Running head: Cross-lingual Lexical Similarity in Translation

Semantic Access in Number Word Translation: the Role of Cross-Lingual Lexical Similarity

Wouter Duyck[1]

Marc Brysbaert[2]

[1]Ghent University, Belgium

[2]Royal Holloway University of London, UK.

Correspondence Address:

Wouter Duyck

Department of Experimental Psychology

Ghent University

Henri Dunantlaan 2

B-9000 Ghent (Belgium)

E-mail: wouter.duyck@UGent.be

Tel.: 32-9-264-64-35 - Fax: 32-9-264-64-96

Abstract

The revised hierarchical model of bilingualism (e.g., Kroll & Stewart, 1994) assumes that second language (L2) words primarily access semantics through their first language (L1) translation equivalents. Consequently, backward translation from L2 to L1 should not imply semantic access but occurs through lexical wordform associations. However, recent research with Dutch-French bilinguals showed that both backward and forward translation of number words yields a semantic number magnitude effect (Duyck & Brysbaert, 2004), providing evidence for strong form-to-meaning mappings of L2 number words. In two number word translation experiments with Dutch-English-German trilinguals, the present study investigated whether semantic access in L1-L2 and L1-L3 number word translation depends on lexical similarity of the languages involved. We found that backward translation from these more similar language pairs to L1 still yields a semantic magnitude effect, whereas forward translation does not, in contrast with the Dutch-French results of Duyck and Brysbaert (2004). We argue against a dual route model of word translation and suggest that the degree of semantic activation in translation depends on lexical form overlap between translation equivalents.

Keywords: bilingualism, translation, conceptual, semantic mediation, number word, lexical overlap

Semantic Access in Number Word Translation: the Role of Cross-lingual Lexical Similarity

In most models of monolingual language processing, a distinction is made between form-level (lexical) word representations, and semantic representations. Because readers are able to extract meaning from written text very fast and without too much effort, these models often have strong mappings between native language words' forms and their underlying meaning. In the literature on bilingualism however, the influential revised hierarchical model (RHM) of Kroll and Stewart (1994, see Figure 1), does not postulate such strong form-to-meaning mappings for second language (L2) words (at least not at moderate L2 proficiency levels). Because these words forms are often acquired by associating them with their corresponding translation equivalents, it is assumed that L2 words may only access semantics through their L1 counterparts. As a consequence, the RHM assumes that backward translation from L2 to L1 occurs through word-word associations at the lexical form level, without access to semantics. Contrastingly, because of the strong form-to-meaning mappings of L1 words, forward translation from L1 to L2 is expected to yield semantic activation. According to the developmental assumption of the RHM, this asymmetric lexicosemantic organization is expected to disappear in high levels of L2 proficiency. Evidence the RHM comes for instance from Sholl, Sankaranarayanan and Kroll (1995), who found that forward translation is significantly facilitated by prior presentation of translation target's pictures. This semantic priming effect does not occur in backward translation, suggesting a lexical-form backward translation process. Evidence for the developmental assumption of the RHM comes from Talamas, Kroll, and Dufour (1999). They found greater interference of semantically related distractors in a translation recognition task when the participants were highly proficient in L2, whereas less proficient bilinguals suffered more interference from form-related words. For a more detailed review of the findings supporting the different assumptions of the RHM, we refer to Kroll and colleagues (Kroll & de Groot, 1997; Kroll & Tokowicz, 2005).

Insert Figure 1 about here

Although this asymmetry hypothesis has provided an influential view on bilingual lexicosemantic organization during the last decade, recently a few findings have been reported that may not easily be explained within the existing theoretical framework of the RHM. First, using a bilingual semantic Stroop task with unbalanced Dutch-English bilinguals, La Heij, Hooglander, Kerling and Van der Velden (1996) found that congruent colour words (for which the ink colour corresponded to the word), were translated faster than incongruent colour words, in both directions of translation. The RHM does not predict such a semantic effect in backward translation of L2 colour words, because they are not assumed to have strong form-to-meaning mappings. In further experiments, these authors also found that both backward and forward translation is facilitated by pictures (e.g. a table) semantically related to the target (e.g. CHAIR) (for similar results, see Bloem & La Heij, 2003). Secondly, using a translation recognition task, Altarriba and Mathis (1997) reported that monolinguals who were trained on a set of English – Spanish word pairs, made more errors on both lexically and semantically related false translations than on unrelated words. According to the RHM's developmental hypothesis, the semantic translation distracter effect should not occur so early after L2 word form acquisition. Also, Altarriba and Mathis reported a bilingual Stroop effect similar to that found by La Heij et al. (1996) using the same L2 training procedure as in their first experiment.

In the picture experiments above, as argued by Kroll and De Groot (1997), the context provided by the distractor pictures may have artificially boosted activation in the semantic system during translation, whereas the RHM was designed to explain out-of-context L2 (and L1) word translation. This criticism does not apply to the symmetric semantic translation effects obtained by Duyck and Brysbaert (2004; see also Duyck & Brysbaert, 2002; Duyck, Lagrou, Gevers, & Fias, in press). Using an out-of-context number word translation task with Dutch-French bilinguals, they found a semantic effect of number magnitude in both directions of translation. It took longer to translate number words representing larger quantities (e.g. *huit* and *acht* [8]) than number words representing smaller quantities (e.g. *deux* and *twee* [2]), independent of lexical effects of word frequency. Interestingly, these effects were also obtained for (so-called 'Estonian') number words that were learned only a few minutes before the translation task.

Because lexicosemantic organization in the RHM is only depending on general L2 proficiency, and does not differ across word types, Duyck and Brysbaert (2004) argued that these findings are problematic for the model as a whole, even though there is independent evidence that the linking between new symbols and meanings is particularly fast for numerical stimuli (Logan & Klapp, 1991; Tzelgov, Yehene, Kotler, & Alon, 2000). Therefore, Duyck and Brysbaert (2004) proposed an extension of the RHM that differs from its predecessor in two ways (see Figure 2). First, translation is not assumed to be a dual-route process that follows either the lexical-form or the semantic route. Instead, translation output is the result of activation coming form both semantic and lexical form representations, which may be more or less activated. Secondly, the strength of form-to-meaning mappings through which this activation is forwarded, is not solely depending on L2 proficiency (as in the RHM), but is also a function of word type. Consequently, backward (and forward) translation may imply a different amount of semantic activation depending on word-level variables. For instance, because number words (or colour words, see the results of La Heij et al., 1996) have a well confined meaning that (virtually) completely overlaps across languages, they will develop strong L2 form-to-meaning mappings earlier than words that have a more diffuse meaning with language-specific connotations (e.g. abstract words). Consequently, their translation will trigger semantic activation more easily, and earlier in the word learning process, than other types of words will. Abstract words for example (e.g., Figure 2, *duty*), are represented by a more fuzzy set of semantic features and may have slightly different meanings in different languages, yielding smaller semantic translation effects (see the concreteness effects on word translation reported by De Groot and colleagues, e.g., de Groot, 1992; de Groot & Comijs, 1995; de Groot & Poot, 1997; de Groot, Dannenburg, & van Hell, 1994).

In this model of Duyck and Brysbaert (2004), not only the connection weights between the lexical-form and the semantic level, but also between L1 and L2 lexical form representations may differ as a function of word characteristics. It is assumed that these connections are stronger for words with a large form overlap (e.g., Figure 2: *ball* - *bal* for an English-Dutch bilingual) than for words with a small lexical form overlap (e.g., Figure 2: *duty* - *plicht* for an English-Dutch bilingual). This lexical form overlap assumption, also present in the Distributed Representations model of Van Hell and De Groot (1998), provides an explanation for the occasional finding that translation equivalents with a large form overlap (cognates) are easier to translate and show less evidence for semantic mediation in translation, than words with no form overlap For instance, De Groot (1992) found that correlations between semantic variables and translation RTs are smaller for noncognates than for cognates. Similarly, De Groot and Comijs (1995) reported that noncognates showed stronger manifestations of semantic access in forward translation, relative to cognates (for similar cognate manipulations, see for example de Groot et al., 1994). Also, in L2 word production, it has often been found that cognates are produced faster than noncognates, a finding which is often attributed to facilitative effects between form representations (e.g., Costa, Santesteban, & Cano, 2005; see also Costa, Caramazza, & Sebastian-Galles, 2000; Gollan & Acenas, 2004). Similar cognate facilitation effects in bilingual aphasia were recently reported by Kohnert (2004). These cognate findings offer support for the assumption that cross-lingual form overlap may indeed influence bilingual word production and translation.

Insert Figure 2 about here

The Present Study

The aim of the present study is twofold. First, it attempts to provide additional, more direct, evidence for the lexical form overlap assumption of Duyck and Brysbaert's (2004) model, manipulating cross-lingual similarity in the same number word translation paradigm. To this end, we will investigate translation performance of Dutch-English-German trilinguals, because English (e.g., *ten*) and German (e.g., *zehn*) number words have more form overlap with their Dutch translation equivalents (e.g., *tien*) than the French number words (e.g., *dix*) investigated by Duyck and Brysbaert (2004). According to the lexical overlap assumption, these more cross-lingually similar, English and German translation equivalents should receive more activation through lexical-form connections, so that this Dutch-English-German translation study should yield smaller semantic number magnitude effects than the Dutch-French study of Duyck and Brysbaert (2004). Because the RHM assumes that semantic activation during translation is depending on translation direction, and not on word-level variables such as lexical form overlap, it would not predict different translation patterns for Dutch-French number words on the one hand (Duyck & Brysbaert, 2004), and Dutch-English or Dutch-German (this study) on the other hand. Also, because L2 and L3 (number) words are not assumed to have strong form-to-meaning mappings, the RHM does not predict a number magnitude effect for backward translation for any language pair (especially not for L3, because it only predicts semantic backward translation effects in high levels of proficiency). The second objective of this study is to test the generalizability of the bilingual model of Duyck and Brysbaert (2004) for multilinguals. Because virtually all translation studies have only investigated translation between L1 and L2, this is one of the first studies to assess semantic mediation in translation between L1 and L3. (but see de Groot & Hoeks, 1995; Francis & Gallard, 2005).

Experiment 1

Based on the reasoning above, we predict that the present experiment will yield semantic effects of number magnitude in the translation conditions, but not in the naming conditions, similar to Duyck and Brysbaert (2004). Because cross-lingual lexical form similarity is larger for the present Dutch-English stimuli however, it may be the case that number magnitude effects are weaker, or even disappear, in certain translation conditions.

Method

Participants

Thirteen first-year university students participated for course requirements. All of them were native Dutch speakers and reported English as their L2. All participants started to learn English in a school setting around the age of 14-15 (formal English courses are mandatory at that age in the Belgian school system), and live in a L1 dominant environment, speaking Dutch at home, at school, with friends, etc. All of them are regularly exposed to their L2 (English) through Belgian popular media and entertainment (music, internet, films, television, etc.). Like almost everybody in Belgium, all participants also have some knowledge of German and French, which were reported respectively as their L3 and L4. Participants were asked to rate their L1 to L4 proficiency on a 7-point Likert scale ranging from ‘very bad’ to ‘very good’ after the actual experiment. Means are reported in Table 1. Mean self-reported L1 (M = 6.77) and L2 proficiency (M = 5.31) differed significantly. Also, L2, L3 and L4 proficiency differed significantly from each other (all ps < .001).

Insert Table 1 about here

Materials

All materials were identical to Duyck and Brysbaert (2004). Stimuli were presented on a standard 15” VGA color monitor, as yellow characters on a black background. Stimulus presentation was computer driven by a PC equipped with a voice key which was connected through the gameport. All Arabic digits, Dutch and English number words representing quantities from 1 to 12 were selected as stimuli. Mean similarity between Dutch and English translation equivalents (M = 406.13) was significantly higher (p < .01) than for the Dutch-French pairs (M = 146.06) used by Duyck and Brysbaert (2004), according to the word similarity measure described by Van Orden (1987)[1].

Design

Similar to Duyck and Brysbaert (2004), the experiment had a 2 (Naming Language: L1 versus L2) x 3 (Stimulus format: Arabic numbers, L1 number words and L2 number words) x 12 (Number Magnitude) full factorial design. All variables were manipulated within-subjects.

Procedure

The procedure was also identical to Duyck and Brysbaert (2004). All participants completed two blocks (L1 naming and L2 naming) of 360 trials. The order of these blocks was counterbalanced across participants. Within each block, 10 series of 36 randomly ordered trials were presented, corresponding to every number magnitude from 1 to 12 in each of the three stimulus formats (Arabic, L1, L2). Hence, the participants did not know which stimulus format would appear before the beginning of each trial. Only naming language was blocked[2]. Each trial started with the presentation of a fixation stimulus ('*') for 500 ms. This was replaced by the target, which remained visible until pronunciation of the target triggered the voice key. The Inter Trial Interval (ITI) was 1000 ms. The experiment lasted for about 45 minutes, including a little break.

Results

Variance Analysis

The proportion of invalid trials due to naming errors or faulty time registration was 3.37%. These trials were excluded from all analyses. Also, outlier RTs that deviated more than 2.5 standard deviations from a participant's mean RT for a given naming language, were excluded from the analyses (0.78 % of the data). An ANOVA was performed with Naming Language, Stimulus Format, and Number Magnitude as repeated measures factors. The dependent variable was the mean RT across correct trials. Mean RTs as a function of Naming Language, Stimulus Format and Number Magnitude are presented in Figure 3. The backward translation condition can be found in the left part of the figure, whereas forward translation is plotted in the right part of the figure.

Insert Figure 3 about here

Similar to Duyck and Brysbaert (2004), the effect of Naming Language was not significant, $F(1, 12) = 1.81$, $MSE = 28637$, $p > .20$. Naming in Dutch took 509 ms, while English naming took 494 ms. This confirms that L2 proficiency was quite high. Planned comparisons showed that backward translation was significantly slower ($M = 548$ ms) than forward translation ($M = 497$ ms), $F(1, 12) = 12.91$, $MSE = 16252$, $p < .01$, as opposed to the predictions based on the RHM.

The main effects of Stimulus Format ($F(2, 24) = 27.90$, $MSE = 2443$, $p < .001$) and Number Magnitude ($F(11, 132) = 6.46$, $MSE = 2546$, $p < .001$) were significant, but these effects were embedded in an important three way interaction with naming language, $F(22, 264) = 4.21$, $MSE = 441$, $p < .001$. Indeed, as can be seen in Figure 3, the effect of Number Magnitude appears to be present in only some of the Stimulus Format x Naming Language conditions, as expected. These effects of Number Magnitude will be investigated in more detail by means of regression analyses.

Regression Analysis

Following Duyck and Brysbaert (2004); and to assess the effect of number magnitude independent of number frequency[3], regression analyses were performed according to the procedure for repeated measures data described by Lorch and Myers (1990, Method 3), with number magnitude and frequency as predictors (for a detailed statistical explanation of the computational procedure of these tests, see Lorch & Myers, 1990).

The regression weights for the six conditions [i.e. 2 naming language (L1 versus L2) x 3 stimulus formats (Arabic, L1, and L2)] are displayed in Table 2. Most importantly, the regression weight of number magnitude differed significantly from zero in the backward translation condition (English [L2] to Dutch [L1]), $t(12) = 2.88$, $p < .02$. This semantic effect of number magnitude was not significant for forward translation (Dutch [L1] to English [L2]), $t < 1$, $p > .49$. Accordingly, conceptual mediation was significantly larger in backward than in forward translation, $t(12) = 2.92$, $p < .02$. As can be seen in Table 2, the effect of Number magnitude was not significant for L1 and L2 within-language number word naming (all *p*s > .10), as expected.

Insert Table 2 about here

Discussion

The results of Experiment 1 are quite clear. Similar to Duyck and Brysbaert (2004), we obtained a significant semantic effect of number magnitude in backward translation. It took longer to translate L2 number words representing large quantities (e.g. *eight*) than L2 number words representing small quantities (e.g. *two*). These findings strongly suggest conceptual mediation in backward translation and the existence of strong L2 form-to-meaning mappings, since magnitude information is not stored at the lexical-form level. The regression analyses confirmed that this magnitude effect in backward translation was not due to effects of number word frequency. As the frequency effect is usually situated at the lexical level, this is further evidence that the translations were not based on direct wordform associations. As expected, we did not find a semantic effect for within-language number word naming, because, unlike word translation, this is not a semantic task (Fias, 2001). As expected on the basis of the lexical form overlap hypothesis, the semantic Dutch-English translation effects in this study were not so strong as in Dutch-French translation: unlike Duyck and Brysbaert (2004), forward translation did not yield a magnitude effect, suggesting a less semantically mediated translation process for these two similar languages. Further theoretical implications of these findings will be discussed in the *General Discussion.*

Experiment 2

In this second experiment, we will investigate whether the previous findings generalize to German (L3). Because German is also lexically similar to L1, we predict a similar pattern of semantic translation effects.

Method

Participants

Participants were the same Dutch-English-German multilinguals who also participated in Experiment 1. We chose not to manipulate translation language (L2 vs. L3) between subjects to avoid individual difference confounds when comparing L2 and L3 results (see also de Groot & Hoeks, 1995;

Francis & Gallard, 2005). There was a minimum of two weeks between the two experiments. Order of sessions was counterbalanced across participants. Self-reported L3 proficiency was significantly lower than L2 and L1 proficiency, *p*s < .001.

Materials, Design and Procedure

The materials, design and procedure were identical to Duyck and Brysbaert (2004) and to Experiment 1, except that foreign-language number words were now in German (L3). Mean similarity between Dutch and German translation equivalents (*M* = 552.81) was significantly higher (*p* < .01) than for the Dutch-French pairs (*M* = 146.06) used by Duyck and Brysbaert (2004), according to the word similarity measure described by Van Orden (1987). Similarity was not significantly larger for Dutch-German (this Experiment) than for Dutch-English (Experiment 1), *p* > .20.

Results

Variance Analysis

The proportion of invalid trials due to naming errors or faulty time registration was 2.16%. These trials were excluded from all analyses. Following the same criterion as in Experiment 1, outliers were excluded from all analyses (1.8 % of the data). Again, an ANOVA was performed with Naming Language, Stimulus Format, and Number Magnitude as repeated measures factors. Mean RTs are presented in Figure 4.

Insert Figure 4 about here

In contrast with L2 naming in Experiment 1, L3 naming (*M* = 595) was significantly slower than L1 naming (*M* = 499), as indicated by the significant main effect of Naming Language, *F*(1, 12) = 42.52, *MSE* = 50810, *p* < .001. Unlike Experiment 1, planned comparisons showed that backward translation was significantly faster (*M* = 546 ms) than forward translation (*M* = 598 ms), $F(1, 12) = 11.47$, *MSE* = 18544, $p < .01$.

The main effects of Stimulus Format ($F(2, 24) = 11.32$, *MSE* = 2470, $p < .001$) and Number Magnitude ($F(11, 132) = 15.32$, *MSE* = 3809, $p < .001$) were significant, but these effects were again embedded in a three way interaction effect with naming language, $F(22, 264) = 8.31$, *MSE* = 1695, $p < .001$. Indeed, Figure 4 shows that the effect of Number Magnitude is only present in some of the Stimulus Format x Naming Language conditions. Similar to Experiment 1 and to Duyck and Brysbaert (2004), these effects of Number Magnitude will be investigated in more detail by means of regression analyses.

Regression Analysis

Regression analyses were performed according to the procedure for repeated measures data described by Lorch and Myers (1990, Method 3), with number magnitude and frequency[4] as predictors.

The regression weights for the six conditions [i.e. 2 naming language (L1 versus L2) x 3 stimulus formats (Arabic, L1, and L2)] are displayed in Table 4. Similar to Experiment 1, the regression weight of number magnitude differed significantly from zero in backward translation, $t(12) = 6.58$, $p < .001$. Although Figure 4 also shows a small trend for longer RTs with increasing magnitude for forward translation, this effect was not significant, $t < 1$, $p > .60$. Apparently, this small increasing trend was not due to number magnitude, but partly to other factors such as word frequency. As can be seen in Table 3, the effect of Number magnitude was not significant for L1 and L3 within-language number word naming (all *p*s > .18), as expected.

Insert Table 3 about here

It might be interesting to note that Table 3 also shows a strong significant effect of frequency in the backward translation condition. However, the direction of this effect is reversed, with longer RTs for more frequent number words. Figure 4 shows that this reversed frequency effect may be due to very fast

RTs on targets *vier*, *acht* and *elf*, which are low frequent (M = 2.20, or 158 per million words) relative to the other number words (M = 2.60 or 360 per million words). However, these three number words are also the only identical Dutch-German cognates (i.e. have the same spelling) in the stimulus set, which may explain these fast RTs. Indeed, de Groot (1992) has shown that cognates are translated faster and yield less semantic activation than non-cognates. To find out whether the reversed frequency effect obtained in this study is indeed due to a confound with cognate status, we repeated our regression analyses including cognate status as an additional predictor. The results of this analysis are displayed in Table 4.

Insert Table 4 about here

As expected, the effect of cognate status in backward translation was significant, $t(12) = 5.92$, $p < .001$. Identical cognates were translated faster than non-cognates. Also, the reversed frequency effect in Table 3 was indeed an artifact of cognate status. With this predictor included, the reversed frequency effect in backward translation was no longer significant, $t(12) = 1.11$, $p > .28$. Importantly, with cognate status included, the magnitude effect in backward translation was still significant, $t(12) = 6.40$, $p < .001$[5]. Finally, the effect of cognate status in forward translation was also significant, $t(12) = 2.96$, $p < .05$. With this additional predictor, there was still no magnitude effect in forward translation, $t < 1$, $p > .78$. Note that a similar cognate status analysis was not possible for Experiment 1, because there are no Dutch-English identical cognate number words between 1 and 12.

Discussion

The results of Experiment 2 are similar to Experiment 1. Following Duyck and Brysbaert (2004), we obtained a significant semantic effect of number magnitude in backward translation, but not in forward translation. Hence, the observed differences between L3 translation and L2 translation (Experiment 1) do not concern the pattern of semantic mediation, but rather overall naming speed. L3 naming was significantly slower than L1 naming, whereas L2 naming was not, as could be expected from the lower L3 proficiency scores. Consequently, backward translation was now significantly faster, not slower (Experiment 1), than forward translation. The observation that backward translation was faster, but still yielded a number magnitude effect, whereas slower forward translation did not, strongly suggests it is not always appropriate to draw conclusions about semantic access solely from overall translation speed differences. In the RHM for example, forward translation is expected to be slower, because it requires an extra processing step (i.e. semantic access). The present findings suggest that translation speed is not an additive function of processing steps in a dual-route translation model. Further theoretical implications of these findings will be discussed in the *General Discussion.*

General Discussion

According to the RHM of Kroll and Stewart (1994), forward translation is more likely to be conceptually mediated than backward translation, because form-to-meaning mappings are stronger for L1 than for L2. Backward translation may only be semantically mediated for high levels of L2 proficiency. Contrastingly, the model of Duyck and Brysbaert (2004) assumes that L2 form-to-meaning mappings may be strong and develop rapidly for certain types of words, such as number words and colour words (La Heij et al., 1996). In this model, the degree of semantic activation during translation is not a function of translation direction, but the result of parallel activation coming from both lexical-form and semantic representations (see the *Introduction*).

The results obtained in the present study are not in line with the asymmetric lexicosemantic organization proposed by the RHM. We obtained a strong semantic number magnitude effect in backward translation from both L2 and L3. Replicating semantic backward translation effects reported by Duyck and Brysbaert (2004) with Dutch-French bilinguals, this shows that L2 and even L3 number words may strongly and rapidly activate their underlying semantic representation, as assumed in Duyck and Brysbaert's model. Surprisingly, this semantic magnitude effect was not obtained in forward translation (both for L2 and L3), even though the RHM would predict stronger semantic effects in forward than in backward translation.

In general, semantic translation effects were less strong in this study with more lexically similar languages (Dutch-English and Dutch-German) than in the Dutch-French study of Duyck and Brysbaert (2004). This is consistent with the lexical form overlap hypothesis which states that more similar translation equivalents have stronger wordform connections than dissimilar translations. Because translation output in Duyck and Brysbaert's (2004) model is a function of activation coming from both lexical-form and semantic representations; output will be relatively less influenced by semantic activation when two languages are more similar. Looking at specific words, if the form overlap between translation equivalents is at a maximum (the case of identical cognates), translation RTs were significantly faster (Experiment 2, Dutch-German cognates *vier, acht* and *elf*) and did not seem to follow the increase of RTs as a function of number magnitude (see also Footnote 5). This is consistent with earlier findings that cognates are easier to translate and show less evidence for semantic mediation in translation, than words with no form overlap (e.g., de Groot, 1992; de Groot et al., 1994; see Costa et al., 2005, for similar cognate facilitation effects in bilingual word production). It also suggests that translation between two similar languages is not necessarily less semantically mediated, irrespective of the specific words involved. Instead, the influence of cross-lingual lexical-form similarity operates on the word level, not at the language level. Finally, note that the present evidence supporting the lexical overlap assumption is inconsistent with findings reported in the original Kroll and Stewart (1994) study. Using a post-hoc cognate analysis, they found that cognate status (maximal cross-lingual lexical-form overlap) did not interact with the obtained semantic effects in translation.

The present set of findings suggests that lexicosemantic organization for certain types of words may be depending to a larger extent on form overlap than on general L2 (or L3) proficiency. First, similar semantic translation patterns were observed for English (L2) and German (L3), two languages that have equally similar number words relative to Dutch ($p > .20$, see earlier, Van Orden, 1987), even though L2 and L3 proficiency differed significantly from each other (as indicated by different proficiency scores and the fact that L3 naming, but not L2 naming, was slower than L1 naming). Similarly, Duyck and Brysbaert (2004) found similar symmetric semantic translation patterns for balanced and unbalanced Dutch-French bilinguals, who also differed in L2 proficiency. Secondly, this study with lexically more similar language pairs yielded different semantic translation patterns than Duyck and Brysbaert (2004), who tested two lexically dissimilar languages (Dutch-French), and did obtain symmetric semantic mediation. This difference emerged even though L2 proficiency was probably not lower in the present study than in Duyck and Brysbaert (2004). Using exactly the same procedure, apparatus and materials, mean L2 (English) naming speed in this study was 494 ms, which is faster than the mean L2 (French) naming speed of the balanced (M = 510 ms) and unbalanced (M = 546) bilinguals in Duyck and Brysbaert (2004). So, this dissociation between proficiency and lexical similarity manipulations suggests that the crucial determinant for lexicosemantic organization and resulting translation patterns may be cross-lingual lexical similarity between translation equivalents, rather than L2 proficiency. Of course, further research is needed to generalize the present observations for number words to other types of words. At present, our conclusion may therefore only be that any future model of bilingual lexicosemantic organization will have to include a possible influence of cross-lingual lexical form overlap, to account for the findings above.

Even though cross-lingual lexical similarity may be more influential for bilingual lexicosemantic organization than L2/L3 proficiency, one could argue that the obtained semantic number magnitude effects in backward translation are still consistent with the developmental assumption of the RHM, if one assumes high L2/L3 proficiency specifically for number words. Indeed, even though general L3 proficiency was quite low in the present study, participants were still very fast in processing L3 number words. Duyck and Brysbaert (2004) already argued that such an account, implying word-level influences on lexicosemantic organization, would indeed constitute a useful extension of the traditional RHM. In this view, it may also be interesting to note that Duyck and Brysbaert (2004) already suggested that strong form-to-meaning mappings for number words may also be a consequence of their early age of acquisition. Indeed, it is plausible to assume that L2/L3 words that are acquired earlier, such as number words, develop stronger form-to-meaning mappings than later acquired L2/L3 words.

The second aim of this study was to investigate semantic mediation in L1-L3 translation. To our knowledge, there have only been two studies that have looked at L3 translation (but see Goral, Levy, Obler, & Cohen, 2006, for a trilingual aphasia study). First, De Groot and Hoeks (1995)investigated translation with Dutch-English-French trilinguals. However, they only looked at forward translation, so a comparison between their results and ours may not be complete. Secondly, Francis and Gallard (2005) looked at L1-L3 translation by English-Spanish-French trilinguals. Using a repetition priming paradigm, they found that both backward and forward L1-L3 translation was semantically mediated. The absence of translation through wordform associations in the study of Francis and Gallard is consistent with the lexical overlap hypothesis, as English and French are two languages from a different origin (Germanic vs. Roman) and are therefore quite dissimilar, just as Dutch and French, which also yielded symmetric conceptual mediation in Duyck and Brysbaert (2004).

To conclude this discussion, it may be worthwhile to speculate a little bit on the reason why it was exactly *forward* translation that did not yield semantic mediation, and not backward translation. The general pattern of results is consistent with the cross-lingual form overlap hypothesis. Semantic translation effects were weaker with two more similar language pairs (Dutch-English and Dutch-German), than in the study of Duyck and Brysbaert (2004), using less lexically similar languages (Dutch-French). However, even if one would allow word-level cross-lingual form overlap effects within the architecture of the RHM, one would expect that semantic effects would especially be weaker for backward translation, which was clearly not the case. In our model, translation performance is the outcome of relative activation coming from both lexical-form and semantic representations. So, theoretically it does not exclude the possibility that forward translation yields less semantic activation than backward translation. Of course, this general principle does not provide a detailed account of the specific pattern of results obtained, and only an implemented version of our connectionist architecture may reveal the patterns of activation expected for specific stimuli. At least, the data show that the assumption for ubiquitous greater semantic activation in forward translation is wrong. As such, these data are important for any future model of bilingual lexicosemantic organization. However, future research is needed to provide a definite answer to this issue.

To summarize, using the number word translation paradigm with lexically similar languages, we have obtained evidence for semantic access in backward translation, but not in forward translation. In combination with earlier symmetric semantic translation effects from Duyck and Brysbaert (2004), obtained with lexically more dissimilar languages, this suggests that lexicosemantic organization may be influenced by lexical form overlap between translation equivalents.

References

Altarriba, J., & Mathis, K. M. (1997). Conceptual and Lexical Development in Second Language Acquisition. *Journal of Memory and Language, 36*(4), 550-568.

Baayen, R., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database.* [CD-ROM]. Philadelphia: University of Pennsylvania: Linguistic Data Consortium.

Bloem, I., & La Heij, W. (2003). Semantic Facilitation and Semantic Interference in Word Translation: Implications for Models of Lexical Access in Language Production. *Journal of Memory and Language, 48*(3), 468-488.

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, 26*(5), 1283-1296.

Costa, A., Santesteban, M., & Cano, A. (2005). On the Facilitatory Effects of Cognate Words in Bilingual Speech Production. *Brain and Language, 94*(1), 94-103.

de Groot, A. M. B. (1992). Determinants of Word Translation. *Journal of Experimental Psychology: Learning Memory and Cognition, 18*(5), 1001-1018.

de Groot, A. M. B., & Comijs, H. (1995). Translation Recognition and Translation Production - Comparing a New and an Old Tool in the Study of Bilingualism. *Language Learning, 45*(3), 467-509.

de Groot, A. M. B., Dannenburg, L., & van Hell, J. G. (1994). Forward and Backward Word Translation by Bilinguals. *Journal of Memory and Language, 33*(5), 600-629.

de Groot, A. M. B., & Hoeks, J. C. J. (1995). The Development of Bilingual Memory: Evidence From Word Translation by Trilinguals. *Language Learning, 45*(4), 683-724.

de Groot, A. M. B., & Poot, R. (1997). Word Translation at Three Levels of Proficiency in a Second Language: the Ubiquitous Involvement of Conceptual Memory. *Language Learning, 47*(2), 215-264.

Duyck, W., & Brysbaert, M. (2002). What Number Translation Studies Can Teach Us About the Lexico-Semantic Organisation in Bilinguals. *Psychologica Belgica, 42*(3), 151-175.

Duyck, W., & Brysbaert, M. (2004). Forward and Backward Number Translation Requires Conceptual Mediation in Both Balanced and Unbalanced Bilinguals. *Journal of Experimental Psychology: Human Perception and Performance, 30*(5), 889-906.

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.

Duyck, W., Lagrou, E., Gevers, W., & Fias, W. (in press). Roman Digit Naming: Evidence for a Semantic Route in Translation. *Experimental Psychology.*

Fias, W. (2001). Two Routes for the Processing of Verbal Numbers: Evidence from the SNARC Effect. *Psychological Research,*

Francis, W. S., & Gallard, S. L. K. (2005). Concept Mediation in Trilingual Translation: Evidence From Response Time and Repetition Priming Patterns. *Psychonomic Bulletin & Review, 12*(6), 1082-1088.

Gielen, I., Brysbaert, M., & Dhondt, A. (1991). The Syllable-Length Effect in Number Processing Is Task- Dependent. *Perception & Psychophysics, 50*(5), 449-458.

Gollan, T. H., & Acenas, L. A. R. (2004). What Is a Tot? Cognate and Translation Effects on Tip-of-the-Tongue States in Spanish-English and Tagalog-English Bilinguals. *Journal of Experimental Psychology: Learning Memory and Cognition, 30*(1), 246-269.

Goral, M., Levy, E. S., Obler, L. K., & Cohen, E. (2006). Cross-Language Lexical Connections in the Mental Lexicon: Evidence From a Case of Trilingual Aphasia. *Brain and Language, 98*(2), 235-247.

Kohnert, K. (2004). Cognitive and Cognate-Based Treatments for Bilingual Aphasia: a Case Study. *Brain and Language, 91*(3), 294-302.

Kroll, J. F., & de Groot, A. M. B. (1997). Lexical And Conceptual Memory in the Bilingual: Mapping Form to Meaning in Two Languages. In de Groot, A. M. B. and Kroll, J. F. (Eds.), *Tutorials In Bilingualism: Psycholinguistic Perspectives.* (pp. 201-224). Mahwah, NJ: Lawrence Erlbaum Publishers.

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.

Kroll, J. F., & Tokowicz, N. (2005). Models of Bilingual Representation and Processing: Looking Back and to the Future. In Kroll, J. F. and de Groot, A. M. B. (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches.* (pp. 531-553). New York: Oxford University Press.

La Heij, W., Hooglander, A., Kerling, R., & Van der Velden, E. (1996). Nonverbal Context Effects in Forward and Backward Word Translation: Evidence for Concept Mediation. *Journal of Memory and Language, 35*(5), 648-665.

Logan, G. D., & Klapp, S. T. (1991). Automatizing Alphabet Arithmetic .1. Is Extended Practice Necessary to Produce Automaticity. *Journal of Experimental Psychology: Learning Memory and Cognition, 17*(2), 179-195.

Lorch, R. F., & Myers, J. L. (1990). Regression-Analyses of Repeated Measures Data in Cognitive Research. *Journal of Experimental Psychology: Learning Memory and Cognition, 16*(1), 149-157.

Sholl, A., Sankaranarayanan, A., & Kroll, J. F. (1995). Transfer Between Picture Naming and Translation - a Test of Asymmetries in Bilingual Memory. *Psychological Science, 6*(1), 45-49.

Talamas, A., Kroll, J. F., & Dufour R. (1999). From Form to Meaning: Stages in the Acquisition of Second-Language Vocabulary. *Bilingualism: Language and Cognition, 2*(1), 45-58.

Tzelgov, J., Yehene, V., Kotler, L., & Alon, A. (2000). Automatic Comparisons of Artificial Digits Never Compared: Learning Linear Ordering Relations. *Journal of Experimental Psychology: Learning Memory and Cognition, 26*(1), 103-120.

van Hell, J. G., & de Groot, A. M. B. (1998). Conceptual Representation in Bilingual Memory: Effects of Concreteness and Cognate Status in Word Association. *Bilingualism: Language and Cognition, 1*, 193-211.

Van Orden, G. C. (1987). A Rows Is a Rose - Spelling, Sound, and Reading. *Memory & Cognition, 15*(3), 181-198.

Author Note

Wouter Duyck, Department of Experimental Psychology, Ghent University, Ghent, Belgium. Marc Brysbaert, Royal Holloway University of London, UK.

This research was made possible by the Research Foundation Flanders, of which the first author is a post-doctoral fellow.

Correspondence concerning this article should be addressed to Wouter Duyck, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 B-Ghent, Belgium. E-mail should be sent to wouter.duyck@UGent.be.

Footnotes

1. Van Orden's (1987) word similarity measure is defined as 10[(50F + 30V + 10C)/A] + 5T + 27B + 18E) with F = number of pairs of adjacent letters in the same order, shared by pairs; V = number of pairs of adjacent letters in reverse order, shared by pairs; C = number of single letters shared by word pairs; A = average number of letters in the two pairs; T = ration of shorter word to longer word; B = 1 if first two letters are the same, else B = 0; E = 1 if last two letters are the same, else E = 0

2. Duyck and Brysbaert (2004) showed that the semantic number magnitude effects in both translation directions also emerge if both stimulus format and naming language are blocked.

3. Gielen, Brysbaert, and Dhondt (1991) reported a significant correlation between number magnitude and number frequency ($r = -.621$, $p < .01$). So, because smaller numbers are more frequent, it is possible that any effect of number magnitude in the data is a confounded effect of number frequency. Therefore, we included the digit frequency measures as reported by Gielen et al. (1991) in our regression analyses. For the English and German number words, we included the log CELEX frequencies per million words (Baayen, Piepenbrock, & Van Rijn, 1993), obtained by the WordGen stimulus selection program (Duyck, Desmet, Verbeke, & Brysbaert, 2004). Correlations between Dutch-English, Dutch-German and English-German frequency measures for the included number words were respectively .69, .99 and .70.

4. Similar to Experiment 1 (English), German word frequency was extracted from the CELEX lexical database (Baayen et al., 1993), using the WordGen program (Duyck et al., 2004).

5. Note that separate regression analyses for backward translation of non-cognate and cognate number words yielded a number magnitude regression weight that differed significantly from zero for non-cognate number words ($B = 9.77$, $t(13) = 5.99$, $p < .001$), but not for cognate words (a negative weight of $B = -2.15$, *ns*). These analyses should be interpreted with caution though, because there were too few cognates (three: *vier*, *acht*, *elf*) to include frequency as an additional control predictor in these separate Lorch and Myers (1990) regression analyses.

Table 1. Mean self-reported proficiency for L1 to L4 (Experiments 1 and 2). Standard deviations are indicated between brackets.

	L1 (Dutch)	L2 (English)	L3 (German)	L4 (French)
Experiment 1	6.77 (0.44)	5.31 (0.95)	4.23 (1.09)	3.15 (1.07)

Table 2. The regression equations for the six naming language x stimulus format conditions (Experiment 1) according to the procedure described by Lorch and Myers (1990) (* $p < .05$). BT = backward translation, FT = forward translation

Naming Language	Stimulus Format		Intercept	Number Magnitude	Frequency
L1 Naming (Dutch)	Arabic Numbers	*Y* =	546	- 1.07 NM	- 0.12 F*
	L1 Number Words (Dutch)	*Y* =	468	+ 1.32 NM	+ 4.60 F
	L2 Number Words (English) [BT]	*Y* =	515	+ 4.39 NM*	+ 1.26 F
L2 Naming (English)	Arabic Numbers	*Y* =	572	- 2.24 NM	- 0.16 F*
	L1 Number Words (Dutch) [FT]	*Y* =	500	- 1.22 NM	+ 1.97 F
	L2 Number Words (English)	*Y* =	486	+ 0.69 NM	+ 7.3417 F

Table 3. The regression equations for the six naming language x stimulus format conditions (Experiment 2) according to the procedure described by Lorch and Myers (1990) (*** p < .001). BT = backward translation, FT = forward translation

Naming Language	Stimulus Format		Intercept	Number Magnitude	Frequency
L1 Naming (Dutch)	Arabic Numbers	*Y* =	462	+ 0.21 NM	+ 0.02 F
	L1 Number Words (Dutch)	*Y* =	446	+ 2.18 NM	+ 7.49 F
	L2 Number Words (German) [BT]	*Y* =	425	+ 9.23 NM***	+ 24.79 F***
L3 Naming (German)	Arabic Numbers	*Y* =	832	- 2.05 NM	- 0.52 F***
	L1 Number Words (Dutch) [FT]	*Y* =	619	+ 1.03 NM	- 11.63 F
	L3 Number Words (German)	*Y* =	571	+ 1.64 NM	- 5.99 F

Table 4. The regression equations (Lorch & Myers, 1990) for the two translation conditions with cognate status as a predictor (Experiment 2) (* $p < .05$; *** $p < .001$).

		Intercept	Number Magnitude	Frequency	Cognate Status
Backward Translation	*Y* =	499	+ 8.89 NM***	+ 5.89 F	- 100.81 C***
Forward Translation	*Y* =	643	+ 0.59 NM	- 18.05 F*	- 26.27 C**

Figure Captions

Figure 1. The revised hierarchical model of bilingual memory (as published in Kroll & De Groot, 1997). Solid lines represent stronger links than dotted lines.

Figure 2. Duyck and Brysbaert's (2004) model of bilingual lexicosemantic organization with varying semantic overlap and differently weighted lexico-semantic and intralexical connections. Solid lines represent stronger links than dotted lines. Depicted words and semantic representations are illustrative examples for Dutch-English bilinguals.

Figure 3. Mean naming RTs of Experiment 1 (Dutch-English) by Naming Language (L1 vs. L2), Stimulus Format (Arabic digits, L1 number words, L2 number words) and Number Magnitude (1 to 12). Straight lines represent best linear fit according to a least squares criterion.

Figure 4. Mean naming RTs of Experiment 2 (Dutch-German) by Naming Language (L1 vs. L2), Stimulus Format (Arabic digits, L1 number words, L2 number words) and Number Magnitude (1 to 12). Straight lines represent best linear fit according to a least squares criterion.

Figure 1

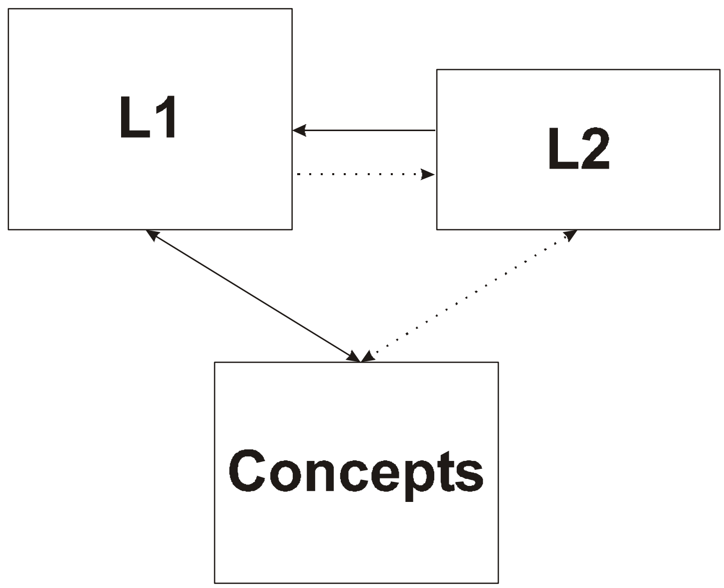


Figure 2

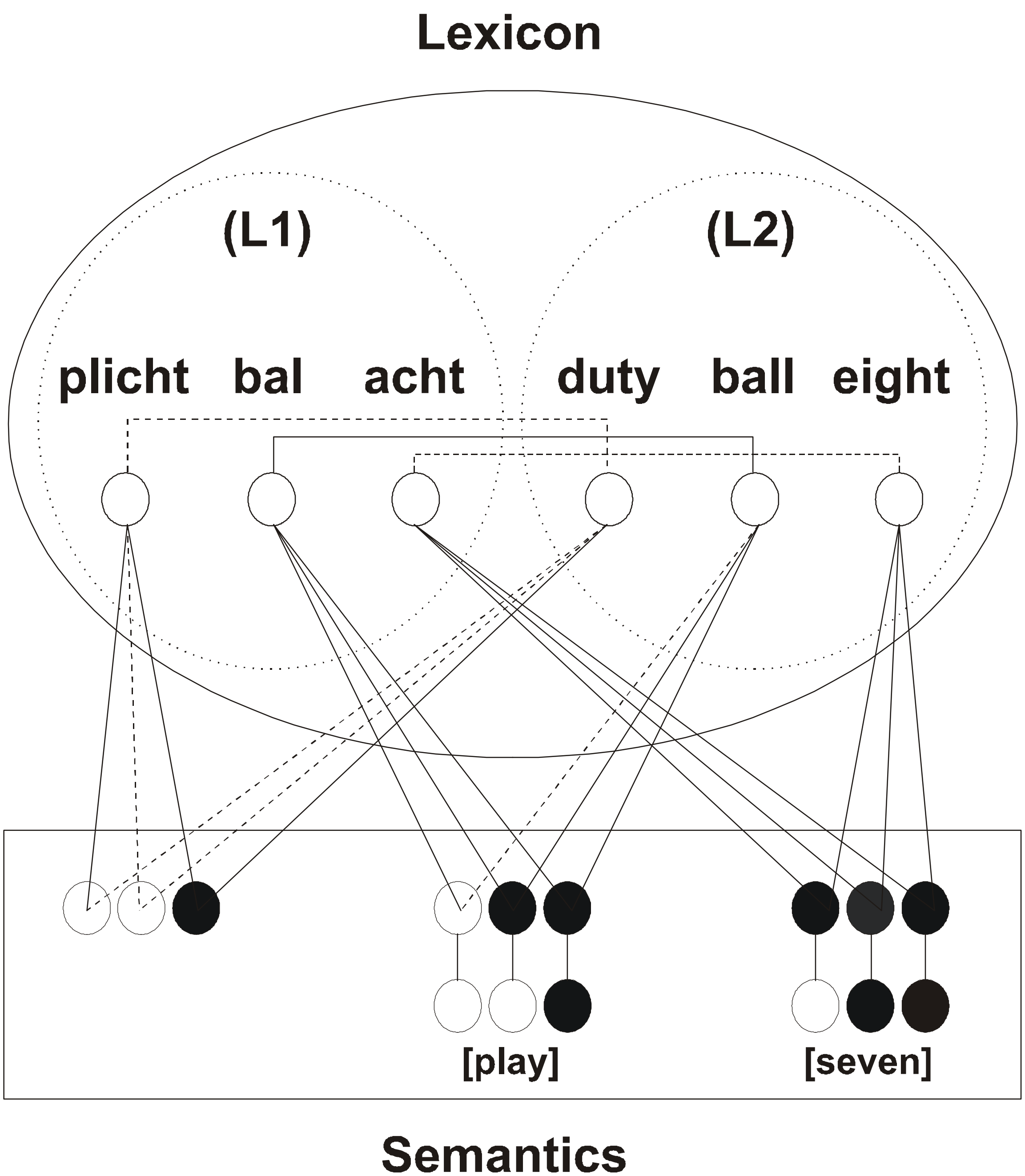

Lexicon
(L1)
(L2)
plicht
bal
acht
duty
ball
eight
[play]
[seven]
Semantics

Figure 3

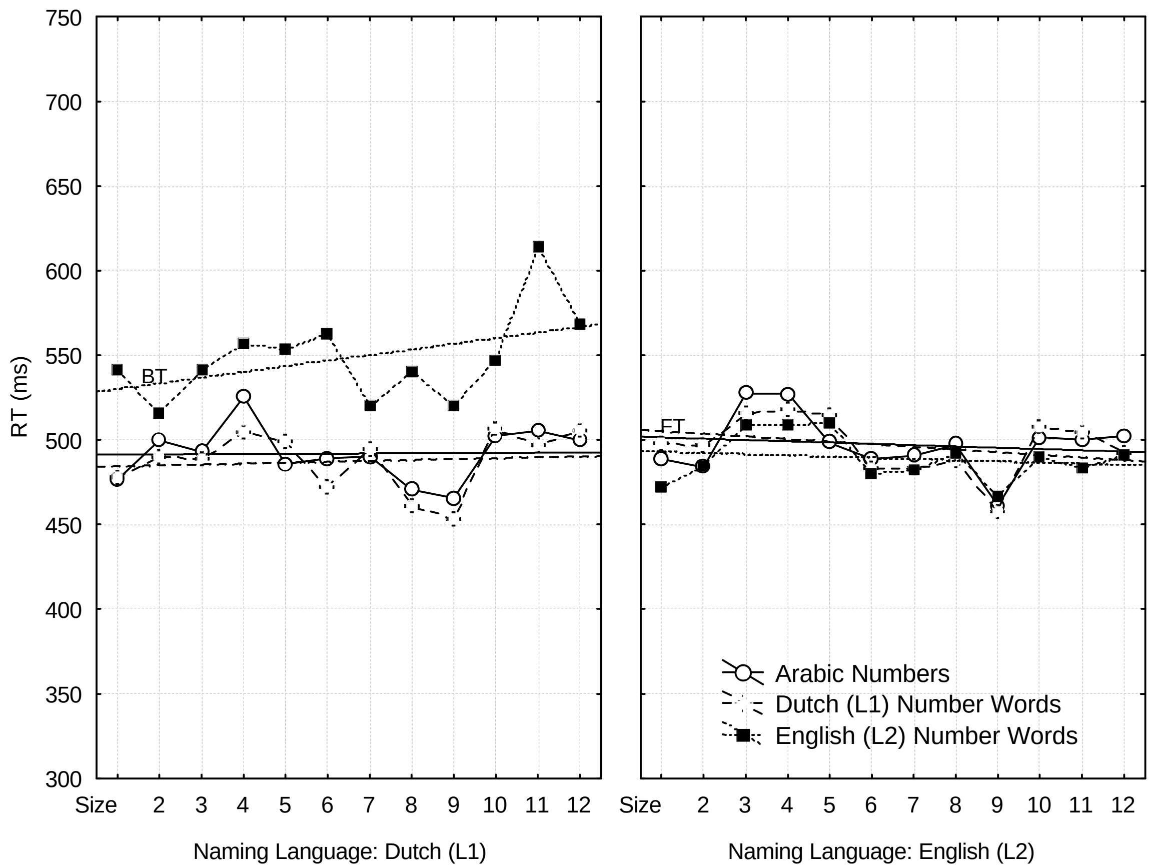
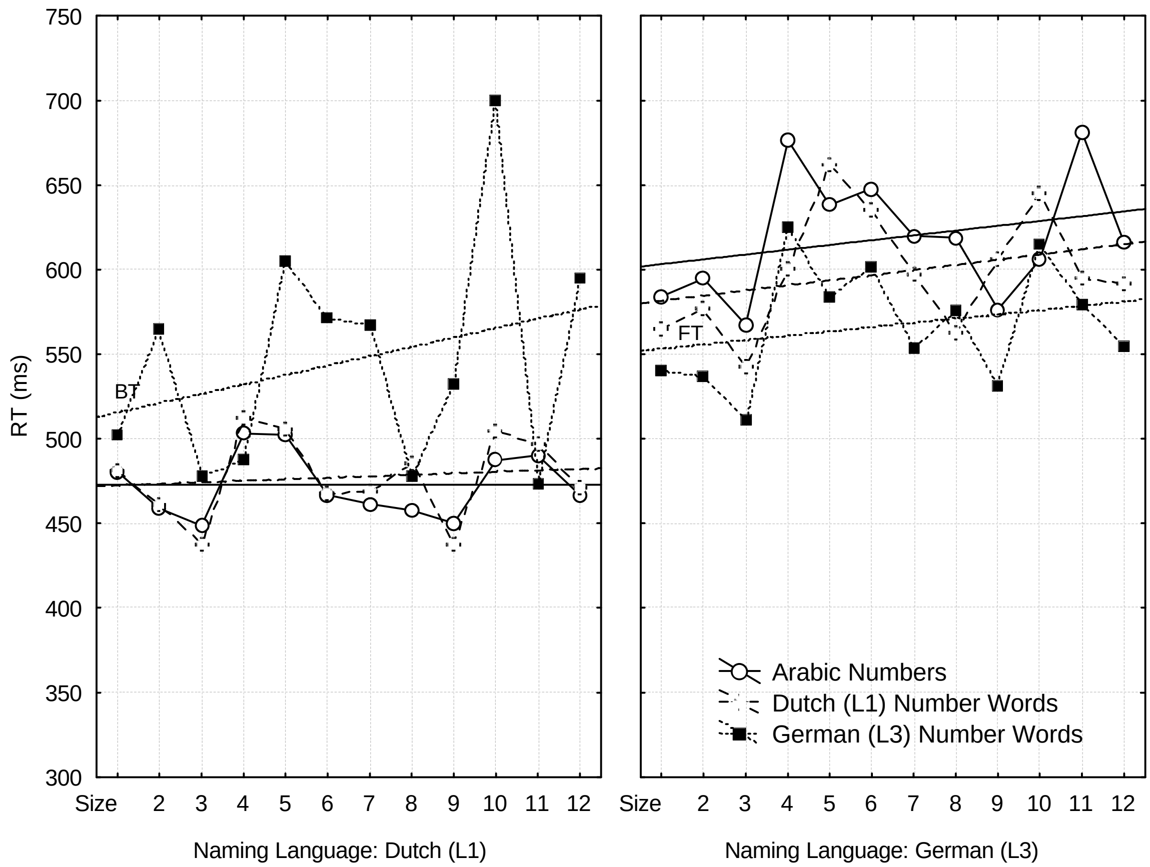

750
700
650
600
550
500
450
400
350
300
RT (ms)
BT
FT
Size 2 3 4 5 6 7 8 9 10 11 12
Naming Language: Dutch (L1)
Naming Language: English (L2)
Arabic Numbers
Dutch (L1) Number Words
English (L2) Number Words

Figure 4



750
700
650
600
550
500
450
400
350
300
RT (ms)
BT
FT
Size
2
3
4
5
6
7
8
9
10
11
12
Naming Language: Dutch (L1)
Naming Language: German (L3)
Arabic Numbers
Dutch (L1) Number Words
German (L3) Number Words