Language Learning 47:2, June 1997, pp. 215–264

Word Translation at Three Levels of Proficiency in a Second Language: The Ubiquitous Involvement of Conceptual Memory

Annette M. B. de Groot and Rik Poot
University of Amsterdam

Three groups of 20 unbalanced bilinguals, different from one another in second language (L2) fluency, translated one set of words from L1, Dutch, to L2, English (forward translation), and a second set of matched words from L2 to L1 (backward translation). In both language sets we orthogonally manipulated 3 word characteristics: word imageability, word frequency, and cognate status. Similar data patterns were obtained for the 3 proficiency groups, suggesting that word translation was qualitatively the same in all 3. Especially, word imageability affected performance similarly across the 3 proficiency levels. This suggests that our 3 groups did not differ in the degree to which they involved conceptual memory in the translation process. Further, contrary to the "asymmetry" model of word translation, conceptual memory appeared to operate as much in backward translation as in forward translation. The combined data indicate that, at least beyond the

---

We owe much gratitude to the director, ESL teachers, and students of the "Willem Blæu College" in Alkmaar, The Netherlands, for their enthusiastic contribution to this study. We thank Judith Kroll for insightful comments on an earlier version of this article, and Janet van Hell for helpful discussions regarding data collection and analysis.

Correspondence concerning this article may be addressed to A. M. B. de Groot, University of Amsterdam, Department of Psychology, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands.

Internet: pn_groot@macmail.psy.uva.nl

initial stages of L2 fluency, "concept mediation" is a universal process in translating words between 2 languages. We can account for the few, generally small differences between the fluency groups in terms of the idea that some subsets of words lag in L2 knowledge development.

Recently, word translation has become a popular task in studies of bilingualism. Many studies that used this task have had two purposes. First, they aimed to reveal the variables that play a role in word-translation performance. This purpose has been most prominent in a number of studies that looked at the role of *word type* in word translation by varying a large number of word characteristics (De Groot, 1992a; De Groot & Comijs, 1995; De Groot, Dannenburg, & Van Hell, 1994; Murray, 1986; Sánchez-Casas, Davis, & García-Albea, 1992; Sánchez-Casas, Suárez-Buratti, & Igoa, 1992). Other studies have investigated the role of three other variables in word translation, namely, *L2 proficiency* (Abunuwara, 1992; Chen & Leung, 1989; De Groot & Hoeks, 1995; Kroll & Curley, 1988; Potter, So, Von Eckardt, & Feldman, 1984), *translation direction* (De Groot et al., 1994; Kroll & Stewart, 1990, 1994; La Heij, Hooglander, Kerling, & Van der Velden, 1996; Sholl, Sankaranarayanan, & Kroll, 1995; Van Hell & De Groot, 1995), and *learning strategy* (Chen, 1990; Chen & Leung, 1989). These latter three groups of studies, however, have mostly focused on the second goal of word-translation studies: to reveal underlying processing in word translation and the bilingual memory structures in which this processing takes place. The two goals are, in fact, closely intertwined: When research has identified a particular variable as a determinant of word translation, the researchers typically seek a cause for the associated effects. Usually, they have designated variation in bilingual memory organization across different types of words and bilinguals as the source of these effects. In turn, they assume this variation affects the underlying translation processes (see below for clarification).

The "standard" model of bilingual memory explicitly distinguishes between two levels of representation that together contain three memory stores. One level ("form memory") contains the representations of word forms, the second ("conceptual memory") those of word meanings. Studies typically assume word-form memory comprises two sets of elements, one for each of the bilingual's two languages; in other words, word-form memory is language-specific. In contrast, the standard model assumes the elements in conceptual memory are language-independent; that is, they are shared between the two languages (but see, e.g., De Groot, 1992b, 1993; Kirsner, Smith, Lockhart, King, & Jain, 1984; Kroll & De Groot, 1997, for qualifications).

The main conclusion of the studies that manipulated *word type* is that three word dimensions play a crucial role in translation performance: meaning, frequency of use/familiarity, and whether or not the two words in a translation pair share a "cognate" relation. (Cognates are translation-equivalent words similar in form—orthography and/or phonology; noncognates are form-dissimilar translation equivalents.) Concrete words (that, in addition to being concrete, are usually also easy to imagine, to define, and to contextualize; see below), frequently-used words, and words with a cognate translation in the target language are easier to translate (faster, with fewer omissions and errors) than abstract words, uncommon words, and noncognates, respectively. These effects occur both in "translation production" and in "translation recognition." In the former, the participants, presented with words in the "source" language (the language from which they translate), retrieve these words' translations from memory (e.g., De Groot et al., 1994); in the latter, they decide whether pairs of words, each pair consisting of one word in one of their languages and a second word in their other language, are translation equivalents (e.g., De Groot & Comijs, 1995).

One way to explain the effects of concreteness and cognate status is to qualify the standard model's assumption (see above) that conceptual memory is language-independent: Thus, conceptual memory may contain both language-independent and language-specific representations. Alternatively, in terms of the "distributed" view that the units in conceptual memory do not each individually represent a complete concept but more elementary meaning elements instead (e.g., Masson, 1991), one could assume that larger or smaller sets of these elementary units are shared between translation pairs. In terms of these views, either cognates and concrete words, as compared to noncognates and abstract words, are more often stored language-independently in conceptual memory, or share relatively many elementary conceptual units between languages (see, e.g., De Groot, 1992b, 1993, and Van Hell & De Groot, 1996, for more detail). One can explain the frequency effect by postulating stronger connections between the various memory stores for high-frequency words than for low-frequency words (see, e.g., De Groot, 1992a, for more detail).

As to the effect of *L2 proficiency*, some (Chen & Leung, 1989; De Groot & Hoeks, 1995; Kroll & Curley, 1988) but not all (e.g., Potter et al., 1984) of the studies that manipulated this variable have indicated that the translation processes of bilinguals with different levels of L2 proficiency differ qualitatively, suggesting different types of underlying memory structures at different levels of L2 proficiency. More specifically, most studies have assumed that the memory structures in bilinguals of different fluency levels differ in the types and strength of the connections between the various memory stores (and, incidentally, in the size of the L2 word-form store): Bilinguals with high L2 proficiency have relatively strong connections between representations in L2 word-form memory and the corresponding representations in conceptual memory as compared to the connections between corresponding L1 and L2 representations in the two word-form stores. (Some researchers have gone further, proposing that the latter type of connections does not even exist in relatively fluent bilinguals, e.g., Potter et al., 1984.) Bilinguals with a low level of L2 proficiency show the opposite pattern. As a consequence of these differences in the strength of the various types of connections, the researchers assume that word translation proceeds along different routes in bilinguals of different fluency levels. The stronger a connection, the more probably it plays a role in the translation process. Word translation in bilinguals with higher levels of L2 proficiency will thus often involve conceptual memory, employing the strong connection between the form representation of the L2 word and its conceptual representation (shared with its translation equivalent in L1). This process is usually called translation through "concept mediation." In contrast, less fluent bilinguals more often use the direct connection between the corresponding L1 and L2 word-form representations in word translation, a process generally called "word-association" translation. These less fluent bilinguals cannot use the indirect translation route (through conceptual memory) more frequently because, in translating from L1, they will often get stuck upon hitting the weak connection from conceptual memory to the L2 word-form store; this same weak connection often will prevent them from even *accessing* conceptual memory when L2 is the source language.

Concentrating on *translation direction*, some studies have suggested that "forward" translation, that is, from the L1 (stronger) language into the (weaker) L2, qualitatively differs from "backward" translation, in the reverse direction (Kroll & Stewart, 1990, 1994; Sholl et al., 1995). These studies have pointed at two effects to substantiate this view: faster backward than forward translation, and larger effects of semantic manipulations in forward than in backward translation. The semantic manipulation in Kroll and Stewart's study, mixed or blocked presentation of words from a limited set of semantic categories, affected forward but not backward translation. The study by Sholl et al. had a picture-naming task, requiring semantic processing, that preceded the translation task. The picture-naming affected subsequent translation performance on words that had been named before, but only in forward translation. Kroll and her colleagues developed the "asymmetry" model (or "revised hierarchical model") to account for these effects. It attributes them, again, to differences in the strengths of the within- and between-level connections between memory units in word-form memory and those in conceptual memory. The model incorporates strong direct connections from units in L2 word-form memory to the corresponding units in L1 word-form memory, and weak connections between these two memory stores in the reverse direction. Furthermore, it assumes that the connections between the units in L1 word-form memory and the corresponding representations in conceptual memory are strong. The differences in strength of the various connections have the effect that forward translation predominantly passes through conceptual memory whereas backward translation primarily uses the direct connections between the two word-form stores (e.g., Kroll & Stewart, 1994, for more detail). However, other studies (De Groot et al., 1994; Van Hell & De Groot, 1995) have found only weak support for the asymmetry model; yet another study (La Heij et al., 1996) obtained absolutely no support for asymmetrical translation.

Finally, the studies that looked at the role of *learning strategy/environment* (Chen, 1990; Chen & Leung, 1989) have suggested that during the initial stages of L2 vocabulary learning the underlying memory structures and processing operations vary with the nature of the learning operations (whether the target L2 word is presented with the corresponding L1 word or with a picture depicting the word's referent). In the later stages of learning, however, the underlying structures and the processes they support apparently become independent of the initial learning strategy. (See De Groot, 1995, for a more detailed discussion of the relation between each of these four variables and bilingual memory organization.)

Our present translation-production study is the first we know of to manipulate 3 of the above 4 variables simultaneously: word type, L2 proficiency, and translation direction. This may disclose possible interactions between these 3 variables. The fourth variable, learning environment, we controlled for: The 3 groups of participants, differing in L2 proficiency, share the same background of L2 learning. We manipulated word type on each of the above-mentioned 3 word characteristics known to affect word-translation performance: meaning, frequency/familiarity, and cognate status. We chose word imageability as a representative meaning variable; it reflects the extent to which a word's referent evokes a mental image. It highly correlates with 3 other variables that also concern meaning aspects of words: word concreteness, context availability, and definition accuracy (e.g., De Groot et al., 1994; Schwanenflugel, Harnishfeger, & Stowe, 1988).

Of the 3 word-type variables we manipulated, word imageability is particularly important, because it demonstrates the involvement of conceptual memory in the translation process. Word imageability concerns an aspect of a word's meaning, and word meaning is stored in conceptual memory; hence, an effect of word imageability indicates that conceptual memory somehow plays a role in the translation process. (See Kroll & Stewart, 1994, La Heij et al., 1996, and Sánchez-Casas, Suárez-Buratti, & Igoa, 1992, for other, but conceptually similar, ways to ascertain that concept-mediation has taken place.) The size of the effect demonstrates the degree to which conceptual memory is involved (a null-effect suggesting that conceptual memory is bypassed). The view that word translation (and the underlying memory structures) differs between bilinguals of higher and lower levels of proficiency in L2, and in the way suggested above, would receive support if imageability's effect was larger the higher the L2 proficiency. In contrast, an observed imageability effect that turned out to be immune to the proficiency manipulation would suggest either of two possibilities: (a) The view that bilinguals with different levels of L2 fluency have different memory structures, and consequently differ qualitatively in word translation, is incorrect; (b) the view expressed under (a) is basically correct but requires the qualification that the underlying memory structures and the translation processes they support differ only between beginning L2 learners on the one hand and those beyond the initial stages of L2 proficiency on the other. The participants in all 3 of the fluency groups would then be past the critical initial stage of learning. In other words, the data would fit the view that, rather than increasing gradually, conceptual involvement during word translation is basically an all-or-none process.

The imageability manipulation also allows us to test the idea (see above) that forward translation primarily operates through conceptual memory, whereas backward translation more often uses the direct links between the form representations of a word and of its translation in the bilingual's other language. If this idea is correct, the imageability effect should be larger in forward than in backward translation. The combined effects of both L2 proficiency (larger imageability effects with higher proficiency levels) and translation direction (larger imageability effects in forward than in backward translation) should mean that especially forward translation will show significant differences between proficiency groups.

In a number of earlier studies we manipulated word type and discussed its effects on translation production (De Groot, 1992a; De Groot et al., 1994) and translation recognition (De Groot & Comijs, 1995) at length. Our present analysis therefore primarily focuses on the remaining 2 variables, L2 proficiency and translation direction. Regarding translation direction, we particularly look at whether and, if so, *how* the imageability effect varies across the 2 levels of this variable. We have also manipulated this variable before (De Groot et al., 1994; De Groot & Comijs, 1995), but between participants, instead of within participants as here, and in only one proficiency group. Within-participant manipulation of translation direction should be more sensitive than between-participants manipulation. Regarding the L2 proficiency variable the question is whether any of the other variables differently affects the different proficiency groups; in other words: Does translation performance differ qualitatively between bilinguals of various L2 proficiency levels? Finally, most of our earlier translation studies manipulated the word-type variables in a correlational design; the present study orthogonally manipulates *all* variables. The orthogonal design lends itself much better to detecting interactions between the variables.

Method

Participants

There were 60 participants. Forty were pupils of a secondary school in Alkmaar, The Netherlands; 20 of these were from Grade 3 (typically 15 years old), the remaining 20 from Grade 5 (typically 17 years old). The third group of 20 participants consisted of first-year psychology students from the University of Amsterdam. All 60 participants had Dutch as their L1. At the time of testing, the Grade 3 pupils had been school-trained in English for about 2.5 years, 3 hours per week. Grade 5 pupils had received about 4.5 years of school training in English, again 3 hours a week. Finally, the university students had been school-trained in English for about 6 years, again 3 hours a week. Furthermore, their university training, which had lasted at least 5 months at the time of testing, had involved predominantly English text books. All participants had been exposed to English outside school in informal settings (watching television, listening to music, etc.) from childhood. Finally, during their last 2 years at primary school (at age 11–12) they had received some preparatory training in English.

Prior to testing, we asked each participant to assess his or her comprehension and production skills in English on a 7-point scale. A 7 had to be ticked if they considered the skill equally well developed as their corresponding skill in Dutch; a 1 if they thought the skill totally lacking. A 3 (proficiency) by 2 (type of skill: comprehension vs. production) by 20 (participants) ANOVA on the obtained scores showed a main effect of proficiency, $F(2,57) = 5.45$, $p < .01$. Collapsed across the comprehension and production scores, the university students' score (5.26) was higher than the Grade 5 pupils' (4.70), which in turn was higher than the Grade 3 pupils' (4.63). However, a Newman-Keuls test indicated that the difference between the scores of the Grade 5 and Grade 3 pupils was not statistically reliable. Collapsed across the 3 proficiency levels, the score for comprehension was higher than that for production: 5.22 and 4.51, respectively; $F\ (1,57) = 91.71$, $p < .0001$.

The interaction between proficiency level and type of skill was not significant ($F < 1$), suggesting that the difference between comprehension and production competence was the same for all 3 groups.

Materials

One set of 144 Dutch words and a second, matched set of 144 English words served as stimulus materials. All stimulus words were nouns; those in the Dutch set were all different from those in the English set, both in form and in meaning (i.e., the words in the English set were not translations of the Dutch words). All 288 stimulus words were selected from a corpus of 440 Dutch-English translation pairs for which we had earlier collected scores on a number of word characteristics (De Groot et al., 1994). The available scores included those on imageability and log frequency of both the Dutch words and their translations in English, and a score reflecting the cognate status of all 440 translation pairs. We orthogonally varied the 144 words selected for each language set in the present study on these 3 variables, 2 levels per variable. The imageability and log frequency scores for the Dutch words entered the matching procedure for the Dutch language set. Similarly, the imageability and log frequency scores for the English words entered the matching procedure for the English language set.[1] We had originally derived the frequency scores from the CELEX frequency count (see Burnage, 1990, for a description). This count is based on a Dutch corpus of 42.5 million printed words and an English corpus of 18.8 million printed words. Before transforming the absolute frequency scores in the English corpus to log frequency scores, we corrected for the size difference between the Dutch and English corpora by multiplying the absolute frequency scores of the English words by a factor of 2.26 (42.5/18.8). Each of the 8 stimulus groups per language condition thus formed (imageability by log frequency by cognate status; 2 levels for each variable) contained 18 words.

We matched the words in corresponding cells between the 2 languages. For instance, we matched the 18 words in the cell imageable/frequent/cognate of the Dutch language set with the 18 words in the cell imageable/frequent/cognate of the English language set on all 3 variables: imageability, log frequency, and cognate status. We performed 24 *t*-tests (8 stimulus groups by 3 word characteristics) to verify that this between-language matching had been successful. The *p*-values of these *t*-tests varied between .43 and .98, indicating that indeed the matching operation had succeeded. In other words, the imageability, log frequency, and cognate status of the words in each of the 8 Dutch word groups were never statistically different from these same word characteristics in the corresponding English word group. The mean imageability, log frequency, and cognate status scores for each of the 8

Table 1

Imageability (IMA), Log Frequency (FR), and Cognate Status (CS) Scores for All Stimulus Groups.

		Dutch			*English*		
Stimulus group		IMA	FR	CS	IMA	FR	CS
HI-HF-C	*M*	6.44	3.84	5.79	6.44	3.81	5.69
	SD	(.32)	(.38)	(.67)	(.37)	(.29)	(.53)
HI-HF-NC	*M*	6.43	3.85	1.31	6.46	3.83	1.31
	SD	(.24)	(.33)	(.19)	(.32)	(.23)	(.12)
HI-LF-C	*M*	6.34	2.78	5.41	6.40	2.81	5.45
	SD	(.35)	(.22)	(.70)	(.25)	(.29)	(.87)
HI-LF-NC	*M*	6.45	2.71	1.29	6.51	2.69	1.29
	SD	(.28)	(.30)	(.31)	(.39)	(.22)	(.13)
LI-HF-C	*M*	3.59	3.92	5.24	3.59	4.04	5.25
	SD	(.99)	(.44)	(.86)	(.79)	(.44)	(1.02)
LI-HF-NC	*M*	2.97	3.72	1.36	3.11	3.78	1.40
	SD	(.74)	(.34)	(.25)	(.40)	(.30)	(.28)
LI-LF-C	*M*	4.01	2.80	4.74	4.06	2.81	4.61
	SD	(1.27)	(.22)	(1.08)	(1.05)	(.41)	(1.14)
LI-LF-NC	*M*	2.99	2.79	1.23	3.10	2.76	1.27
	SD	(.88)	(.23)	(.10)	(.60)	(.34)	(.17)

Note. HI = High Imageability; LI = Low Imageability; HF = High Log Frequency; LF = Low Log Frequency; C= Cognates; NC = Noncognates.

cells within both language sets appear in Table 1. The scores for each individual stimulus are provided in the Appendix. In addition to the 288 test words, there were 30 practice words. Fifteen of them, Dutch nouns, preceded the Dutch test set; the remaining 15 were English nouns preceding the English test set.

Apparatus and Procedure

We ran the experiment on an Apple Macintosh Plus computer in a normally lit room. The computer screen presented the stimuli in black lowercase letters against a light-grey background. A PASCAL program controlled the stimulus presentation and the recording of the response times (RTs). A microphone that activated a voice-operated switch registered the participants' responses. Participants sat facing the screen at a comfortable reading distance. The experimenter sat to the left of the participant, typing his/her responses on the computer keyboard, and monitoring the workings of the voice switch. The experimenter noted down failures of the voice switch to respond to the participant's response and triggering by another sound.

Prior to the experiment the participants received both oral and written instructions. We asked them to provide their translation responses as quickly as possible, while making as few errors as possible, and to remain silent if they did not know the translation of the stimulus word. We encouraged them not to produce any other noises than the intended response so as not to activate the voice switch inadvertently, and to speak out the intended response clearly and loudly right from the onset so that the voice switch would detect it. We rotated the order in which we presented the 2 language sets between the participants; half the participants received the Dutch words (to be translated into English) before the English words (to be translated into Dutch), the other half in the reverse order. There was a pause of about 5 minutes between the 2 language sets; the complete test session took about 45 minutes. We presented the words within a language set in a random order (different for each participant) and in blocks of 24 words each.

After each block, we allowed the participant a brief rest. The experimenter subsequently initiated the next block by hitting the Return key. Fifteen practice words preceded the test words of each language set.

The sequence of events on each trial was as follows: Prior to the stimulus word, a fixation stimulus (an asterisk) appeared in the middle of the screen for 1 second, slightly above where the word was to appear. Immediately after it disappeared, the stimulus word appeared and remained on the screen until the voice switch registered the start of the participant's response (or any other sound). We measured RT from the onset of the stimulus word. The experimenter then typed the participant's response (what was being typed did not appear on the screen). Finally, the experimenter touched the Return key, 1 second after which the next trial started. The maximum presentation per stimulus was 5 seconds. If this duration expired with no response, the experimenter typed the Dutch word for "none" and called the next trial by hitting the Return key.

Results and Discussion

For each participant in each of the 3 proficiency groups we calculated 16 mean response times, one for each of the 16 conditions formed by the 2 levels of each of the variables: translation direction (Dutch to English, forward, vs. English to Dutch, backward); imageability (high imageability, HI, vs. low imageability, LI); log frequency (high frequency, HF, vs. low frequency, LF); and cognate status (cognates vs. noncognates). We also calculated a mean RT for each test word in both language sets, collapsed across participants. In calculating these means we excluded RTs of error responses (a translation registered by the voice-switch, but incorrect). Overall, errors occurred on 6.61% of the trials. We also excluded RTs associated with faulty (premature or delayed) voice-switch registrations; these occurred on 2.90% of the trials. Finally, we excluded responses that, according to Dutch-English translation dictionaries, could also be considered correct, but that were not the responses we intended (2.26% of the responses). We excluded the latter because their scores on the variable "cognate status," would be flawed; including them would thus have compromised the data. We analyzed the error data and the omission data (no response at all and, consequently, no associated RT-omissions occurred on 10.82% of the trials) separately. The total percentage of trials excluded from the RT analyses (errors + faulty voice-switch registrations + correct translations not intended by the investigators + omissions) was 22.59% overall. The RT analyses thus concerned 77.41% of all trials (university students, 86.62%; Grade 5, 77.96%; Grade 3, 67.67%).

We performed two 3 (proficiency) by 2 (translation direction) by 2 (imageability) by 2 (log frequency) by 2 (cognate status) ANOVAs on the mean RTs, one by participants and one by items. In the analysis by participants we treated proficiency as a between-participants variable, and the remaining 4 variables as within-participant variables. In contrast, in the analysis by items, proficiency served as a within-item variable whereas we treated all others as between-items variables. We performed this same pair of analyses on the error data, that is, on the percentage of translation errors per participant and per item in the various conditions. Finally, we performed this pair of analyses on the omission data.

Instead of discussing all possible interactions between all 5 variables, a rather unwieldy project, we will focus on the central questions: First, what variables determine translation performance? Second, do they do so in the same way across the 3 proficiency levels? Third, are forward and backward translation qualitatively different processes, as assessed by the size of the imageability effect in the 2 translation directions? Thus we report the main effects of all variables, the interaction between proficiency on the one hand and the remaining variables on the other, and, finally, the interaction between imageability and translation direction. Other interactions, as a rule, we do not report, unless a particular outcome is salient in view of data from other, related studies.

Main Effects

On all 3 sets of analyses, that is, those of the RT data, the error data (ER), and the omission data (OS), all main effects except one were significant, at the .05% level or better, on both the analyses by participants (*F1*) and that by items (*F2*). The only exception, the main effect of translation direction on the analysis of omissions, was only statistically reliable by participants and not by items. Table 2 shows a summary of these main effects and the corresponding statistics.

Not surprisingly, the university students generally performed better than the Grade 5 pupils, who in turn did better than the Grade 3 pupils, on all 3 measures: response times, errors, and

Table 2

Mean Response Times (RT = Milliseconds), Error Rates (ER = Percentages), and Omission Scores (OS = Percentages) for All Variables.

	Proficiency		
	University	Grade 5	Grade 3
RT	1095	1170	1347
ER	4.36	6.67	8.65
OS	5.40	9.74	17.20

RT: *F1* (2,57) = 11.68, and *F2* (2,544) = 146.00
ER: *F1* (2,58) = 14.26, and *F2* (2,546) = 32.75
OS: *F1* (2,58) = 22.39, and *F2* (2,546) = 128.60

	Translation direction	
	Forward	Backward
RT	1141	1267
ER	5.18	7.93
OS	11.43	10.14

RT: *F1* (1,57) = 34.63, and *F2* (1,272) = 19.11
ER: *F1* (1,58) = 62.46, and *F2* (1,273) = 7.12
OS: *F1* (1,58) = 5.34, and *F2* < 1

Table 2 (continued)

Mean Response Times (RT = Milliseconds), Error Rates (ER = Percentages), and Omission Scores (OS = Percentages) for All Variables.

	Imageability	
	high	low
RT	1143	1265
ER	4.17	8.94
OS	6.41	15.16
RT: *F1* (1,57) = 67.73, and *F2* (1,272) = 28.12		
ER: *F1* (1,58) = 104.63, and *F2* (1,273) = 21.93		
OS: *F1* (1,58) = 198.23, and *F2* (1,273) = 30.68		
	Log frequency	
	high	low
RT	1075	1333
ER	3.26	9.85
OS	4.20	17.36
RT: *F1* (1,57) = 567.19, and *F2* (1,272) = 90.41		
ER: *F1* (1,58) = 175.83, and *F2* (1,273) = 41.55		
OS: *F1* (1,58) = 240.72, and *F2* (1,273) = 70.42		
	Cognate Status	
	Cognates	Noncognates
RT	1101	1307
ER	4.88	8.24
OS	7.67	13.89
RT: *F1* (1,57) = 148.46, and *F2* (1,272) = 42.03		
ER: *F1* (1,58) = 49.54, and *F2* (1,273) = 10.58		
OS: *F1* (1,58) = 129.53, and *F2* (1,273) = 15.15		

omissions. Newman-Keuls tests showed that all but one of the differences between the 3 groups on all 3 dependent variables were statistically reliable ($p < .05$ or better). The exception, the RT difference between the university students and the Grade 5 pupils, failed to reach significance on the analysis by participants. Furthermore, in agreement with findings reported before, translation performance was better for words easy to imagine than for words difficult to imagine, for more frequent words than for less frequent ones, and for cognates than for noncognates. The response time, error, and omission data all converge on these conclusions. Finally, in agreement with an earlier study (De Groot et al., 1994), backward translation produced more errors but fewer omissions than forward translation. However, the effect of translation direction on response time does not agree with all earlier reports. For example, Kroll and Stewart's (1994) asymmetry model of word translation predicts shorter response times for backward than for forward translation. As Table 2 shows, we obtained the opposite result here.

Interactions Between Proficiency and the Remaining Variables

Table 3 presents, for all 3 dependent variables, all first-order interactions between proficiency on the one hand and each of the remaining independent variables on the other. Below each column we indicate whether or not that particular interaction was statistically significant.

The interaction of proficiency with translation direction was significant on the RT and ER analyses, but not on the OS analyses: $F1\ (2,57) = 3.60, p < .05$, and $F2\ (2,544) = 18.10, p < .0001$ for RT; $F1\ (2,58) = 5.49, p < .01$, and $F2\ (2,546) = 3.69, p < .05$ for ER; $F1 < 1$, and $F2 < 1$ for OS. As Table 3 shows, the effect of translation direction on RT was much larger for the 2 lower-proficiency groups than for the university students. In fact, simple effects analyses testing the effect of translation direction in each of the proficiency groups indicated that for the university students the effect of direction was not significant ($p > .10$ on the analyses both by participants and by items, as opposed to $p < .001$ on both analyses for the other 2 groups). The effect of translation direction on errors, with backward translation producing more errors than forward translation, increased with decreasing proficiency levels.

However, some of the above findings are qualified by higher-order interactions. The analysis of the omission data by participants (but not by items) produced significant second-order interactions between translation direction, proficiency, and imageability, and between the first two of these and log frequency. In turn, both of these second-order interactions were further qualified by the third-order interaction between proficiency, trans-

Table 3

Mean Response Times (Milliseconds), Error Rates (Percentages), and Omission Scores (Percentages) for All Conditions Formed by the Interactions Between Proficiency and the Remaining Variables.

	Response times							
Proficiency level	Translation direction		Imageability		Frequency		Cognate status	
University	FW	1071	high	1042	high	959	C	987
	BW	1120	low	1149	low	1232	NC	1204
Effect		49		107		273		217
Grade 5	FW	1100	high	1120	high	1046	C	1066
	BW	1240	low	1220	low	1294	NC	1274
Effect		140		100		248		208
Grade 3	FW	1253	high	1267	high	1220	C	1251
	BW	1441	low	1427	low	1474	NC	1443
Effect		188		160		254		192
	F1:s;F2:s		F1:ns;F2:s		F1:ns;F2:ns		F1:ns;F2:ns	

	Error rates							
Proficiency level	Translation direction		Imageability		Frequency		Cognate status	
University	FW	3.74	high	2.38	high	1.76	C	2.99
	BW	4.98	low	6.34	low	6.95	NC	5.73
Effect		1.24		3.96		5.19		2.74
Grade 5	FW	5.17	high	4.20	high	3.47	C	4.90
	BW	8.16	low	9.13	low	9.86	NC	8.44
Effect		2.99		4.93		6.39		3.54
Grade 3	FW	6.63	high	5.94	high	4.55	C	6.74
	BW	10.66	low	11.35	low	12.74	NC	10.56
Effect		4.03		5.41		8.19		3.82
	F1:s;F2:s		F1:ns;F2:ns		F1:s;F2:s		F1:ns;F2:ns	

Proficiency level	Omission scores: Translation direction		Imageability		Frequency		Cognate status	
University	FW	6.25	high	3.11	high	1.43	C	3.92
	BW	4.55	low	7.69	low	9.37	NC	6.88
Effect		1.70		4.58		7.94		2.96
Grade 5	FW	10.52	high	5.83	high	3.13	C	7.08
	BW	8.96	low	13.65	low	16.35	NC	12.40
Effect		1.56		7.82		13.22		5.32
Grade 3	FW	17.50	high	10.28	high	8.06	C	12.01
	BW	16.91	low	24.13	low	26.35	NC	22.40
Effect		0.59		13.85		18.29		10.39
		F1:ns;F2:ns		F1:s;F2:s		F1:s;F2:s		F1:s;F2:s

Note. FW = Forward; BW = Backward; C = Cognates; NC = Noncognates.

lation direction, imageability, and log frequency (significant by participants only), *F1* $(2,58) = 4.10, p < .05$, and *F2* $(2,546) = 1.27, p > .10$. This interaction points out that for one stimulus type in 2 of the proficiency groups the common finding (Tables 2 and 3) that forward translation leads to more omissions than backward translation does not materialize: Grade 5 and Grade 3 pupils produced *more* omissions when translating low-frequent, low-imageability words backward (24.17% and 38.61%, respectively) than when translating them forward (20.28% and 30.28%, respectively). Tables 4 to 6 clearly show this effect (split up for cognates and noncognates). They present the mean RT and percentages of errors and omissions for all cells, organized around the imageability effect; each table summarizes the data for one proficiency group.

The previous finding suggests that in backward translation the lower-proficiency groups perform particularly poorly on a narrowly circumscribed subset of the stimulus materials. We can draw the same conclusion from the second-order interaction between proficiency, translation direction, and cognate status on the analyses of the error data—*F1* $(2,58) = 5.06, p < .01$, and *F2* $(2,546) = 5.04, p < .01$, which indicates that these groups err in backward translation of noncognates in particular. This interaction was qualified further by the third-order interaction between

Table 4

Mean Response Times (in Milliseconds), Error Rates (in Percentages), and Omission Scores (in Percentages) for All Translation Conditions. University Students.

	High frequency words						Low frequency words					
	Cognates			Noncognates			Cognates			Noncognates		
	RT	ER	OS	RT	ER	OS	RT	ER	OS	RT	ER	OS
Forward												
HI	773	0.4	0.1	937	1.0	0.8	1047	2.5	4.1	1362	4.9	10.5
LI	914	1.6	2.4	1101	3.8	5.4	1116	4.4	9.1	1316	11.3	17.7
Effect	141	1.2	2.3	164	2.8	4.6	69	1.9	5.0	-46	6.4	7.2
Backward												
HI	814	0.5	0.2	1020	1.6	0.6	1079	1.7	3.0	1304	6.5	5.7
LI	976	1.9	0.6	1133	3.3	1.4	1178	11.0	11.8	1458	13.4	13.2
Effect	162	1.4	0.4	113	1.7	0.8	99	9.3	8.8	154	6.9	7.5

Note. HI = High Imageability; LI = Low Imageability.

Table 5

Mean Response Times (in Milliseconds), Error Rates (in Percentages), and Omission Scores (in Percentages) for All Translation Conditions. Grade 5 Students.

	High frequency words						Low frequency words					
	Cognates			Noncognates			Cognates			Noncognates		
	RT	ER	OS	RT	ER	OS	RT	ER	OS	RT	ER	OS
Forward												
HI	856	0.6	0.3	961	1.4	3.3	1139	6.1	11.4	1329	6.4	16.1
LI	934	4.4	4.4	1125	4.4	8.1	1113	7.5	15.6	1340	10.6	25.0
Effect	78	3.8	4.1	164	3.0	4.8	-26	1.4	4.2	11	4.2	8.9
Backward												
HI	888	0.6	0.6	1170	2.2	0.6	1196	3.9	4.7	1420	12.5	9.7
LI	1119	4.2	1.4	1316	10.0	6.4	1284	11.9	18.3	1528	20.0	30.0
Effect	231	3.6	0.8	146	7.8	5.8	88	8.0	13.6	108	7.5	20.3

Note. HI = High Imageability; LI = Low Imageability.

Table 6

Mean Response Times (in Milliseconds), Error Rates (in Percentages), and Omission Scores (in Percentages) for All Translation Conditions. Grade 3 Students.

	High frequency words						Low frequency words					
	Cognates			Noncognates			Cognates			Noncognates		
	RT	ER	OS	RT	ER	OS	RT	ER	OS	RT	ER	OS
Forward												
HI	990	0.8	0.8	1090	1.4	5.8	1294	6.9	15.0	1389	9.7	26.4
LI	1084	3.9	10.6	1374	5.6	20.8	1302	12.5	21.9	1505	12.2	38.6
Effect	94	3.2	9.8	284	4.2	15.0	8	5.6	6.9	116	2.5	12.2
Backward												
HI	1100	1.4	0.6	1328	3.1	1.9	1419	7.8	11.7	1529	16.4	20.0
LI	1280	6.7	4.7	1518	13.6	19.2	1540	13.9	30.8	1814	22.5	46.4
Effect	180	5.3	4.1	190	10.5	17.3	121	6.1	19.1	285	6.1	26.4

Note. HI = High Imageability; LI = Low Imageability.

proficiency, translation direction, cognate status, and imageability, although the latter interaction was only significant on the analysis by participants, $F1\ (2,58) = 4.57, p < .05$, and $F2\ (2,546) = 1.84, p > .10$. The two lower-proficiency groups performed especially poorly in backward translation of noncognates low on imageability (cf. Tables 4 to 6).

Imageability interacted significantly with proficiency when RT was the dependent variable, but only on the analysis by items. Furthermore, it interacted significantly with proficiency on the OS analyses, both by participants and by items. The interaction between imageability and proficiency was not reliable on the ER analyses: $F1\ (2,57) = 1.60, p > .10$, and $F2\ (2,544) = 4.96, p < .01$ for RT; $F1 < 1$, and $F2 < 1$ for ER; $F1\ (2,58) = 19.12, p < .0001$, and $F2\ (2,546) = 18.67, p < .0001$ for OS. Table 3 shows that imageability's effect on RT was particularly large for the lowest-proficiency group, and that the difference between HI and LI words in terms of percentage of omissions increased with decreasing proficiency level.

A number of these first-order interactions need qualification, however: On the analyses of the RT data the second-order interaction between proficiency, imageability, and cognate status was statistically reliable on the analysis by participants and approached significance on the analysis by items, $F1$ (2,57) = 3.08, $p = .05$, and $F2$ (2,544) = 2.90, $p = .06$. This interaction indicated that for the Grade 3 pupils one group of materials, low-imageability noncognates, was particularly problematic. Although the imageability effect (i.e., the difference in processing time between HI and LI words) for cognate materials in the Grade 3 participants was about as large as the imageability effect for both cognates and noncognates in the 2 other groups (all 5 effects ranged between 94 ms and 118 ms), for noncognates it was considerably larger (218 ms). This interaction between proficiency, imageability, and cognate status was also significant (by participants) and marginally significant (by items) on the analyses of omissions, $F1$ (2,58) = 5.75, $p < .01$, and $F2$ (2,546) = 2.46, $p = .09$. Once more, low-imageability noncognates turned out to be exceptionally hard for the Grade 3 pupils. They gave rise to an omission percentage of 31.25, as compared to 17.01 on low-imageability cognates (Table 6).

The imageability effect suggests the involvement of conceptual memory in word translation; a comparison of the effect's size across the 3 groups should inform us of the degree of conceptual-memory involvement at different levels of L2 proficiency. On the basis of current views on translation processes in bilinguals of various levels of L2 expertise, we expected the imageability effect would be largest for the highest-proficiency group. The most noteworthy aspect of the present data is that, if anything, the imageability effect was largest in the *lowest* proficiency group. The analyses substantiated this observation, but qualifiedly: The differential influence of imageability in the 3 proficiency groups held for a particular subset of words, namely, noncognates. Table 6 suggests that infrequent noncognates were especially problematic for the Grade 3 pupils. (However, this was not substantiated statistically: The third-order interaction between proficiency, imageability, cognate status, and frequency was not significant.) It appears that the lowest-proficiency group exploits conceptual memory in word translation to the same extent as do the higher-proficiency groups. In the General Discussion section of this paper we will propose a view of word translation that can account for the slightly larger imageability effects in the lowest-proficiency group without having to draw the contradictory conclusion that they in fact concept-mediate *more* than do the more fluent groups.

Log frequency interacted significantly with proficiency on the analyses with ER and OS as the dependent variables, but not on those with RT as the dependent variable: $F1 < 1$, and $F2 < 1$ for RT; $F1\ (2,58) = 3.08, p = .05$, and $F2\ (2,546) = 3.71, p < .05$ for ER; $F1\ (2,58) = 12.45, p < .0001$, and $F2\ (2,546) = 23.68, p < .0001$ for OS. The percentage difference in errors and omissions between words of high and low frequency became larger with decreasing L2 proficiency.

Finally, the interaction of proficiency with cognate status was statistically reliable on the OS analyses, but not on the RT and ER analyses: $F1 < 1$, and $F2 < 1$ for RT; $F1 < 1$, and $F2 < 1$ for ER; $F1\ (2,58) = 16.05, p < .0001$, and $F2\ (2,546) = 12.61, p < .0001$ for OS. The effect of cognate status on percentages of omissions became larger with lower proficiency levels.

Translation Direction and the Imageability Effect

The interaction between translation direction and imageability provides a test of the idea that conceptual memory is involved more in forward than in backward translation, a basic assumption in Kroll and her colleagues' asymmetry model of word translation (e.g., Kroll & Stewart, 1994). If their model is correct, the imageability effects should be larger in forward than in backward translation. Tables 4 to 6 show that, contrary to this prediction, these effects often appear *smaller* in forward translation. The statistical analyses support this observation. The analyses by participants (though not by items) all pointed at a significant interaction between translation direction and imageability: $F1\ (1,57) = 11.45$, $p < .01$, and $F2\ (1,272) = 2.02, p > .10$ for RT; $F1\ (1,58) = 18.11, p <$

.001, and $F2\ (1,273) = 2.07, p > .10$ for ER; $F1\ (1,58) = 15.05, p <$ $.001$, and $F2\ (1,273) = 1.22, p > .10$ for OS. Forward translation showed imageability effects of 88 ms (on the RT data), 3.34% (on the error data), and 7.09% (on the omission data); the corresponding effects for backward translation were 156 ms, 6.20%, and 10.42%. Tables 4 to 6 furthermore suggest that the interaction between translation direction and imageability in the RT data occurs because the low-frequency words often fail to produce an imageability effect in forward translation. Indeed, the second-order interaction between translation direction, imageability and frequency was statistically reliable on the analysis by participants of these data, $F1\ (1,57) = 5.76, p < .05$, and $F2\ (1,272) = 1.13, p > .10$. Collapsed across the proficiency groups, the low-frequency words translated forward showed a nonsignificant imageability effect of only 22 ms. The imageability effects in the remaining conditions, forward/high-frequent, backward/high-frequent, and backward/low-frequent, were 155 ms, 171 ms, and 143 ms, respectively. Not only does conceptual memory play a role in backward translation, but for one type of word (low-frequency words) it may even be involved *more* in backward than in forward translation.[2]

The Role of Word Length

We did not systematically manipulate the length of the stimulus and response words in this study. Yet, this variable has been shown to play a role in translation performance. Our earlier work (De Groot, 1992a; De Groot et al., 1994) has shown correlations of about .25 to .35 between stimulus and response length on the one hand and translation-production RT on the other. Furthermore, this earlier work indicated that length correlates negatively with both imageability and cognate status: Words easy to imagine and cognates tend to be shorter than words hard to imagine and noncognates.

We carried out additional analyses on the present data to find out whether or not our results needed qualifying in view of a possible effect by this length variable. First we performed an ANOVA with one variable (word group) and 4 levels: the lengths of the stimulus and response words, by both forward and backward translation. This analysis showed a significant effect of word group, $F(3,572) = 2.78, p < .05$. Subsequent Newman-Keuls analyses indicated no significant differences between the average lengths of the 4 word groups. However, a Duncan test pointed out that the lengths of the response words in forward and backward translation differed significantly from one another ($p < .05$); however, the difference in length was very small. The (English) response words in forward translation averaged 5.38 letters, the (Dutch) response words in backward translation 5.90 letters. The lengths of the stimulus words in forward and backward translation (Dutch and English words, respectively, with average lengths of 5.63 and 5.92 letters), did not differ significantly. Further analyses indicated that the significant length difference between the response words in forward and backward translation was due to the small number of words 10 or more letters long in the complete set (28 words out of 576—4 x 144). There were more 10-plus-letter words in the backward response words than in the forward response words (9 and 4, repectively). Removing all 28 of these long words from the analysis of word length across the 4 word groups meant that the length difference between the groups disappeared: $F(3,544) = 1.74, p > .10$. Furthermore, when we calculated new mean RTs for the 3 proficiency groups and for both directions of translation, excluding all 28 words of 10-plus letters, the critical finding that backward translation took longer than forward translation, especially for the lower-proficiency groups, remained: The differences between backward and forward translation time, with backward translation always taking longer, were 25 ms, 107 ms, and 194 ms, for the university students, the Grade 5 pupils, and the Grade 3 pupils, respectively (cf. Table 3). All in all, the small difference in length of the response words in forward and backward translation does not seem to underlie the large overall differences in RT.

In a final set of analyses, we calculated partial correlations between RT on the one hand and cognate status and the imageability of both the English and the Dutch words on the other, partialing out the length of the words. We performed these analyses for all 3 proficiency groups and for both translation directions. In all, we calculated 36 partial correlations.[3] Of these, only one dropped to a statistically nonsignificant value—the partial correlation between backward translation RT and the Dutch words' (the response terms') imageability when partialing out the length of the English words (the stimulus words) for the university students. Especially important, all the partial correlations between imageability and forward as well as backward translation RT for the 2 lower-proficiency groups remain significant. This finding converges with the earlier conclusion that concept mediation also occurs with lower levels of L2 proficiency, both in forward and in backward translation.

A Reanalysis of Earlier Data

One of the earlier translation studies performed in our laboratory (De Groot et al., 1994) had indeed weakly supported Kroll's asymmetry model (e.g., Kroll & Stewart, 1994), though the asymmetries predicted by the model only surfaced in a few of the relevant comparisons and were small. However, whenever an asymmetry occurred, it followed the direction predicted by Kroll's model (a larger role of semantic variables, including imageability, in forward translation), and *never* the opposite direction. The participants in that earlier study were comparable to those in our present highest-proficiency group, so why would a weak directional asymmetry emerge in that study but not in our present one, not even with participants drawn from the same population?

As we pointed out before (see Materials), we selected the stimuli used in the present investigation from the larger word set tested by De Groot et al. (1994). In other words, the stimulus set used in the 1994 study included words not tested here. Translation processes may differ for different types of words (e.g., De Groot, 1993). Therefore, the suggestion that in terms of translation asymmetries the present study does not replicate De Groot et al.'s (1994) results requires re-analyzing those earlier data associated with the 288 words re-used here. Only if the asymmetries in the earlier study remain for this smaller data set are we dealing with a clear incompatibility.

Therefore, we selected from the data of Experiment 2 of the 1994 data set all RT-data associated with the present 288 stimulus words. One reason to use Experiment 2 rather than Experiment 1 of the 1994 study for comparison was that Experiment 2 had evidenced more pronounced directional asymmetries than Experiment 1.[4] On these data we performed 2 analyses similar to those of the present study, one by participants and one by items: 2 (translation direction) by 2 (imageability) by 2 (log frequency) by 2 (cognate status). We did not include proficiency as a variable because all the 1994 participants had come from one population (first-year university students in psychology). A further difference between these re-analyses and those reported above was that we now treated translation direction as a between-participants variable (because each participant had performed either forward or backward translation). For the rest, the analyses were as reported above. We report only the outcomes relevant to the above discussion.

Again, imageability exerted a strong effect on translation performance, $F1\ (1{,}50) = 67.00, p < .0001$, and $F2\ (1{,}272) = 19.71, p < .0001$. Translation took 1077 ms and 1203 ms for high- and low-imageability words, respectively. But, particularly relevant in the present context, the direction of translation did not modulate the imageability effect: The interaction between imageability and translation direction was not significant, $F1\ (1{,}50) = .76, p > .10$, $F2 < 1$, suggesting that forward and backward translation are conceptually mediated to the same extent. Further, translation direction produced a significant effect, $F1\ (1{,}50) = 4.36, p < .05$, and $F2\ (1{,}272) = 9.52, p < .01$, with forward translation producing the *shorter* RT (forward, 1092 ms; backward, 1188 ms). This replicates the pattern for the lower-proficiency groups in the present study. Interestingly, the effect of translation direction was qualified by an interaction with frequency, $F1\ (1{,}50) = 5.30, p < .05$, $F2\ (1{,}272) = 4.63, p < .05$. High-frequency words were translated equally fast in both directions (984 ms and 1035 ms, a nonsignificant difference), but low-frequency words took longer to translate backward (forward, 1199 ms; backward, 1341 ms).

In sum, in this reanalysis of a subset of the 1994 data, imageability's asymmetrical effect in the 2 translation directions, predicted by the asymmetry model, disappeared. The data suggested as large a role for conceptual memory in backward translation as in forward translation. Furthermore, this reanalysis introduced for participants with an L2 fluency comparable to our present highest-fluency group the finding obtained for the present 2 lower-proficiency groups, namely, that backward translation (of low-frequency words) takes longer than forward translation.

General Discussion

Word Type

The findings that the imageability, cognate status, and frequency of the stimulus words affect translation performance replicate results from similar and other versions of the translation task (e.g., De Groot, 1992a; De Groot & Comijs, 1995; De Groot et al., 1994; Kroll & Stewart, 1994; Murray, 1986; Sánchez-Casas, Davis, & García-Albea, 1992). Previous studies have also reported effects of the former 2 of these variables on performance in other bilingual tasks than translation.

Imageability. An effect of imageability has also appeared in bilingual word association (Kolers, 1963; Taylor, 1976; Van Hell & De Groot, 1996) and in cross-language semantic priming (Jin, 1990). As De Groot (1993) argued, one way to explain all these imageability effects is to assume that concrete words share more (or larger parts; see our introductory section and De Groot, 1992b) of their conceptual representations between languages than do abstract words (but see, e.g., Paivio & Desrochers, 1980, for an alternative view).[5] (De Groot, 1993, discussed the details about this argument as well as the exact nature of the various bilingual tasks and the imageability effect therein.)

Cognate Status. Previous studies have obtained an effect of cognate status not only in word translation, but also in, again, bilingual word association (Taylor, 1976; Van Hell & De Groot, 1996), cross-language semantic priming (De Groot & Nas, 1991), and cross-language repetition priming (Davis, Sánchez-Casas, & García-Albea, 1991; Sánchez-Casas, Davis, & García-Albea, 1992). Like that of imageability, the effect of cognate status could be explained by cognates' and noncognates' sharing different amounts of conceptual representations between languages: Cognates may share more (or larger parts) of their conceptual representations between languages than do noncognates. De Groot (1992a) has suggested a second interpretation, namely, that the effect of cognate status in word translation may be due to stronger direct connections between the L1 and L2 word-form representations of cognate translations than between those of noncognate translations. Consequently, translation of a cognate would often proceed by use of this direct connection between the word-form representations of a translation pair (word-association translation), bypassing conceptual memory. As De Groot argued at the time, analogously to the analysis pursued here, variables that reflect meaning aspects of words, such as imageability, should, in that case, affect the translation of cognates less than that of noncognates. Our analyses show this is not *generally* the case (Tables 4 to 6), although the data for Grade 3 participants appear consistent with this prediction (which is supported by the second-order interaction between proficiency, imageability, and cognate status). One could conclude that bilinguals with a low level of proficiency in L2 rely more on word-association translation than do more fluent bilinguals (Kroll & De Groot, 1997, offer some support for this position). However, the Grade 3 students produced as large an imageability effect for cognates as did the 2 higher-fluency groups, a contrary finding. (The size of the imageability effect for noncognates—considerably larger for the Grade 3 students than for the other 2 groups—caused the interaction between proficiency, imageability, and cognate status). Thus, concept-mediation apparently occurs not only in noncognate but also in cognate translation. Kroll and Stewart (1994, p. 164) drew the same conclusion, though restricted to forward translation.

Word Frequency. Earlier word-translation studies (e.g., De Groot, 1992a) have attributed the effect of our third stimulus manipulation, word frequency, to differences between frequent and infrequent words regarding the strength of the various connections within the bilingual memory structures that represent these words. Departing from the standard model of bilingual memory set forth in our introduction (two language-specific word-form memory stores, one language-independent conceptual memory), we assume the effect arises from stronger connections between the L1 and L2 word-form representations, on the one hand, and the conceptual representation shared by L1 and L2, on the other, for frequent words than for infrequent words. If these indirect (via conceptual memory) connections between the L1 and L2 word-form representations are stronger for frequent words, the translation of frequent words should relatively often exploit them; in other words, the translation of frequent words should relatively often be conceptually mediated. In contrast, infrequent words will more often be translated via the direct connection between the two word-form representations.

If these hypotheses are valid, and if our analysis that the (degree of) conceptual memory's involvement in word translation can be deduced from the occurrence of an imageability effect (and its size) is correct, then frequent words should show larger imageability effects than infrequent words. The RT data on forward translation substantiated this prediction; only high-frequency words demonstrated a statistically reliable imageability effect, replicating De Groot's earlier data (1992a, Experiment 1). The backward-translation data, however, showed equal imageability effects for both high- and low-frequency words, suggesting an equally large involvement of conceptual memory for both types of words in backward translation.

Proficiency

Performance was on the whole better—on all 3 performance measures—with higher levels of L2 proficiency than with lower levels. This finding, though hardly remarkable, does demonstrate that our study's design and procedure were sensitive enough to detect proficiency-based differences.

Our main question about proficiency was whether it affects the degree of conceptual involvement in word translation, as assessed by the size of the imageablity effect. As discussed in the introductory section, the view of word-association translation and concept-mediation translation by bilinguals with lower and higher levels of L2 proficiency, respectively, predicts larger (or even exclusive) imageability effects with higher levels of proficiency. In this light, the pattern of data that we obtained (Table 3) *is* remarkable. The analyses of response times and errors showed equally large imageability effects across the 3 proficiency groups. (In fact, the least proficient group showed the largest effect on the response-time analyses; but the difference in effect size between the groups was not reliable on the analysis by participants). In addition, analyses of the omission data showed that the imageability effect increased not with increasing but with decreasing proficiency. In other words, our data controvert the idea that, while translating words, bilinguals with less L2 fluency concept-mediate less than those with more L2 fluency. Except for the omission data, the data patterns were similar across the 3 proficiency groups, suggesting that word translation was qualitatively the same for all.

Does this conclusion invalidate the common view on how L2 proficiency relates to bilingual memory and processing? Not necessarily: As stated above, our demonstration (for RT and errors) that the imageability effect is immune to differences in L2 proficiency could be explained differently. The conceptual involvement in word translation may not develop gradually with increasing fluency in L2, but may instead be an all-or-none process. Bilinguals may rely on the word-association links between the L1 and

L2 word-form representations (not only in word translation, but also in L2 processing in general) *only* during the very initial stages of L2 learning, then "leap" at some point to conceptual-based translation. All 3 of our proficiency groups, even the least fluent, might have been beyond the early word-association stage.

The present data in fact fit those from Potter et al. (1984), who also found evidence of equal reliance on concept mediation for 2 bilingual groups at different stages of L2 development. Their lower-proficiency participants were American high school students who had studied French in school for 2 or 3 years. However, Chen and Leung's (1989) less fluent participants, who had also been trained in the non-native test language (L2 for one group, but L3 for another) for about 2 to 3 years, showed a data pattern consistent with word-association translation. Chen and Leung attributed the apparent inconsistency between their study and Potter et al.'s (1994) to the presumably higher L2 fluency of the participants in the latter study, despite an about equal duration of school training in the target language. The least proficient bilinguals in Kroll and Curley's (1988) study also demonstrated a data pattern consistent with word-association translation. These non-fluent participants were American undergraduates who had participated in German classes for less than 30 months; hence, the average duration of their L2 training may have been well below the 2 to 3 years' training of Potter et al.'s non-fluent participants. All these studies, including our present one, concur with research on the role of expertise in problem solving. When "prenovices," novices, and experts are solving, for instance, problems in physics or software design, novices and experts operate in qualitatively the same way, whereas prenovices employ qualitatively different strategies (see VanLehn, 1989, for a review). However, our data do not agree with De Groot and Hoeks' (1995) finding of null-effects for the imageability manipulation in Dutch-English-French trilinguals who translated from L1, Dutch, to L3, French, despite the fact that the participants had been trained in French for 6 years. This inconsistency led De Groot (1995, pp. 164–165) to suggest that recent use or disuse of a foreign language may play a role in how bilinguals (or multilinguals) exploit the relevant memory structures. De Groot and Hoeks' participants had not used French for quite some time. Could prolonged disuse have made them revert from concept mediation to word association?[6]

In addition to the participants' exact level of L2 proficiency and, possibly, whether or not they have recently used an L2, one further factor may determine whether or not nonfluent bilinguals exhibit concept mediation: namely, the exact nature and demands of the task. Contrary to their prediction, Dufour and Kroll (1995) obtained support for concept mediation in nonfluent bilinguals. They used a semantic categorization task, where the participants indicated on every trial whether the referent of the test word was a member of the category to which a previously presented word referred. The authors suggested (p. 177) that concept mediation may have occurred for the nonfluent group because semantic categorization (unlike, for instance, word translation) does not require participants to produce words in L2.

Translation Direction

As for translation direction, we can account for two of its main effects in terms of a number of characteristics of unbalanced bilingualism. Consistent with earlier results (De Groot et al., 1994), in our present study, forward translation gave rise to more omissions and fewer errors than did backward translation. In 1994 we attributed these findings to differences in the size of the L1 and L2 vocabularies of the participants (who, like the participants in the present study, were all unbalanced bilinguals) and to differences in the "quality" of the knowledge stored in the two vocabularies: The L2 vocabulary of unbalanced bilinguals is likely to be smaller than their L1 vocabulary, and the knowledge pertaining to the L2 words that *are* stored is probably less well established and more often incorrect than L1 lexical knowledge. The absence of a stimulus' translation term in the other lexicon will result in an omission, whereas a weakly established or faulty representation will cause a problem of understanding (uncertainty as to the stimulus word's meaning and/or assigning to it an incorrect meaning) and, as a consequence, a translation error. Because of the differences between the L1 and L2 vocabularies, L1 to L2 translation will more often result in an omission than L2 to L1 translation (the L2 lexicon is smaller), and L2 to L1 translation will lead to more errors (L2 elements are relatively often misunderstood). Recently, La Heij et al. (1996) presented an elegant account of the processes underlying word-translation performance from which the present ideas about the source of errors and omissions in translation ensue naturally (see below).

Kroll and colleagues' asymmetry model (Kroll, 1993; Kroll & Stewart, 1994; Sholl et al., 1995) could also quite easily account for the effect of translation direction on omissions and errors. However, their model cannot handle the main effect of translation direction on response time. The asymmetry model predicts faster backward than forward translation. Indeed, this finding has been obtained repeatedly (e.g., Kroll, 1993, for a list of five studies that all demonstrated the effect). Yet, we obtained the opposite result (qualified however by proficiency; the group with the highest L2 fluency showed a null-effect of translation direction on response time). Kroll's model predicts faster backward translation because it assumes that backward translation exploits the direct link between the L2 and L1 word-form representations (word-association translation) whereas forward translation proceeds through conceptual memory via two links: the one from the L1 word-form representation to the conceptual representation shared by L1 and L2, and the link from the latter representation to the L2 word-form representation (translation through concept-mediation). Because this second translation route is longer, it should take more time to complete.

Our findings of either a null-effect of translation direction on response times (the highest-proficiency group) or of faster responses in *forward* translation (the 2 lowest-proficiency groups) are not unique. The former was also obtained by De Groot et al.

(1994, Experiment 1), La Heij et al. (1996, Experiment 4), and Van Hell and De Groot (1995); the latter occurred for De Groot et al. (1994, Experiment 2), La Heij et al. (1996, Experiment 3), and Swaak (1992). All of these studies' participants were (Dutch) university undergraduates, as were those in our present highest-proficiency group. In sum, a number of studies do not support the prediction of the asymmetry model that backward translation is faster than forward translation; in fact, they have shown the opposite result with bilinguals of various levels of L2 fluency.

Our data also controvert the second prediction of the asymmetry model, that conceptual memory is implicated more in forward translation than in backward translation: In our study, the imageability effect was not smaller but *larger* in backward translation than in forward translation, although an interaction between imageability, translation direction and frequency on the RT analysis suggested that this finding only held for low-frequency words. High-frequency words showed an equally large imageability effect in the two translation directions.

All in all, our data do not support the main predictions of the asymmetry model of word translation. Kroll and De Groot (1997) have discussed why the data of the various studies do not converge better. For one instance, the different studies have used different stimulus sets. That this factor is indeed relevant is suggested by our reanalysis of De Groot et al.'s (1994) data: In a subset of the earlier stimuli, the asymmetries observed in 1994 disappeared. Kroll and De Groot provided some evidence that differences between bilingual groups and between tasks used in the various studies are also critical. We will not repeat their discussion here, but merely conclude that the directional asymmetries as observed by Kroll and her colleagues (Kroll, 1993; etc.) may be tied to limited conditions, the exact nature of which further research has yet to determine. Instead of speculating on this nature, we will focus on a parsimonious, simple view of word translation advanced by La Heij et al. (1996) and (in a different terminology) by Snodgrass (1993), a view that may account for many of the reported patterns of translation data without assuming qualitatively different translation processes in the two directions of translation.

Concept Mediation as the Standard Translation Process

La Heij et al. (1996) assumed that *both* forward and backward translation are largely conceptually mediated. They suggested that one can decompose the translation process in either direction into 2 processing stages: Stage 1, determining the meaning of the presented word ("concept activation"); and Stage 2, retrieving the response word on the basis of the conceptual information activated in Stage 1 ("word retrieval"). When bilinguals with a complete mastery of their L1 face a word in their L1, Stage 1, concept activation, will generally proceed smoothly (being a well-practiced skill), but Stage 2, word retrieval in the L2, may often be a problem. In other words, in forward translation Stage 2 is probably the more vulnerable of the two. In contrast, when these same bilinguals face a word in the L2, to be translated into the L1 (backward translation), Stage 1 is the more vulnerable stage. These views—that concept activation is relatively easy with an L1 word stimulus and that word retrieval is relatively easy with an L2 stimulus—are in fact two of the basic assumptions of Kroll and colleagues' asymmetry model as well (e.g., Kroll & Stewart, 1994). La Heij et al. and the asymmetry model share a third basic assumption: that unbalanced bilinguals' L2 lexicon is smaller than their L1 lexicon. However, whereas the asymmetry model developed this set of basic assumptions into the proposal of qualitatively different forward and backward translation processes, La Heij et al. proposed that translation in both directions is effectuated through concept mediation. Although Snodgrass (1993) labeled the 2 stages, concept activation and word retrieval, differently ("encoding" and "production," respectively), her views on word translation are basically the same as La Heij et al.'s.

This parsimonious account of word translation can deal with many of the relevant findings in word-translation research. As already mentioned, it can handle the finding that forward translation gives rise to more omissions than backward translation, whereas backward translation results in more errors. Reaction time differences between forward and backward translation become explicable in terms of differences in the relative difficulty of concept activation and word retrieval in the 2 directions of translation. If concept activation of an L2 stimulus is relatively more difficult than word retrieval of the same L2 word, then backward translation will take longer than forward translation. In the event that L2 word retrieval is exceptionally hard, then forward translation will take longer. Finally, when the difficulty of concept activation of an L2 word equals that of L2 word retrieval, translation direction will not affect RT. In this view, our data would suggest that for the highest-proficiency group concept activation and word retrieval in L2 were in balance, whereas the two lower-proficiency groups found activating concepts for L2 words particularly problematic.

A number of higher-order interactions in our data become explicable if we assume that the lower-proficiency groups found some narrowly circumscribed subsets of the stimulus materials exceptionally problematic in backward translation: As suggested earlier, the interaction between proficiency, translation direction, imageability, and log frequency on the omission data indicated that omissions occurred when the 2 lesser-fluency groups translated infrequent words with low imageability from L2 to L1. Thus, concept activation for this subset of L2 words apparently lags in the development of the L2 lexicon. In fact, this effect overruled the general tendency (see above) that forward translation produced more omissions than backward translation. Similarly, the interaction between proficiency, translation direction, and cognate status on the error data suggests that the less-fluent participants experienced exceptional problems with backward translation of noncognates. All in all, these findings indicate that (a) for less-fluent bilinguals (at least those of the type we tested) activating concepts for L2 words is particularly problematic, and (b) the degree to which they experienced problems in backward translation (and presumably when processing L2 words in other tasks as well) depends on specific characteristics of the stimulus words. For some types of words, progress appears slower than for other types. Finally, our reanalysis of De Groot et al.'s (1994) data indicates that participants with an L2 fluency level comparable to our most-fluent group still find concept activation of particular L2 stimulus words exceptionally difficult, as suggested by their longer response times for backward than forward translation of infrequent words.

Assuming that progress in L2 learning proceeds at a different speed for different types of words also explains two of the remaining reliable interactions in our data: namely, the interactions between proficiency, imageability, and cognate status in the response-time and omission data. Noncognates of low imageability are particularly problematic for the least-fluent group (but now in both directions).

In all, the simple set of ideas regarding word translation advanced by La Heij et al. (1996) and others (Snodgrass, 1993, but for the most part also by Kroll & Stewart, 1994) has considerable explanatory power. Just one finding in our data we cannot obviously account for in terms of the view that translation via concept mediation is a universal process, at least in all bilinguals beyond the very initial stages of word translation: No significant imageability effect occurred in forward translation of infrequent words (a similar result, for a different set of stimuli, also characterize De Groot's Experiment 1, 1992a). The specific backward-translation condition in Kroll and Stewart's (1994) study may have incidentally instantiated the (limited) circumstances required for translation that is not conceptually mediated.[7] We will refrain from speculating on exactly what critical characteristic these two conditions shared that apparently causes translation to bypass conceptual memory. Instead, we will conclude with the assertion that when translating words in either direction, bilinguals of various L2 fluency levels apparently access and exploit conceptual memory representations at least most of the time. Exactly which underlying trait causes a deviation from this pattern (in backward translation of categorized vs. mixed lists, see Note 7; at extremely low levels of fluency in L2, not tested here; in translating infrequent words from L1 to L2) remains to be established.

Revised version accepted 10 October 1996

Notes

[1]Both imageability and log frequency correlate highly across languages. The words in the present set of 144 Dutch words correlate .96 with their translations in English on imageability and .83 on log frequency. The words in the present set of 144 English words correlate .95 with their Dutch translations on imageability and .80 on log frequency.

[2]As can be seen in Tables 4 to 6, the average RTs for some cells in this study (e.g., for noncognates with low imageability and frequency values in the Grade 3 data set) rest on less than 50% of the maximum number of responses; we removed the rest of the responses from the analyses because they were, for instance, errors or omissions. A number of stimulus words in these cells in fact evoked no response or very few, from 1 to 4 out of the maximum of 20 within a proficiency group. The responses that *were* generated to these apparently excessively hard stimulus words may be unreliable. We therefore ran a second set of RT analyses, analogous to those reported in the text, excluding all data associated with stimulus words that generated fewer than 5 out of the 20 responses within a group. The pattern of results emerging from these analyses was essentially the same as that for the reported set of analyses: The same main and interaction effects were significant in the 2 sets of analyses, and all effects went in the same direction in both sets.

[3]Six of this total set of 36 partial correlations partialed out the length of the Dutch words from the correlations between RT and cognate status (2 for each proficiency group—one for forward translation RT and a second for backward translation RT); a similar set of 6 partialed out the length of the English words from the correlations between RT and cognate status; 12 partialed out the length of the Dutch words from either the correlations between RT and the imageability of the Dutch words or from the correlations between RT and the imageability of the English words (again for both directions of translation and for all 3 proficiency groups); finally, the last 12 were analogous to the previous 12, but now partialing out the length of the English words.

[4]A practical reason for not using Experiment 1's data from the 1994 study was that its forward data, collected well before the backward data and on a

different computer (see De Groot et al., 1994, for details), were stored on diskettes that our present computer facilities could not read.

[5]As mentioned in the introduction, word imageability and word concreteness are highly correlated. Hence, the terms "word imageability" and "word concreteness" are often used interchangeably, and words rated high or low on the imageability variable are often simply called "concrete" and "abstract", respectively.

[6]Dutch and French belong to different subgroups of the Indo-European family of languages, namely, Germanic and Romance, respectively. One could hypothesize that languages only share conceptual representations, and consequently enable translation through concept mediation, if they belong to the same subgroup of languages within a family of languages. That Dutch and French do not could then be the source of the null-effect of imageability when Dutch words are translated into French. This hypothesis, however, is invalidated by the fact that both Potter et al. (1984) and Chen and Leung (1989) obtained a data pattern suggesting concept mediation despite the fact that the languages they tested belonged to different language families.

[7]That Kroll and Stewart's (1994) data suggested the absence of conceptual involvement in backward translation whereas the present study clearly suggested its presence is not necessarily contradictory. These two studies assessed the involvement of conceptual memory in different ways: Kroll and Stewart had their participants translate words referring to exemplars from a limited set of semantic categories. They presented these words blocked by category, or mixed. They deduced conceptual involvement in translation from the effect of this blocking-mixing manipulation. They regarded any effect of this manipulation (as occurred in forward translation) as a signature of translation through concept mediation, and the absence of such an effect (as in backward translation) as support for word-association translation. Although conceptually the same, in actual practice this manipulation was very different from ours, and required the presentation of very different stimulus materials. Consequently, the conditions that give rise to word-association translation in their study and in ours may be totally different from one another.

References

Abunuwara, E. (1992). The structure of the trilingual lexicon. *European Journal of Cognitive Psychology, 4,* 311–322.

Burnage, G. (1990). *CELEX: A guide for users.* Nijmegen: SNN.

Chen, H.-C. (1990). Lexical processing in a non-native language: Effects of language proficiency and learning strategy. *Memory & Cognition, 18,* 279–288.

Chen, H.-C., & Leung, Y.-S. (1989). Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 316–325.

Davis, C. W., Sánchez-Casas, R. M., & García-Albea, J. E. (1991). *Bilingual lexical representation as revealed using the masked priming procedure.* Unpublished manuscript, St. Louis University, Madrid.

De Groot, A. M. B. (1992a). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1001–1018.

De Groot, A. M. B. (1992b). Bilingual lexical representation: A closer look at conceptual representations. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 389–412). Amsterdam: Elsevier.

De Groot, A. M. B. (1993). Word-type effects in bilingual processing tasks: Support for a mixed representational system. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 27–51). Amsterdam-Philadelphia: John Benjamins.

De Groot, A. M. B. (1995). Determinants of bilingual lexicosemantic organisation. *Computer Assisted Language Learning, 8*, 151–180.

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*, 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*, 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*, 683–724.

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*, 90–123.

Dufour, R., & Kroll, J. F. (1995). Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory & Cognition, 23*, 166–180.

Jin, Y.-S. (1990). Effects of concreteness on cross-language priming in lexical decisions. *Perceptual and Motor Skills, 70*, 1139–1154.

Kirsner, K., Smith, M. C., Lockhart, R. S., King, M. L., & Jain, M. (1984). The bilingual lexicon: Language-specific units in an integrated network. *Journal of Verbal Learning and Verbal Behavior, 23*, 519–539.

Kolers, P. A. (1963). Interlingual word associations. *Journal of Verbal Learning and Verbal Behavior, 2*, 291–300.

Kroll, J. F. (1993). Accessing conceptual representations for words in a second language. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 53–81). Amsterdam-Philadelphia: John Benjamins.

Kroll, J. F., & Curley. J. (1988). Lexical memory in novice bilinguals: The role of concepts in retrieving second language words. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues: Vol. 2, Clinical and educational implications* (pp. 389–395). Chichester: John Wiley & Sons.

Kroll, J. F., & De Groot, A. M. B. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (169–199). Hillsdale, NJ: Erlbaum.

Kroll, J. F., & Stewart, E. (1990, November). *Concept mediation in bilingual translation.* Paper presented at the 31st Annual Meeting of the Psychonomic Society, New Orleans.

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*, 149–174.

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*, 648–665.

Masson, M. E. J. (1991). A distributed memory model of context effects in word identification. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 233–263). Hillsdale, NJ: Erlbaum.

Murray, D. J. (1986). Characteristics of words determining how easily they will be translated into a second language. *Applied Psycholinguistics, 7*, 353–372.

Paivio, A., & Desrochers, A. (1980). A dual-coding approach to bilingual memory. *Canadian Journal of Psychology, 34*, 388–399.

Potter, M. C., So, K.-F., Von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23*, 23–38.

Sánchez-Casas, R. M., Davis, C. W., & García-Albea, J. E. (1992). Bilingual lexical processing: Exploring the cognate/non-cognate distinction. *European Journal of Cognitive Psychology, 4*, 293–310.

Sánchez-Casas, R. M., Suárez-Buratti, B. & Igoa, J. M. (1992, September). *Are bilingual lexical representations interconnected?* Paper presented at the fifth conference of the European Society for Cognitive Psychology, Paris.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language, 27*, 499–520.

Sholl, A., Sankaranarayanan, A., & Kroll, J. F. (1995). Transfer between picture naming and translation: A test of asymmetries in bilingual memory. *Psychological Science, 6,* 45–49.

Snodgrass, J. G. (1993). Translating versus picture naming: Similarities and differences. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 83–114). Amsterdam-Philadelphia: John Benjamins.

Swaak, J. (1992). Untitled. Unpublished master's thesis, University of Amsterdam.

Taylor, I. (1976). Similarity between French and English words: A factor to be considered in bilingual language behavior? *Journal of Psycholinguistic Research, 5,* 85–94.

Van Hell, J. G., & De Groot, A. M. B. (1995). *Disentangling context availability and concreteness in lexical decision and word translation.* Manuscript submitted for publication.

Van Hell, J. G., & De Groot, A. M. B. (1996). *Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association.* Manuscript submitted for publication.

VanLehn, K. (1989). Problem solving and cognitive skill acquisition. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 527–579). Cambridge, MA: MIT Press.

Appendix

Dutch set	*Imageability Score*	*Log Frequency*	*Cognate Status Score*
HI-HF-C			
boot (boat)	6.68	3.45	6.63
hand (hand)	6.62	4.64	6.89
zee (sea)	6.72	3.78	5.47
kapitein (captain)	5.69	3.62	5.84
trein (train)	6.77	3.54	5.79
moeder (mother)	6.56	4.40	5.42
mond (mouth)	6.52	3.98	4.74
brood (bread)	6.38	3.47	4.95
maan (moon)	6.60	3.44	4.68
water (water)	6.48	4.19	6.63
kat (cat)	6.68	3.48	6.58
vader (father)	6.24	4.39	5.47
vinger (finger)	6.76	3.82	6.42
grond (ground)	5.96	4.17	5.47
voet (foot)	6.56	3.98	5.84
prins (prince)	6.16	3.46	6.32
haar (hair)	6.60	3.61	5.68
politie (police)	5.92	3.61	5.32

Dutch set	*Imageability Score*	*Log Frequency*	*Cognate Status Score*
HI-HF-NC			
been (leg)	6.62	3.90	1.32
paard (horse)	6.65	3.82	1.26
dorp (village)	6.12	3.76	1.21
berg (mountain)	6.68	3.37	1.16
wang (cheek)	6.15	3.45	1.32
kamer (room)	6.00	4.19	1.26
bloem (flower)	6.68	3.60	1.21
fles (bottle)	6.46	3.68	1.32
raam (window)	6.35	3.87	1.16
jongen (boy)	6.44	4.18	1.32
dame (lady)	6.20	3.58	1.84
boer (farmer)	6.44	3.63	1.16
gezicht (face)	6.12	4.33	1.11
meisje (girl)	6.32	4.18	1.32
vrouw (woman)	6.50	4.58	1.74
stoel (chair)	6.84	3.81	1.37
vogel (bird)	6.58	3.61	1.26
tuin (garden)	6.56	3.70	1.16
HI-LF-C			
insekt (insect)	5.84	2.87	6.63
appel (apple)	6.84	2.99	6.16
roos (rose)	6.65	3.08	6.26
vos (fox)	6.60	2.50	4.68
koren (corn)	5.65	2.36	5.21
wiel (wheel)	6.27	2.95	5.68
zeep (soap)	6.31	2.83	5.11
cirkel (circle)	6.31	3.05	6.37
naald (needle)	6.38	2.82	4.68
peper (pepper)	6.36	2.83	5.63
tijger (tiger)	6.68	2.47	5.63
gitaar (guitar)	6.50	2.42	5.53
zeil (sail)	6.12	2.80	5.58
honing (honey)	5.73	2.72	5.21
tabak (tobacco)	6.16	2.75	5.32
bruid (bride)	6.20	2.66	4.95
leeuw (lion)	6.72	2.99	3.79
muis (mouse)	6.80	2.95	4.89
HI-LF-NC			
landkaart (map)	6.12	2.15	1.26
varken (pig)	6.46	2.98	1.16
handschoen (glove)	6.42	2.74	1.21
knoflook (garlic)	6.32	2.68	1.08

Dutch set	Imageability Score	Log Frequency	Cognate Status Score
haai (shark)	6.23	2.14	1.21
handdoek (towel)	6.56	2.84	1.08
begrafenis (funeral)	6.28	3.00	1.24
aardbei (strawberry)	6.80	2.31	1.32
kraan (tap)	6.48	2.80	1.42
kikker (frog)	6.62	2.57	1.37
plafond (ceiling)	6.04	3.04	1.21
schaar (scissors)	6.96	2.52	2.47
mier (ant)	6.80	2.49	1.12
sap (juice)	5.88	2.65	1.16
konijn (rabbit)	6.54	2.98	1.21
pijl (arrow)	6.44	2.84	1.21
eend (duck)	6.38	3.01	1.42
laars (boot)	6.72	3.05	1.08
LI-HF-C			
woord (word)	4.73	4.40	5.89
naam (name)	4.68	4.25	5.58
dag (day)	3.69	4.60	4.74
wereld (world)	5.08	4.28	4.58
jaar (year)	4.36	4.69	5.11
regel (rule)	3.65	3.77	4.00
kritiek (criticism)	2.62	3.49	3.96
inzicht (insight)	1.92	3.63	5.40
informatie (information)	2.65	3.66	6.16
figuur (figure)	4.31	3.59	6.00
taak (task)	3.08	3.80	4.84
tijd (time)	3.08	4.66	4.26
zoon (son)	4.46	3.90	5.42
kwaliteit (quality)	2.65	3.51	5.08
principe (principle)	1.73	3.51	6.32
leider (leader)	4.36	3.48	5.95
plan (plan)	3.60	3.93	6.89
wens (wish)	4.00	3.44	4.16
LI-HF-NC			
gelegenheid (opportunity)	2.35	3.65	1.28
gedachte (thought)	2.73	4.07	2.12
mening (opinion)	2.15	3.66	1.40
smaak (taste)	3.36	3.52	1.42
schoonheid (beauty)	3.77	3.27	1.16
vrede (peace)	3.64	3.34	1.53
geval (case)	1.96	4.36	1.48
vraag (question)	2.58	4.30	1.21
leeftijd (age)	2.77	3.64	1.08

Dutch set	*Imageability Score*	*Log Frequency*	*Cognate Status Score*
teken (sign)	3.68	3.59	1.32
voordeel (advantage)	2.48	3.46	1.16
bewijs (proof)	2.35	3.43	1.20
deel (part)	3.23	4.22	1.64
boodschap (message)	4.08	3.44	1.32
vijand (enemy)	2.77	3.41	1.16
onderwijs (education)	4.56	3.84	1.52
voorbeeld (example)	2.15	3.99	1.16
gevaar (danger)	2.85	3.69	1.37
LI-LF-C			
seizoen (season)	4.92	2.93	5.63
hitte (heat)	5.04	3.10	4.53
heiden (heathen)	2.12	2.58	4.52
spraak (speech)	3.20	2.38	3.88
schaamte (shame)	2.80	2.86	4.32
vloed (flood)	4.88	2.53	5.79
leugen (lie)	2.56	3.03	3.04
inkt (ink)	5.42	2.53	5.53
ijzer (iron)	5.50	2.88	3.11
pond (pound)	4.24	2.81	5.16
schandaal (scandal)	2.27	2.60	5.20
saus (sauce)	5.54	2.81	6.11
hel (hell)	5.08	2.97	6.89
wijsheid (wisdom)	2.68	3.04	4.04
paniek (panic)	3.56	3.08	5.84
eed (oath)	2.48	2.66	3.84
datum (date)	4.80	2.99	4.28
zus (sister)	5.15	2.63	3.63
LI-LF-NC			
offer (sacrifice)	3.81	2.97	1.16
nederlaag (defeat)	2.72	2.75	1.16
schepping (creation)	3.36	2.94	1.24
koorts (fever)	4.36	2.96	1.08
reuk (smell)	2.96	2.36	1.42
zwakte (weakness)	2.36	2.46	1.36
hartstocht (passion)	3.84	2.87	1.20
grap (joke)	3.80	3.01	1.32
herstel (recovery)	2.80	2.99	1.24
wraak (revenge)	3.12	2.88	1.24
gunst (favour)	2.36	2.98	1.21
noodlot (fate)	1.69	2.71	1.16
misdaad (crime)	3.69	2.96	1.44
becht (confession)	2.69	2.34	1.08

Dutch set	*Imageability Score*	*Log Frequency*	*Cognate Status Score*
deugd (virtue)	1.31	2.84	1.24
hertog (duke)	3.92	3.01	1.11
uitstel (delay)	1.58	2.64	1.28
boosheid (anger)	3.38	2.56	1.26
English set			
HI-HF-C			
steen (stone)	6.65	3.72	5.00
vriend (friend)	6.11	4.18	5.79
fruit (fruit)	6.61	3.47	6.63
bad (bath)	6.71	3.41	5.84
zon (sun)	6.65	3.80	5.32
zomer (summer)	5.75	3.70	5.21
straat (street)	6.58	4.11	5.37
schouder (shoulder)	6.50	3.73	5.84
koffie (coffee)	6.82	3.58	6.16
meester (master)	5.61	3.47	5.37
muziek (music)	6.14	3.72	5.63
huis (house)	6.96	4.47	5.16
boek (book)	6.54	4.26	6.58
rivier (river)	6.86	3.71	5.89
hart (heart)	6.08	3.81	6.05
dokter (doctor)	6.46	3.86	6.37
dochter (daughter)	6.64	3.71	5.47
veld (field)	6.29	3.90	4.79
HI-HF-NC			
ziekenhuis (hospital)	6.75	3.70	1.24
oorlog (war)	5.75	4.16	1.26
jurk (dress)	6.69	3.84	1.37
weg (road)	6.65	3.99	1.21
lichaam (body)	6.62	4.17	1.16
muur (wall)	6.42	3.95	1.21
vlees (meat)	6.56	3.48	1.47
auto (car)	6.77	4.15	1.37
verf (paint)	5.92	3.66	1.32
horloge (watch)	6.75	4.07	1.11
golf (wave)	6.54	3.68	1.44
jas (coat)	6.31	3.45	1.26
winkel (shop)	6.68	3.79	1.26
stad (town)	6.04	3.95	1.53
boom (tree)	6.46	3.89	1.26

English set	*Imageability Score*	*Log Frequency*	*Cognate Status Score*
huid (skin)	6.71	3.63	1.47
boerderij (farm)	6.62	3.69	1.21
leger (army)	6.00	3.67	1.37
HI-LF-C			
pen (pen)	6.85	3.07	7.00
kalf (calf)	6.38	2.62	6.11
parel (pearl)	6.57	2.78	5.00
dief (thief)	6.11	2.57	5.79
neger (negro)	6.46	2.98	4.95
perzik (peach)	6.46	2.41	3.20
bakker (baker)	6.46	2.80	5.58
peer (pear)	6.04	2.42	6.32
lam (lamb)	6.50	3.02	5.79
rijst (rice)	6.57	3.03	5.11
kroon (crown)	6.11	3.09	5.11
viool (violin)	6.72	2.33	4.37
slaaf (slave)	6.12	3.15	5.63
katoen (cotton)	5.92	3.04	4.74
sok (sock)	6.38	2.88	6.42
vlam (flame)	6.57	3.07	5.11
piraat (pirate)	6.46	2.33	5.95
bijbel (bible)	6.57	2.91	5.95
HI-LF-NC			
aap (monkey)	6.75	2.86	1.05
druif (grape)	5.96	2.60	1.42
kaak (jaw)	5.75	2.97	1.56
citroen (lemon)	6.86	2.79	1.47
kussen (pillow)	6.86	2.91	1.21
schildpad (turtle)	6.62	2.17	1.16
slager (butcher)	6.08	2.55	1.32
kogel (bullet)	6.12	2.97	1.16
fiets (bike)	7.00	2.64	1.26
mouw (sleeve)	6.00	2.86	1.16
lepel (spoon)	6.75	2.87	1.42
handtekening (signature)	6.14	2.61	1.16
snor (moustache)	6.93	2.86	1.37
wortel (carrot)	6.50	2.50	1.37
vliegtuig (airplane)	6.86	2.37	1.37
vlinder (butterfly)	6.73	2.59	1.21
duif (pigeon)	6.58	2.60	1.21
paraplu (umbrella)	6.77	2.72	1.32
LI-HF-C			
wil (will)	3.08	5.31	6.58

English set	Imageability Score	Log Frequency	Cognate Status Score
waarde (worth)	2.57	3.63	4.08
einde (end)	4.35	4.38	4.63
gevoel (feeling)	3.64	3.93	2.76
ding (thing)	3.46	4.68	6.12
vorm (form)	4.42	4.25	6.37
natuur (nature)	4.69	3.88	5.37
manier (manner)	3.00	3.59	5.05
mijl (mile)	3.77	3.84	6.21
methode (method)	2.89	3.76	6.00
invloed (influence)	2.46	3.66	4.16
reden (reason)	2.79	4.14	4.08
resultaat (result)	3.46	4.04	5.08
prijs (price)	4.71	3.81	5.79
publiek (public)	4.69	4.15	5.84
kans (chance)	2.77	3.85	4.76
actie (action)	4.43	4.26	5.53
cultuur (culture)	3.46	3.52	6.16
LI-HF-NC			
taal (language)	3.35	4.09	1.24
inhoud (contents)	3.36	3.48	1.16
belofte (promise)	2.96	3.66	1.56
richting (direction)	4.07	3.67	1.60
verschil (difference)	3.04	3.78	1.32
poging (attempt)	2.54	3.87	1.12
plicht (duty)	2.62	3.54	1.16
waarheid (truth)	2.88	3.73	1.20
toekomst (future)	2.92	3.90	1.44
ervaring (experience)	3.25	4.01	1.80
wetenschap (science)	3.43	3.64	1.16
geheugen (memory)	2.73	3.55	1.24
oplossing (solution)	3.00	3.47	1.24
zorg (care)	3.39	3.94	1.47
eeuw (century)	3.08	3.95	1.28
verandering (change)	2.69	4.66	1.48
mogelijkheid (possibility)	3.00	3.54	1.44
meerderheid (majority)	3.73	3.48	2.24
LI-LF-C			
hoogte (height)	5.00	3.27	4.63
heg (hedge)	4.75	2.96	4.00
draad (thread)	3.88	2.92	3.32
zweet (sweat)	5.27	3.24	4.95
stank (stench)	2.65	2.25	3.52
zilver (silver)	5.42	3.33	6.53

English set	*Imageability Score*	*Log Frequency*	*Cognate Status Score*
verraad (betrayal)	2.58	2.51	3.15
bod (bid)	2.39	2.84	4.48
groenteboer (greengrocer)	4.92	1.72	3.00
klimaat (climate)	4.36	3.12	6.00
warmte (warmth)	3.50	2.98	6.20
ritme (rhythm)	3.85	2.98	5.21
walvis (whale)	5.22	2.62	3.60
moordenaar (murderer)	5.20	2.70	4.64
daad (deed)	3.07	2.68	5.28
geluk (luck)	2.62	3.26	4.68
domein (domain)	4.39	2.66	6.24
raadsel (riddle)	3.93	2.47	3.48
LI-LF-NC			
dwang (compulsion)	2.04	2.45	1.08
erfenis (inheritance)	2.81	2.59	1.52
geduld (patience)	3.39	2.87	1.17
eerlijkheid (honesty)	2.96	2.55	1.32
lafaard (coward)	3.73	2.45	1.44
spijt (regret)	2.96	3.09	1.00
wrok (grudge)	2.65	2.25	1.42
ruzie (quarrel)	4.04	2.98	1.32
onschuld (innocence)	2.88	2.85	1.64
geweten (conscience)	2.68	3.01	1.24
verdrag (treaty)	2.89	2.85	1.32
vloek (curse)	3.35	2.95	1.16
misbruik (abuse)	3.50	2.97	1.12
nadeel (disadvantage)	2.71	2.83	1.12
schatting (estimation)	2.21	1.86	1.12
huurder (tenant)	4.25	3.08	1.37
wreedheid (cruelty)	3.85	2.89	1.36
wanhoop (despair)	2.82	3.11	1.16