	*wing*	vleugel	1.80	10	15
mouse	muis	1.28	8	9	*shark*	haai	1.32	8	1
nest	nest	1.30	14	19	*duck*	eend	1.32	15	7
net	net	1.60	15	19	*pig*	varken	1.66	14	3
owl	uil	0.85	4	17	*bra*	beha	0.78	6	4
pan	pan	1.51	21	3	*bat*	vleermuis	1.34	22	4
pumpkin	pompoen	0.30	1	3	*mailbox*	brievenbus	0.30	0	9
ring	ring	2.18	10	8	*rock*	steen	2.14	12	8
rocket	raket	1.20	6	8	*waiter*	ober	1.36	4	15
sheep	schaap	1.62	7	1	*snake*	slang	1.43	5	1
shell	schelp	1.72	5	1	*skirt*	rok	1.54	4	1
shoe	schoen	1.92	6	1	*leaf*	blad	1.94	7	12
shoulder	schouder	2.15	0	1	*picture*	afbeelding	2.25	0	3
sock	sok	1.30	12	1	*deer*	hert	1.08	12	7

star	ster	2.05	8	1	*king*	koning	2.01	9	22
step	step	2.22	6	1	*bird*	vogel	2.03	4	4
swan	zwaan	0.90	7	1	*kite*	vlieger	0.70	6	22
sword	zwaard	1.26	2	1	*glove*	handschoen	1.32	3	11
thumb	duim	1.49	1	5	*queen*	koningin	1.72	3	23
walnut	walnoot	0.70	0	15	*napkin*	servet	0.85	0	19
wheel	wiel	1.73	0	15	*knife*	mes	1.69	0	22
wolf	wolf	1.11	3	15	*frog*	kikker	0.95	3	10
worm	worm	1.30	8	15	*doll*	pop	1.43	11	7
MEAN		1.69	7.21	7.83			1.66	7.29	9.25

		Fillers		
English	Dutch	log freq/ million	neighbours (N)	rank 1st letter
ant	mier	1.08	7	6
arrow	pijl	1.15	0	6
ax	bijl	0.95	13	6
boot	laars	1.61	17	4
box	doos	2.03	13	4
branch	tak	2.00	2	4
butcher	slager	1.00	1	4
button	knoop	1.54	2	4
cage	kooi	1.28	12	2
candle	kaars	1.23	2	2
car	auto	2.55	19	2
cheese	kaas	1.51	0	2
cherry	kers	0.95	3	2
chest	borstkas	1.71	3	2
chimney	schoorsteen	1.04	0	2
cloud	wolk	1.81	2	2
curtains	gordijnen	0.00	0	2
desk	bureau	1.98	2	7
dog	hond	2.09	14	7
donkey	ezel	1.18	1	7
eagle	arend	1.00	0	14
egg	ei	1.94	5	14
factory	fabriek	1.77	0	10
farm	boerderij	1.99	7	10
flower	bloem	2.02	2	10
girl	meisje	2.66	3	11
gun	geweer	2.02	17	11
key	sleutel	1.96	6	22
mirror	spiegel	1.73	0	9
monkey	aap	1.28	1	9
parrot	papegaai	0.60	1	3
peanut	pinda	0.78	0	3
pencil	potlood	1.34	0	3
rabbit	konijn	1.30	0	8

scarf	sjaal	1.11	5	1
scissors	schaar	0.70	0	1
smoke	rook	1.98	3	1
unicorn	eenhoorn	0.00	0	18
window	raam	2.32	1	15
woman	vrouw	2.94	1	15
MEAN		1.50	4.13	6.63