Language Learning 48:1, March 1998, pp. 31–69

Effects of Learning Method and Word Type on Acquiring Vocabulary in an Unfamiliar Language

Lorella Lotto
University of Padova

Annette M. B. de Groot
University of Amsterdam

This study examined the roles of learning method, word frequency, and cognate status in the learning of 80 Italian words by 56 adult Dutch learners previously unfamiliar with Italian. We contrasted 2 learning methods: word learning, where the Italian word was presented with its translation in Dutch, and picture learning, where it was presented with a picture depicting its referent. At test, either pictures or the Dutch words constituted the cues for recall of the Italian words. Recall was tested twice: once after 3 learning trials per stimulus, and a second time after an additional 3 learning trials. Two measures served as dependent variables: retrieval times and recall scores. The results show (a) that word learning resulted in better

Lorella Lotto, Department of General Psychology, University of Padova, Italy; Annette M. B. de Groot, Department of Psychology, University of Amsterdam, The Netherlands.

We would like to thank Volkert Balk for programming the experiments.

Correspondence concerning this article may be addressed to Lorella Lotto, Dipartimento di Psicologia Generale, Università di Padova, via Venezia 8, 35131 Padova, Italy, or to Annette M. B. de Groot, Department of Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands.

performance than picture learning; (b) that performance was better when the study and test conditions were congruent than when they were incongruent; and (c) that cognates and high-frequency words were easier to learn than noncognates and low-frequency words. Particularly noteworthy is that after 6 learning trials performance had not yet become independent of learning method. We discuss the implications of these results for bilingual memory representation and for sequencing curricula for foreign-language learning.

What is the most effective method to acquire vocabulary in a new language, and what types of words are the easiest to learn in the new language? The goal of this study was to obtain partial answers to these two questions by contrasting two acquisition methods and manipulating two word characteristics, word frequency and cognate status, in a second language (L2) word learning study. The answers inform theories on how multilinguals mentally represent and process their multiple languages and, at a more practical level, views on how to construct foreign-language learning programs. The view that a relationship may exist between the way a new language is learned and bilingual memory organization and processing goes back to at least 1954, when Ervin and Osgood suggested that different acquisition contexts lead to different bilingual memory structures (Ervin & Osgood, 1954). A relation between word type and bilingual-memory representation is suggested by an increasing number of studies that show word-type effects in a number of bilingual word-processing tasks, such as cross-language priming (e.g., Cristoffanini, Kirsner, & Milech, 1986; De Groot & Nas, 1991; Jin, 1990), word translation (e.g., De Groot, 1992; De Groot, Dannenburg, & Van Hell, 1994; Kroll & Stewart, 1994; Sánchez-Casas, Davis, & García-Albea, 1992), and cross-language word association (Kolers, 1963; Taylor, 1976; Van Hell & De Groot, 1997). Different bilingual memory representations for different types of words may explain these effects of word type (De Groot, 1993). The second of the above claims, that

knowing what vocabulary acquisition method is the most effective and what words are easiest to learn may inform the construction of foreign-language (FL) curricula, goes without saying.

Chen (1990), prompted by the results of an earlier study by Chen and Leung (1989), also investigated the role of learning method in FL vocabulary acquisition and bilingual memory representation. Chen and Leung tested three models of bilingual memory representation that were (re)introduced in the bilingual literature by Potter, So, Von Eckardt, and Feldman (1984): the "word-association" model, the "concept-mediation" model, and the "intermediate" or "developmental" model (see Figure 1). The word-association and concept-mediation models (the top two models in Figure 1) both postulate two levels of representation, one lexical (storing the forms of words) and one conceptual (storing the words' meanings). Also, they both postulate that words in two different languages are represented in two separate lexicons, whereas concepts are thought to be stored in an amodal, non-linguistic form, in a single system shared between the two languages. However, the two models differ with respect to the connections they assume to exist between the two levels of representation. The word-association model proposes that the form representations of second-language (L2) words are connected directly with the corresponding first-language (L1) words' form representations, and that only the form representations of L1 words are connected directly to the representations of the associated concepts in the conceptual system. This means that L2 words only gain access to concepts indirectly, through L1 mediation (cf. Weinreich's, 1953/1974, "subordinative" system). In contrast, the concept-mediation model assumes direct connections between the shared representations in the conceptual system and the corresponding representations in each of the two lexical stores. As a consequence, not only L1 words but also L2 words gain access to conceptual memory directly (cf. Weinreich's, 1953/1974, "compound" system). The third model, the developmental model, combines the two proposals: It assumes that the word-association model (the topmost model in Figure 1) holds for bilinguals with relatively low

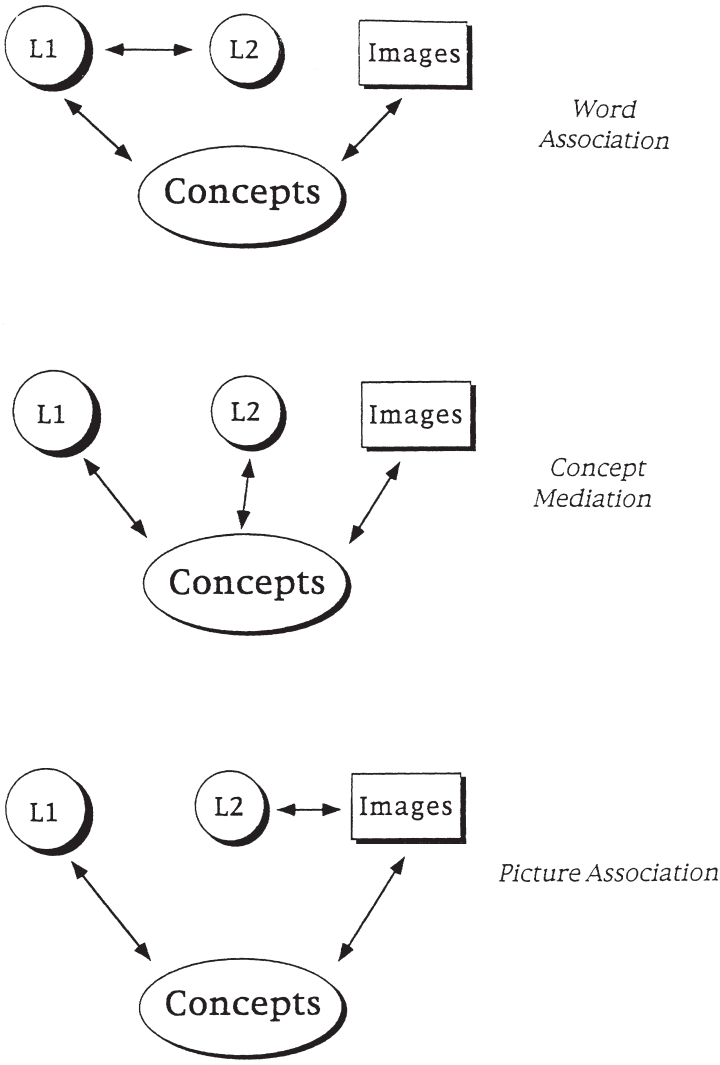

Figure 1. The word association and concept mediation models (adapted from Potter et al., 1984) and the picture association model (adapted from Chen, 1990). In the word association model the representations of second-language words and the corresponding first-language words are connected. In the concept-mediation model words in both languages directly access shared representations in conceptual memory. In the picture-association model the representations of second-language words and the corresponding images are connected.

L2 proficiency, whereas the concept-mediation organization (the middle model in Figure 1) holds for more proficient bilinguals.

Following Potter et al. (1984), Chen and Leung (1989) tested these views by comparing participants with different degrees of L2 proficiency (and of different ages) on a number of tasks. The most crucial comparison involved the participants' performance on picture naming in L2 on the one hand, and translating L1 words into L2, on the other hand. The word-association model predicts that translating L1 into L2 takes less time than picture naming in L2 because the route to the response is shorter in the first task than in the second. According to this model, translation from L1 to L2 involves tracing the link between the corresponding L1 and L2 representations in the lexical store, bypassing conceptual memory. However, picture naming in L2 comes about by tracing the longer route from the "images" store (see Figure 1), via conceptual memory and the L1 lexical store to the L2 lexical store. In contrast, the concept-mediation model predicts equally long response times (RTs) for the two tasks, because the route to the response is equally long in both cases, both involving the conceptual system (see Figure 1).

The data patterns of the *adult* bilinguals in Chen and Leung (1989) supported the above developmental model: At low levels of L2 proficiency, translating from L1 to L2 was faster than picture-naming in L2, whereas at relatively high levels of L2 proficiency the two tasks took equally long. Similar support for the developmental model has been obtained by Kroll and Curley (1988) and, using a different technique, by De Groot and Hoeks (1995), whereas the data in Potter et al.'s seminal study (1984) had supported the concept-mediation model for both of the proficiency groups tested (see also De Groot & Poot, 1997). As Chen and Leung and Kroll and Curley suggested, the difference between their patterns of results and Potter et al.'s (1984) study may have been due to the relatively high L2 proficiency of the "low-proficient" bilinguals in the latter study. In other words, it is possible that Potter et al.'s low-proficient participants had already passed the word-association stage at the time of testing.

One particular finding by Chen and Leung (1989) is of special interest to us here, namely, that low-proficient bilingual *children* showed a pattern of data that fitted neither the word-association model nor the concept-mediation model: They were in fact faster at picture-naming in L2 than at translating L1 words in L2. Chen and Leung explained this result by assuming that child bilinguals' memories do not contain direct links between the lexical representations of L1 words and their translations in L2 (as in the word-association model), but rather direct links between the L2 lexical representations and the corresponding image representations instead. These links are exploited when pictures have to be named in L2. The corresponding model, the "picture-association" model, is depicted in the bottom part of Figure 1. Given such a picture-association memory organization, L2 words gain access to concepts indirectly, via the corresponding image representations. In other words, Chen and Leung proposed that, like adult beginners, child beginners use a "mediator" when accessing the meaning of an L2 word; however, the children's mediator is not the corresponding L1 word, but a pictorial representation of it.

At first sight, age may appear to be the critical factor underlying the differential results obtained for nonfluent adult and child bilinguals. However, as pointed out by Chen (1990, 1992), learning strategy is likely to differ between adult and child L2 learners. Whereas adults are typically taught a new language with the aid of L1 words, children are usually taught a new language with the help of concrete media, such as pictures or even real objects (Chen, 1992). If so, learning strategy rather than age may underlie the different mediation patterns obtained for children and adults at an initial stage of L2 acquisition. Chen (1990) reported a study designed to resolve this indeterminacy. Its critical feature was that the participants were now matched on age, while learning strategy was experimentally manipulated.

Whereas Chen and Leung (1989), Kroll and Curley (1988), and Potter et al. (1984) had all tested participants who already possessed a certain degree of proficiency in the L2, Chen (1990, Experiment 3) set out with participants with absolutely no

knowledge of the L2 under consideration: Cantonese-speaking Chinese undergraduates learned French words from scratch. The participants were split up in two groups, a word-learning group and a picture-learning group. During the study phase the word-learning group received stimuli that each consisted of a French word and its translation in Chinese (French-Chinese pairs), whereas the picture-learning group received stimuli each consisting of a French word and the corresponding picture (French-picture pairs). At test the participants were asked to name pictures in L2 and to translate words from L1 to L2. To trace the development of the patterns of processing, and, thus, of the memory representations, the experiment repeated the sequence of first study then test a number of times. Chen found that in an early stage of learning each group performed better in response to the kind of stimulus with which the French word had been associated during study: The word-learning group was faster in translating from L1 to L2 than in naming pictures in L2, whereas the picture-learning group showed the opposite pattern. However, the test session that followed more extensive training showed that both groups were equally fast in translating from L1 to L2 and in picture naming in L2, a data pattern that supports the concept-mediation model for both groups. Chen concluded that both proficiency and learning strategy play a role in bilinguals' lexical processing.

Given the results of the studies by Chen and Leung (1989) and Kroll and Curley (1988), which both suggested that bilinguals only start to conceptually mediate their L2 after about two years of L2 classroom experience, a surprising aspect of Chen's (1990) study was that the participants already showed a concept-mediation pattern after about 30 minutes of learning an unfamiliar language. The results of the combined studies suggest that it is easier to acquire a concept-mediation pattern in an experimental setting than in a natural classroom setting. Chen (1990, p. 287) wondered why this would be so, and suggested that it might be due to the relatively small number of words to be learned in an experimental setting. In his study only 20 French

words had to be learned, the same 20 in all learning sessions. In the present study we investigated the role of, among other variables, learning method in a study that involved the acquisition of considerably more than 20 words: The participants, with Dutch as their L1, attempted to learn the L2 Italian equivalents of 80 Dutch words in an experimental setting. An interesting question is whether under these circumstances the word- and/or picture-mediation patterns persist longer.

In addition to studying the role of learning method, the present investigation sought to study the role of word type in L2 vocabulary acquisition. The stimulus words were manipulated on two dimensions, word frequency and "cognate status." The first of these variables concerns differences between words in terms of how often they are encountered in language comprehension and used in language production. The second variable involves differences between words in terms of how similar they are in form—phonology and/or orthography—to their translation equivalents in a target L2. Words with similarly formed translation equivalents in the L2 are called "cognates"; words with dissimilar translations are called "noncognates." Both word frequency and cognate status have been investigated in several studies that focused on the organization of bilingual memory in people who, in addition to their L1, already possess at least some (and often a lot of) knowledge of the target L2. Both variables have shown robust effects (on both response times and errors) in these studies (see, e.g., Cristoffanini et al., 1986; De Groot & Nas, 1991; Sánchez-Casas et al., 1992; Taylor, 1976; Van Hell & De Groot, 1997, for effects of cognate status; De Groot, 1992, for effects of word frequency). These effects have typically been interpreted in terms of differences between the bilingual memory representations of different types of words. In contrast to these "representation" studies (that investigate the structure of bilingual memory representations that already existed prior to the experiment), apparently few if any acquisition studies have manipulated word frequency and/or cognate status and have *at the same time* used retrieval time as one of the measures of learning. The typical

acquisition study restricts itself to percentage correct as the measure of learning (e.g., Ellis & Beaton, 1993a). Yet, fluent use of an L2 not only requires that the L2 words be known, but also that they can be accessed and retrieved rapidly from memory. If not, working memory will be continuously overloaded and L2 processing will suffer many breakdowns. The present study looks at learning both in terms of access and retrieval time and in terms of recall scores.

An obvious expectation for the cognate status variable is that performance will be better for cognates than for noncognates. One basis for this assumption is that the form shared between the two translation-equivalent terms will constitute a strong retrieval cue the moment one of the two terms is presented for translation into the other term. Some authors have suggested that cognates share a representation in bilingual memory whereas noncognates are represented separately from one another (e.g., Kirsner, Lalor, & Hird, 1993; Sánchez-Casas et al., 1992). This view suggests a second source of a privileged status for cognates during learning: Setting up a new memory entry may be more demanding than inserting a new but similar element into an already existing one. A third cause of more successful performance for cognates than for noncognates could be that when a cognate has to be learned, in a way *less* has to be learned than when a noncognate has to be learned, due to the partial form overlap between the translation equivalents.

The most salient relation within cognate pairs that noncognate pairs lack is their similarity of form. This form similarity, effective either during study, or at test, or on both occasions, is the basis of our prediction that cognates will turn out easier to learn than noncognates. In view of this it will be particularly interesting to see whether an effect of cognate status also materializes when during study the L2 word is associated with a picture instead of its word equivalent in L1, or when at test a picture instead of the L1 word is presented as the cue to produce the L2 word. After all, the picture does not share any form similarity with the L2 name for the picture. A comparison of the effect of cognate status

obtained in the different learning and test conditions of this study (see the Procedure section) may thus provide insight into the sources of the effect, and it may clarify the processing operations that occur during study and test.

An effect of word frequency may be expected for a number of reasons. Numerous studies have shown lexical access and word retrieval to come about faster for high-frequency words than for low-frequency words (Balota & Chumbley, 1984, 1985). When a high-frequency L1 word is presented at test, the first step in the complete retrieval route, accessing the stimulus word's representation in memory, will therefore be fast as compared to the case of a low-frequency word. This benefit in access should also show in the overall translation response time. Research has also shown word frequency to influence picture naming in L1. Perhaps name retrieval is faster for high-frequency words (see Gregg, 1976, for a discussion); however, access to the meaning representation of the picture may also be faster for a picture representing a common object than for a picture depicting an uncommon one. For this reason one might expect that picture naming in L2 will also be influenced by word frequency.

But equally as interesting as its effects on response times are whatever effects of frequency that might turn up in the recall scores. As pointed out by Ellis and Beaton (1993a), close analogues to L2 vocabulary-acquisition experiments are the paired-associate learning experiments in memory research. These experiments have often produced an effect of word frequency, with common words producing higher recall scores than uncommon words (see Gregg, 1976, for a discussion). We may thus expect here that more Italian words will be learned for high-frequency Dutch words and the corresponding pictures than for low-frequency Dutch words and the associated pictures. Such a finding would be of particular interest, given that we presented the Italian names for the high- and low-frequency words equally often in the experiment. Any effect of frequency on the recall scores could therefore not be attributed to differential presentation frequency of the Italian

words. Instead, differences in concept familiarity might underlie such an effect.

Method

Pretests

Prior to the actual learning experiment we performed two computerized pretests. The first involved a picture-naming-in-L1 task that was to produce picture-agreement norms. This picture-naming pretest provided a means to know that a picture in the picture-learning condition of the main experiment would give rise to the L1 word presented on the corresponding trial in the word-learning condition. The second pretest involved the assessment of the cognate relation between Dutch words and their translations in Italian. The 20 participants tested in these pretests were all different from those tested in the main experiment, but were drawn from the same population. In the picture-naming pretest, we presented 224 pictures. We asked the participants to name the pictures in their L1, Dutch. We registered response times (measured from the onset of the picture) and errors. In the cognate-rating study we presented the corresponding 224 Dutch-Italian word pairs and asked the participants to rate each pair on a 7-point scale on how similar the words in that pair were.

We calculated a mean similarity (cognate-) rating and the corresponding standard deviation for each Dutch-Italian word pair in the cognate-rating study. For each picture we also calculated the mean RT and standard deviation, collapsed across all of the participants who had provided the same response to that picture (which at the same time was the intended response). We considered good pictures those to which at least 15 participants (75%) gave the same response. However, the picture-agreement scores for the pictures selected for presentation in the main experiment were considerably higher. Only 2 of the ultimately

selected pictures had an agreement score of 75%; their average picture-agreement score was 94.5%.

From the total set of 224 word pairs we selected 4 groups of 20 word pairs each: high-frequency cognate pairs, high-frequency noncognate pairs, low-frequency cognate pairs, and low-frequency noncognate pairs. Pairs with mean cognate ratings higher than 4 we regarded as cognate pairs. Noncognate pairs had mean ratings lower than 4. We regarded pairs of which the Dutch term had a frequency of occurrence of 11 or more in a familiar Dutch word-frequency count (Uit den Bogaart, 1975; corpus size: 620,000 words) high-frequency pairs; pairs in which the Dutch terms had a frequency of occurrence of 10 or less we regarded as low-frequency pairs.[1] Additional selection constraints were: that the cognates and noncognates within the 2 frequency conditions matched each other on frequency; that the number of letters contained by the Italian words was statistically equally large in the 4 groups of word pairs; and that the picture-agreement scores were equally large across the 4 word groups. Table 1 shows the

Table 1

Average Picture-Naming RT (in msec), Cognate Rating, Word Frequency, Length of the Italian Words, and Picture-Agreement Scores (in Percentages) for the Four Selected Stimulus Groups

	Cognates		Noncognates	
High frequency	*M*	*SD*	*M*	*SD*
Picture-naming RT	896	165.3	905	169.4
Cognate rating	5.54	0.83	1.95	0.39
Word frequency	47.70	70.26	46.45	62.79
Italian length	6.30	1.75	6.05	1.32
Picture agreement	96.0	7.88	93.5	6.71
Low frequency				
Picture-naming RT	993	178.7	1005	78.9
Cognate rating	5.76	0.83	1.82	0.58
Word frequency	3.20	3.38	3.20	3.33
Italian length	6.80	1.99	6.90	1.48
Picture agreement	93.5	7.09	95.0	5.38

means of the relevant stimulus characteristics for the 4 groups of selected word pairs.

Our aim was to have: (a) no difference on RT in the picture-naming test between the 2 high-frequency groups (cognates vs. noncognates), nor between the 2 low-frequency groups (again, cognates vs. noncognates); (b) no difference between the 2 cognate groups nor between the 2 noncognate groups in their ratings on the cognate-rating test; (c) equal length of the Italian words in the 4 word groups; (d) equal picture-agreement scores in the 4 word groups; and (e) no difference between cognates and noncognates in frequency of occurrence within the 2 frequency groups. Analyses of variance on the data showed that nearly all of these criteria were met. The only exception was that the second criterion was not met perfectly: The cognate ratings turned out to be slightly lower for high-frequency cognates than for low-frequency cognates ($p < .05$). However, as can be seen in Table 1, the direction of the effect was such that it would work against an effect of frequency (the low-frequency words had the higher cognate ratings). So any effect of word frequency to emerge would not be attributable to this small confounding.[2]

Main Experiment

Participants

Sixty-four first-year psychology students from the University of Amsterdam, with Dutch as their L1, participated. They received course credit for participation. We excluded the data from 8 of them, each with a mean test accuracy lower than 60% (averaged across the 2 sessions; see below), from the statistical analyses; because of their high error rates, these participants left cells in the design with not a single RT value. All the remaining 56 participants claimed not to have any prior knowledge of Italian. All of them, however, had considerable knowledge about English, and all had received training in other FLs (often French and/or

German, occasionally Latin) as well. In other words, at least some of the participants may have had some indirect knowledge of Italian, via one or more related languages (e.g., French and Latin).

Materials and Apparatus

The stimuli were based on the 80 pictures and the corresponding words selected from the 2 pretests. The 80 word stimuli consisted of the 4 different groups of words described earlier: high-frequency cognates, high-frequency noncognates, low-frequency cognates and low-frequency noncognates; 20 words per group.[3] The complete set of stimuli, together with, for every individual stimulus, the word-frequency value (Uit den Boogaart, 1975), the cognate rating, and the picture-agreement score, are reported in the Appendix. We created 2 conditions of 80 stimuli each from these materials. In the first condition (the word-learning condition) each stimulus consisted of a Dutch word and the corresponding Italian word. In the second condition (the picture-learning condition) each stimulus consisted of a picture and the corresponding Italian word.

The experiment was run on an Apple Macintosh computer to which a voice-operated switch was connected. Pictures and words appeared against a light grey background in the centre of the computer screen. Stimulus duration, inter-stimulus interval, duration of the fixation stimulus preceding a stimulus (see Procedure), and RT recording were all controlled by an Authorware program. Pictures subtended a visual angle of about 3.5° horizontally by 3.5° vertically, and words subtended a visual angle of 0.6° per letter horizontally by 0.7° vertically.

Procedure

We randomly divided the participants into 4 groups of 14. Each participant was run individually in 2 experimental sessions. Each experimental session included a learning phase and a test phase. On each trial in the test phase, the participants received

a stimulus (a picture or a Dutch word) that, in the same or in a different form (see below), had appeared in association with the corresponding Italian word in the learning phase. The participants' task was to come up with the corresponding Italian word (that is, to name the picture in Italian, or to translate the Dutch word into Italian). In other words, the experiment tested productive, not receptive, language learning. We made this choice because the critical comparison in Chen's (1990) study, the study that we set out to extend here, concerned picture naming in the newly learned L2 versus translating from L1 into the L2. Furthermore, we would have no way to test receptive learning of the FL, here Italian, in a picture condition analogous to the receptive condition with Italian words as test stimuli and Dutch words as the intended responses: In a condition with pictures as test stimuli to be named in Dutch, the FL would not be implicated at all.

In the learning phase, every participant in each of the 4 groups first received 3 times over the same set of 80 stimuli. For 2 of the 4 groups of participants, these stimuli consisted of the 80 Dutch-Italian word pairs selected from the pretests. These 2 groups constituted the word-learning condition. The two remaining groups received the corresponding 80 picture-word stimuli, each consisting of a picture and its name in Italian. These 2 groups constituted the picture-learning condition. We randomized the entire set of stimuli across participants, and within participants across the 3 presentation rounds, so that each participant in each presentation round received the stimuli in a unique presentation order. This procedure ensured that possible order effects (due to, e.g., fatigue) would affect all stimuli to the same extent. We presented the stimuli at a rate of 8 sec each; they were preceded by a fixation stimulus that appeared on the centre of the screen for 1 sec. Following the fixation stimulus and prior to the presentation of the next stimulus pair, the screen was empty for 100 msec.

After all the stimuli had appeared thrice, the test phase started. In the test phase, we presented one of the 2 word-learning groups and one of the 2 picture-learning groups with stimuli congruent with the training they had received in the learning

phase. Hence, the participants in the congruent word-learning group received (Dutch) words as stimuli (Group W-W), and those in the congruent picture-learning group received pictures as stimuli (Group P-P). In contrast, the remaining 2 groups received stimuli incongruent with the materials they had been trained on: The incongruent word-learning group received pictures as stimuli (Group W-P), whereas the incongruent picture-learning group received Dutch words (Group P-W). All participants in all 4 groups had to produce the equivalent Italian word in response to each of the stimuli, picture or Dutch word. We encouraged them to produce their responses as quickly and as accurately as possible. Responses were registered by a microphone that activated a voice-operated switch. We measured reaction time from the onset of the stimulus. We instructed the participants to type in their response immediately after having produced it. Subsequently, they had to press the RETURN key, immediately after which the next trial started. Participants sat facing the screen at a comfortable viewing distance in a sound attenuated room. During the test phase, the experimenter sat to the right of the participant, monitoring the workings of the voice-switch. The experimenter noted down failures of the voice-switch to react to a participant's response. Prior to the learning phase the participant read the instructions, which appeared on the screen in Dutch. In them, we told the participant about the exact nature of the stimuli presented during learning and test. Prior to the learning phase, we presented the participant with 3 stimuli (word-word or picture-word, depending on the condition) for practice. The learning phase lasted about 45 min and the test phase about 15 min. All participants came back into the laboratory the next day; we tested them once more in exactly the same condition they had participated in the day before, with again 3 rounds of learning followed by a test phase. So in all there were 6 learning trials and 2 test trials per word.[4]

Results

Response Times

To reduce variability, we considered all RTs longer than 5,000 msec as outliers and excluded them from the analyses. In addition, we excluded from subsequent analyses scores greater than 3 standard deviations from the mean for that participant in each session. With these procedures, we excluded less than 2% of the RTs from the final analyses.

We analysed the RTs for the remaining correct responses using analysis of variance (ANOVA). The factors included were congruency (between learning and test; congruent vs. incongruent), learning method (word learning vs. picture learning), session (first vs. second), cognate status (cognates vs. noncognates), and frequency (high-frequency words vs. low-frequency words). The ANOVA by participants had congruency and learning method as between-subjects factors, and session, cognate status, and frequency as within-subjects factors. The ANOVA by items used congruency, learning method, cognate status, and frequency as between-items factors, and session as a within-item factor.

Four of the main effects were significant both by participants and by items (*F1* for the analysis by participants and *F2* for the analysis by items). These were congruency, $F1\ (1,52) = 7.96$, $p < .01$, and $F2\ (1,304) = 60.80$, $p < .0001$; session, $F1\ (1,52) = 378.24$, $p < .0001$, and $F2\ (1,304) = 753.05$, $p < .0001$; cognate status, $F1\ (1,52) = 292.09$, $p < .0001$, and $F2\ (1,304) = 195.87$, $p < .0001$; and frequency, $F1\ (1,52) = 35.83$, $p < .0001$, and $F2\ (1,304) = 8.90$, $p < .01$. Participants in the congruent condition were faster than those in the incongruent condition (1409 msec vs. 1605 msec, respectively); they were faster in the second experimental session than in the first (Session 1, 1755 msec; Session 2, 1259 msec); participants responded faster to cognates than noncognates (1298 msec vs. 1715 msec); and high-frequency stimuli elicited a response in less time than low-frequency stimuli (1457 msec vs. 1557 msec). The fifth variable, learning method, was statistically

significant by items but only marginally so by participants, *F1* (1,52) = 3.60, $p = .06$, and *F2* (1,304) = 16.39, $p < .0001$. The word-learning condition produced a shorter RT than the picture-learning condition (1441 msec vs. 1573 msec).

These main effects were qualified by a number of 2-way interactions that all involved the cognate status variable. Generally, a particular effect was always smaller for cognates than for noncognates: The interaction between cognate status and session was significant both by participants and by items, *F1* (1,52) = 20.43, $p < .0001$, and *F2* (1,304) = 16.81, $p < .0001$. It showed that the effect of session was larger for noncognates (575 msec) than for cognates (418 msec). Pairwise comparisons showed that the effect of session was significant for both cognates and noncognates. Table 2 reports the interaction means.

The interaction between cognate status and congruency was statistically significant by items and marginally significant by participants, *F1* (1,52) = 3.80, $p = .06$, and *F2* (1,304) = 5.85, $p < .05$. The effect of congruency was larger for noncognates (243 msec) than for cognates (148 msec). Pairwise comparisons showed that the effect of congruency was significant for both cognates and

Table 2

Mean Response Times (in msec) for All Conditions Formed by the Interactions Between Cognate Status and the Remaining Variables

	Cognates	Non-cognates		Cognates	Non-cognates
Session 1	1507	2003	Incongruent	1372	1837
Session 2	1089	1428	Congruent	1224	1594
Effect	418	575	Effect	148	243

	Cognates	Non-cognates		Cognates	Non-cognates
Low-Frequent	1311	1802	Picture Learning	1336	1809
High-Frequent	1286	1628	Word Learning	1260	1622
Effect	25	174	Effect	76	187

noncognates (see Table 2 for interaction means). The interaction between cognate status and frequency was significant by participants and marginally significant by items, *F1* (1,52) = 13.74, $p < .001$, and *F2* (1,304) = 2.94, $.05 < p < .10$. It pointed out that a frequency effect occurred for noncognates (174 msec) but not for cognates (25 msec, $p > .10$). (Table 2 shows the interaction means.) Finally, the interaction between cognate status and learning method was statistically reliable, but only in the analysis by participants, *F1* (1,52) = 5.20, $p < .05$, and *F2* < 1. Again, noncognates showed the largest effect (187 msec). The 76 msec effect of learning method for cognates turned out not to be significant ($p > .10$). (Table 2 presents the interaction means.) However, the data indicate that this latter 2-way interaction needs to be qualified by a 3-way interaction between cognate status, learning method, and congruency. This interaction was marginally significant, but only by participants, *F1* (1,52) = 3.77, $p = .06$; *F2* (1,304) = 2.14, $p > .10$. The interaction means suggested that the above conclusion that no effect of learning method occurred for cognates in fact only holds for cognates in the incongruent condition (word learning: 1397 msec; picture learning: 1347 msec). In the congruent condition, a clear effect of learning method exists also for cognates (word learning: 1123 msec; picture learning: 1325 msec).

One of the higher-order interactions was statistically reliable on both the analysis by participants and that by items, namely, the interaction between learning method, congruency, and session, *F1* (1,52) = 9.11, $p < .01$; *F2* (1,304) = 9.92, $p < .01$. To pinpoint the source of this interaction, we performed 2 separate ANOVAs on the data by participants, one for Session 1, and the second for Session 2. The cell means for both sessions appear in Figure 2, organized the way Chen (1990) presented his data.

The data from Session 1 showed a significant main effect of congruency, *F1* (1,52) = 4.74, $p < .05$ (congruent: 1660 msec; incongruent: 1850 msec), a marginally significant effect of learning method, *F1* (1,52) = 2.81, $p = .10$ (word learning: 1682 msec; picture learning: 1828 msec), and a marginally significant interaction between the two variables, *F1* (1,52) = 3.20, $p = .08$. A

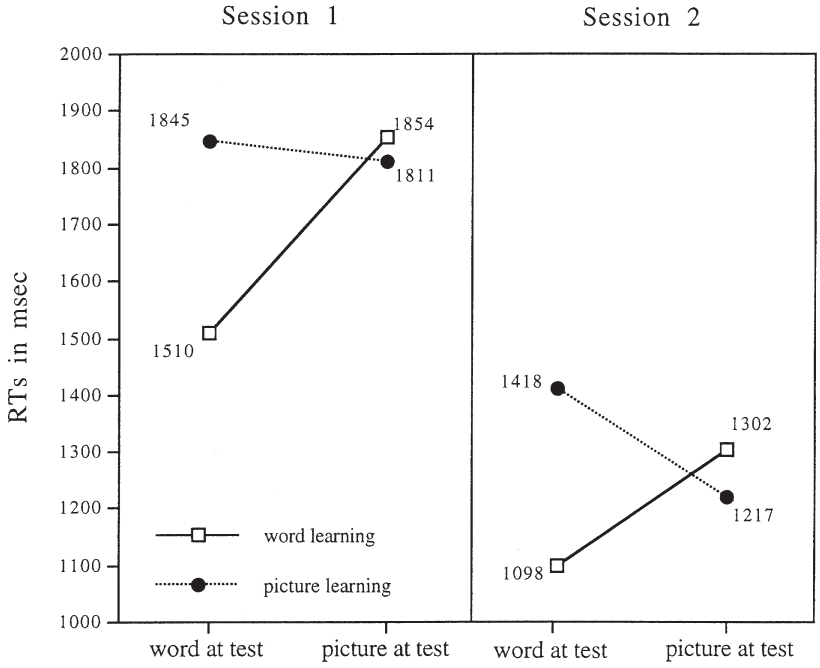

Figure 2. Mean response times for naming pictures in L2 (Italian) and translating words from L1 (Dutch) to L2 in the word-learning and the picture-learning conditions.

Newman Keuls test showed that Condition WW (word learning-word at test; congruent) produced significantly faster responses ($p < .05$) than the three remaining conditions, which did not differ from one another. The data from Session 2 also showed a significant main effect of congruency, $F1$ (1,52) = 12.11, $p < .01$ (congruent: 1157 msec; incongruent: 1360 msec). The main effect of learning method was now significant, $F1$ (1,52) = 4.08, $p < .05$ (word learning: 1200 msec; picture learning: 1317 msec). Finally, there was no trace of an interaction between the two variables ($F1 < 1$).

Recall Scores

We calculated 8 scores for each participant, one for each of the 8 cognate status × frequency × session conditions. Each indi-

vidual score reflected the number of times that particular participant had produced either an erroneous response or none at all (an "omission"). Out of the total number of trials, 4.87% resulted in an omission and 15.86% resulted in an error. In classifying responses as errors we used a very stringent criterion. After pronouncing a response, the participant typed it on the computer keyboard. We treated every typed deviation from the target response as an error, even if the typed response deviated from the target by only one letter. So, for example, we treated "sigaretto" for the target "sigaretta," or "tore" for "torre" as errors. We used this stringent criterion because of the presence of cognates among the materials. If we had used a more lenient criterion, cognates (that by definition already share many letters with the target) might produce hardly any errors. Besides, this way we could prevent ambiguities in scoring. Zooming in on the errors at a more detailed level of analysis showed that 10.79% of the 15.86% errors were small (deviating from the target word by one or two letters only), whereas 5.07% of the errors concerned larger deviations from the targets (e.g., "robinero" for "rubinetto," and "fragello" for "fragola").

We combined the errors and omissions into a single score (per participant, per condition). Similarly, for each individual item we calculated the analogous (errors + omissions) score by summing across the participants in each of the learning method × congruency conditions. On these scores, we performed the same pair of ANOVAs as on the RT data, one by participants and one by items. In what follows, we transform the resulting errors + omissions scores into their counterparts, the recall scores (in percentages); we report these recall scores throughout in the text.[5]

The analyses revealed three main effects that were statistically significant both by participants and by items. These were session, $F1\ (1,52) = 271.58, p < .0001$; $F2\ (1,304) = 730.66, p < .0001$; cognate status, $F1\ (1,52) = 142.80, p < .0001$; $F2\ (1,304) = 78.50, p < .0001$; and frequency, $F1\ (1,52) = 96.63, p < .0001$; $F2\ (1,304) = 16.95, p < .0001$. Participants had higher recall scores in the second session than in the first (Session 1: 66.85%; Session 2: 91.58%); recall was better with cognates than with noncognates (86.47% vs.

71.96%), and recall was better with high-frequency stimuli than with low-frequency stimuli (82.59% vs. 75.85%). The main effect of learning method was significant on the analysis by items, $F2$ (1,304) = 10.35, $p < .01$, but only marginally so on the analysis by participants, $F1$ (1,52) = 3.15, $p = .08$: Word learning produced higher recall scores than picture learning (81.85% vs. 76.58%). The main effect of congruency was not significant on either analysis, $F1 < 1$; $F2$ (1,304) = 1.86, $p > .10$.

The above main effects are qualified by a number of interactions. All but one were qualified by the 4-way interaction between learning method, session, cognate status, and frequency, $F1$ (1,52) = 5.78, $p < .05$; $F2$ (1,304) = 2.59, $p = .10$. The recall scores for all the cells in this interaction are presented in Table 3.

Two of the 2-way interactions involved the learning method variable, namely, the interaction between learning method and session, $F1$ (1,52) = 3.97, $p = .05$; $F2$ (1,304) = 10.69, $p < .01$, and the interaction between learning method and frequency, $F1$ (1,52)

Table 3

Mean Recall Scores (in Percentages) for All Conditions Formed by the Interaction Between Learning Method, Session, Cognate Status, and Frequency

Word-Learning

	Session 1			Session 2		
	Cognates	Noncognates	Effect	Cognates	Noncognates	Effect
HF	85.18	63.75	21.43	96.07	89.64	6.43
LF	77.32	57.68	19.64	93.57	91.61	1.96
Effect	7.86	6.07		2.50	–1.97	

Picture-Learning

	Session 1			Session 2		
	Cognates	Noncognates	Effect	Cognates	Noncognates	Effect
HF	80.36	59.82	20.54	97.32	88.57	8.75
LF	71.25	39.46	31.79	90.71	85.18	5.53
Effect	9.11	20.36		6.61	3.39	

= 20.77, $p < .0001$; *F2* (1,304) = 3.64, $p = .06$. These two interactions indicate that the effect of learning method was larger in Session 1 (8.3%) than in Session 2 (2.3%), and that the learning-method manipulation affected low-frequency words more than high-frequency words (effects of 8.4% and 2.1%, respectively). Post hoc paired comparisons on the analysis by participants showed that the effect of learning method on high-frequency words was not statistically significant. Session interacted not only with learning method but also with cognate status, *F1* (1,52) = 154.79, $p < .0001$; *F2* (1,304) = 93.33, $p < .0001$; and with frequency, *F1* (1,52) = 51.84, $p < .0001$; *F2* (1,304) = 20.15, $p < .0001$. Hence, the effects of frequency and cognate status were larger in Session 1 than in Session 2 (see Table 3). One of the 3-way interactions was significant both by participants and by items, namely, the interaction between session, cognate status, and frequency, *F1* (1,52) = 12.23, $p < .001$; *F2* (1,304) = 5.48, $p < .05$. It showed that the larger effects of frequency and cognate status in Session 1 than in Session 2 were primarily due to the relatively low recall scores of low-frequency noncognates in Session 1. The 4-way interaction between these 3 variables and learning method, already introduced above, further qualified this interaction by pointing out that in particular the picture-learning condition, not the word-learning condition, led to poor performance on low-frequency noncognates (see Table 3). Two of the remaining interactions were significant on the analysis by participants but not on the analysis by items. One of them, between learning method, cognate status and frequency, *F1* (1,52) = 5.92, $p < .05$; *F2* (1,304) = 1.19, $p > .10$, presumably resulted from the poor performance on low-frequency noncognates in the picture-learning condition. The second, between congruency, session, and frequency, *F1* (1,52) = 6.67, $p < .05$; *F2* (1,304) = 2.59, $p > .10$, was presumably due to the especially large difference between the recall scores for high-frequency and low-frequency words in Session 1 of the incongruent condition.

Discussion

We looked at the efficacy of two different learning methods in the acquisition, by adult L1 speakers of Dutch, of words of various types in an FL, Italian. We contrasted picture learning and word learning. In the picture-learning condition, we presented the Italian words in combination with pictures representing the words' referents. In the word-learning condition, we presented the Italian words in combination with their Dutch translation equivalents. At test, we presented the participants with either the pictures or with the Dutch words and required them to produce the corresponding Italian words. In other words, we tested what others (e.g., Ellis & Beaton, 1993a, 1993b) have called "productive" vocabulary learning (as opposed to "receptive" learning, where L1 is the language of the response words). For every individual participant, the conditions during learning and test were either congruent (where either words or pictures were presented both during learning and at test) or incongruent (where picture presentation during learning was followed by word presentation at test, or vice versa). There were 2 test moments, the first after 3 learning trials per stimulus pair, and the second after an additional 3 learning trials per pair. The words to be learned varied on 2 word characteristics: word frequency and cognate status.

Learning Method and Congruency

In terms of retrieval time at test, word learning led to better performance than picture learning, at least in Session 2: The congruent word-learning condition (with words at test) showed shorter RTs than the congruent picture-learning condition (with pictures at test), and the incongruent word-learning condition (with pictures at test) showed shorter RTs than the incongruent picture-learning condition (with words at test). In Session 1, however, the congruent word-learning condition produced the fastest responses, with the remaining 3 conditions, including the incongruent word-learning condition, not showing any differences

between one another (see below). The effect of learning method indicates that presentation of L1-L2 word pairs during learning provides a better opportunity for acquiring L2 words than does the presentation of picture-L2 pairs, at least for our participants. The recall data also suggest that for our participants word learning is a more effective method to acquire an L2 vocabulary than is picture learning: Word learning produced higher recall scores than picture learning (although overall the effect was only marginally significant by participants). Interactions of the learning method variable with both session and word frequency showed that the role of learning method was larger in Session 1 than in Session 2, and that it affected the learning of low-frequency words but not of high-frequency words.

That word learning turned out to be the more effective learning method is particularly interesting in view of the fact that the opposite effect, better learning with pictures or objects, has also been obtained, for instance by Wimer and Lambert (1959; see there for a number of earlier, related studies that yielded conflicting results). Wimer and Lambert's study differed in many aspects from ours. Each of these differences may have contributed to the different result. A recent study by Van Hell and Candia Mahn (1997) may suggest what caused the different outcomes. These authors compared the efficacy of the keyword method on the one hand and rote rehearsal on the other in FL learning by English L1 undergraduates at a university in the United States versus Dutch L1 undergraduates from the population we employed in this study. The American undergraduates learned the Dutch translations of English words; the Dutch undergraduates learned the Spanish translations of Dutch words. In all other respects, the experimental procedures were kept as similar as possible across the 2 experiments. Van Hell and Candia Mahn found that L2 learning by the two groups of participants was differentially affected by the learning-method manipulation: Whereas, in terms of recall scores, the American undergraduates performed equally well in the keyword and rote-rehearsal conditions, the Dutch undergraduates showed better results in the rote-rehearsal

condition. The authors suggested that differential experience in FL learning may underlie this difference in the results from the 2 groups: The Dutch participants in their study, like our participants, had generally had substantial experience in FL learning, and certainly much more than typical American undergraduates (also tested in Wimer & Lambert's study). Possibly increasing experience in FL learning involves a change in the preferred learning method. If so, Van Hell and Candia Mahn's and our studies suggest that when experienced FL learners start to learn vocabulary in yet another FL, learning will be most successful if the new vocabulary is associated with the corresponding L1 words.

A number of studies by Thomas and Wang (see Thomas & Wang, 1996, for an overview) provide further qualification of the general view that keyword learning is a more effective vocabulary-acquisition method than is rote rehearsal. These authors have shown that with their population of participants (American undergraduates) keyword learning resulted in better performance than rote rehearsal when the test phase immediately followed the learning phase. However, when retention was tested 2 days after learning, performance was better in the rote-rehearsal condition.

As to the congruency manipulation, our most salient outcome is that, unlike Chen's (1990) results, after 6 learning trials performance in terms of retrieval time had not (yet) become independent of the learning method. Chen's participants, both in the word-learning and in the picture-learning condition, produced equally long response times when presented with L1 (Chinese) words or with pictures at test, indicating that after 6 learning trials performance no longer depended on the learning method. In contrast, after 6 learning trials, learning method still determined our participants' performance: Those in the word-learning condition were faster at translating Dutch words into Italian than at naming pictures in Italian, whereas those in the picture-learning condition were faster when presented with pictures at test than when presented with Dutch words (Figure 2). These data replicate the pattern of results that Chen obtained when testing after only 3 learning trials. In terms of the models of bilingual memory

presented earlier (Figure 1), after 6 learning trials our participants in the word-learning condition performed in accordance with a word-association model, whereas those in the picture-learning condition responded in accordance with the picture-association model Chen (1990) proposed. In contrast, after the same number of learning trials per stimulus pair, Chen's participants performed according to the concept-mediation model (Figure 1). As pointed out earlier, Chen (1990) wondered why his participants already demonstrated concept-mediation processing after only 30 minutes of L2 training, whereas other studies (Chen & Leung, 1989; Kroll & Curley, 1988) had indicated that it takes about two years of classroom experience in L2 before students start to conceptually mediate their L2. He pointed out that the answer might lie in the small number of L2 words the participants in his study had to learn, 20 words only. The present study, with 80 words to learn, suggests that indeed number of L2 words to be learned may determine how soon after the onset of training concept mediation occurs.

The RT data after 3 trials in this study (see Figure 2) remain to be explained. The data for the word-learning condition again clearly support the word-association model: Translating L1 words into L2 words was faster than picture naming in L2. However, the data for the picture-learning condition do not fit the picture-association model, but rather the concept-mediation one. The conclusion, that participants after 3 learning trials in the picture-learning condition already conceptually mediate their L2, is hard to reconcile with the finding that after 6 learning trials they apparently do not do so. More plausibly, the pattern of data after the first 3 learning trials results from a floor effect in the 3 conditions most difficult for our participants: the two involving picture learning, congruent and incongruent, and the incongruent word-learning condition. As can be seen in Figure 2, response times were quite long in these 3 conditions; statistically they are equally long. An interpretation of interactions in terms of a floor effect is not uncommon (e.g., Papagno, Valentine, & Baddeley, 1991).

Cognate Status and Word Frequency

We also focused on possible effects of cognate status and frequency of the stimulus materials. As have many "representation" studies (see our introductory section), we also obtained these 2 effects, both in terms of retrieval times and in terms of recall scores: Recall scores were higher for cognates than for noncognates and the retrieval times for the former were shorter; similarly, more high-frequency than low-frequency words were recalled and the retrieval times for high-frequency words were shorter. The interactions on the RT analyses of cognate status on the one hand and frequency, congruency, and learning method on the other hand—the effect of all these other variables always being smaller for cognates than for noncognates—suggest that cognates are relatively easy to learn under all circumstances: whether frequent or infrequent, whether learning and testing occur under congruent or incongruent circumstances, and whether the word learning or picture learning method was applied. Furthermore, the interaction between cognate status and session on the reaction-time analysis—the effect of session being relatively small for cognates—indicates that cognates are learned relatively fast.

For an effect of cognate status to occur, the forms of the words did not have to be presented during learning nor at test: The effect also occurred in the picture-learning condition and when pictures were presented at test. The statistical analyses on the recall data pointed out that the effect of cognate status was equally large in all learning method × congruency conditions, indicating that pictures give rise to a cognate effect of the same size as do words. The analyses on the retrieval-time data pointed to the same conclusion. If anything, the effect tended to be larger in one of the conditions with picture presentation (Condition PW). Because the form-relation between translation-equivalent terms probably underlies the effect of cognate status on acquisition, the occurrence of the effect with pictures in learning and at test suggests that the presentation of a picture gives rise to the generation of the form of the corresponding word. In theory, the form concerned could be

either phonological, or orthographic, or both, but one of our findings in particular suggests that only the phonological form is involved, and that this form is generated for the (visually presented) words as well as for the pictures: The effect of cognate status was equally large in the 4 congruency × learning-method conditions. Had the orthographic forms played a role too, we would have obtained larger effects of cognate status in Condition WW, where both during learning and at test the orthographic forms were explicit in the (visually presented) stimuli and therefore the participants would not have to generate them. In sum, the data point out that our participants generated the phonological word forms for both the pictures and the visual word forms. This conclusion ties in nicely with a number of related studies (Baddeley, Papagno, & Vallar, 1988; Papagno et al., 1991; Van Hell & Candia Mahn, 1997) that provided strong support for the role of phonology in learning vocabulary in an L2, not only when the learning material consisted of auditory presented word pairs but also when, as here, presentation was visual. The latter finding in particular suggests that learners generate phonological word forms during learning. Under circumstances that obstruct their generation of these forms, for instance, by articulatory-suppression techniques, learners' performance deteriorates (Baddeley et al., 1988; Papagno et al., 1991).

Briefly, one particular view on monolingual and bilingual word representation can readily account for the effects of cognate status (see Introductory section for other accounts): Kirsner and his colleagues (e.g., Cristoffanini et al., 1986; Kirsner et al., 1993) have proposed that word memory is organized according to morphology: Morphologically related words share a representation in memory. This organizational principle holds not only for words belonging to one language but also for words of different languages (see also Sánchez-Casas et al., 1992). So, a French-English bilingual has one representation containing both the English words "marry," "marriage," "married" and the French words "marier," "mariage," "marié" (example taken from Kirsner et al., 1993). According to this view, learning cognates does not involve creating

a new entry in memory, but rather adding new information to an existing entry. This relatively undemanding process provides an explanation for the higher recall scores and shorter retrieval times for cognates than for noncognates.

A particularly noteworthy aspect of the frequency effects is that we did not manipulate this variable experimentally: High- and low-frequency words appeared equally often during learning (6 times). Furthermore, none of our participants had any (direct) knowledge of Italian before the experiment (though some might have had some indirect knowledge of it via other related languages). These two facts imply that the Italian forms of the high- and low-frequency words must have been equally familiar to our participants after learning. However, familiarity with the Dutch forms did differ between the high-frequency and low-frequency words because, by definition, the former had been experienced more often than the latter before the experiment. A second aspect on which the high- and low-frequency words probably differed pre-experimentally is concept familiarity. Although word frequency and concept familiarity are basically different entities, they tend to be highly correlated. Hence, either the difference in form familiarity of the Dutch words or the difference in concept familiarity underlies the frequency effects. In other words, learners apparently find it easier to assign a new name to a familiar form or concept than to a less familiar one. If concept familiarity were the relevant factor, one might look in one of two directions for an answer why this is so. First, familiar and unfamiliar concepts may provide differential opportunities to elaborate. It is well known that elaboration helps learning (e.g., Craik & Lockhart, 1972). If concept familiarity indeed supports elaboration, familiar concepts should show a learning advantage over unfamiliar concepts, as indeed they did here. Second, high- and low-frequency words differ in terms of the "density" of the information stored in the corresponding memory representations: There is some experimental support for the view that the conceptual representations of high-frequency words contain more informational elements than those of low-frequency words, though the

difference is small (De Groot, 1989). Possibly, dense representations support learning better than sparse representations do,[6] perhaps due to the different opportunities they provide to elaborate: Dense conceptual representations may enable richer elaborations than sparse representations.

Pedagogical Implications

The practical implications of the effects of cognate status and frequency obtained in this study are straightforward: If the goal of an L2 curriculum is to introduce easy words before more difficult words, the teaching of cognates and common words should precede the teaching of noncognates and uncommon words. (Of course, many L2 curricula already take these 2 variables into account. Meara, 1993, and Nation, 1993, discuss L2 acquisition programs that acknowledge the importance of cognate status and frequency, respectively.) Regarding the frequency variable, common words are indeed often introduced before uncommon words. However, the reason to do so is typically not that common words are easier to learn than uncommon words, but that a particular set of high-frequency words provides a considerably larger text coverage than an equally large set of uncommon words does. The sooner a critical coverage is attained through, for instance, direct teaching, the sooner the learners will be ready to start expanding their L2 vocabulary further through reading L2 texts (e.g., Nation, 1993). Our present study provides yet a further reason to begin L2 vocabulary teaching with common words; it is a pleasant coincidence that the words easiest to learn also happen to be the ones most useful to the learner.

To summarize, word-learning appears a more effective means to acquire vocabulary in an L2 than picture-learning, at least when the learners are relatively experienced FL learners. Congruency also affects performance, which, in terms of speed of retrieval, is better when the learning and test conditions match than when they mismatch. Finally, the cognate status and frequency of use of the stimulus materials strongly affect

learning. Interestingly, the effects of cognate status strongly suggest that learners generate the phonological forms of stimuli, both words and pictures, during learning. Regarding the frequency effect, concept familiarity may underlie it: It may be easier for learners to assign a new label to a familiar concept than to an unfamiliar concept.

Limitations of the Study

A number of restrictions of our study constrain the generality of the results and their practical implications. By including the picture-learning condition, we constrained ourselves to the use of concrete, picturable nouns only. A further limitation regarding the stimulus materials is that we defined cognate status exclusively in terms of Dutch-Italian translation pairs, not taking into account the role of any other FLs known to the learners. It remains to be seen in future research to what extent the results generalize to the learning of abstract words and to words defined as cognates on the basis of a broader definition of cognate status than adopted here. Yet another limitation of this study was that we tested only productive learning, even though receptive learning is a far more common technique for testing vocabulary acquisition in the early stages of learning. Furthermore, we tested only one particular group of learners, adult university students, learning words in a completely new FL rather than learning new words in a FL of which they already had some knowledge at the onset of training. Finally, we administered no delayed retention test, so we cannot tell how persistent the obtained effects were. All these restrictions point to possible extensions of future research on vocabulary acquisition in a foreign language.

Revised version accepted 11 July 1997

Notes

[1]We correlated the frequency of the 80 selected Dutch words with the frequency of their Italian translations in a new Italian frequency count (Istituto di Linguistica Computazionale, in press). The correlation between the frequencies in Dutch and Italian was 0.79, $p < .0001$ (cf. the correlation of 0.78 between Dutch and English word frequency, as reported by De Groot et al., 1994). We can therefore conclude that the Italian translation equivalents of the present high-frequency words tended to be high-frequent as well, whereas the Italian translations of the low-frequency words tended to be low-frequent.

[2]Had we been in a position to match the cognate ratings of the cognate stimuli across the 2 frequency groups any better we would have done so. However, because the total set of 224 word pairs with which we started out contained only 45 pairs that shared a cognate relation between Dutch and Italian (according to the criterion that we adopted) and because the other selection constraints also had to be taken into account, there were simply no cognate pairs left from which to select.

[3]"Cognate" and "noncognate" as used here refer only to the perceptual similarity between pairs of Dutch-Italian translation equivalents. One or both terms in a translation pair thus assigned a "noncognate" status (e.g., Italian "cavallo" and Dutch "paard," "horse"; or Italian "coltello" and Dutch "mes," "knife") may well have a cognate translation in other languages known to some of our participants (e.g., the French "cheval" and "couteau," respectively, are cognates of the Italian terms). Similarly, a word assigned a cognate status according to our definition may be a cognate in its relation with other languages as well (e.g., the Italian-Dutch pair "cannone-kanon," both of which are cognates of their translations in English and French). In fact, the Italian-Dutch cognates in our materials are often cognate with both English and French as well, whereas the Italian-Dutch noncognates are mostly also noncognate with English. Furthermore, unlike any of the Dutch words, a number of the Italian terms in the Italian-Dutch noncognate pairs (16 out of the 40) are cognates of the corresponding words in French. The participants in our study may to some extent have exploited their knowledge of other languages while learning the Italian equivalents of the Dutch words. However, we had no reason to suspect that this possibility undermined the conclusions regarding the effects of cognate status and word frequency that we drew from the data: Our cognates were generally more often than the noncognates also cognate with other languages known by the participants in this study, and this held for high- and low-frequency words alike. Furthermore, the groups of high-frequency and low-frequency words did not differ from each other in the number of Dutch-Italian pairs of which one or both terms were also cognates of words in other languages. So even if the participants somehow exploited knowledge of cognate relations with other languages than Italian to some extent, the effects of the word-type manipulations

would still be effects of the targeted manipulations: cognate status and word frequency.

[4]The number of learning trials per stimulus pair in this study was the same as in Chen (1990, Experiment 3), but the number of test trials per word was smaller. Chen included 3 tests, one after one learning trial only, the second after 2 more learning trials, and the third after the final 3 learning trials (Chen, 1990, p. 285). However, he discarded the data from the first test session in his analyses because all participants performed very poorly after the first of the 6 learning trials. In other words, the reported data from Chen's study concerned test performance after 3 learning trials and test performance after 6 learning trials, as in our study.

[5]Sets of analyses performed on the error and omission scores separately generally provided converging support for the conclusions drawn from the present data: The direction of the effects was always the same.

[6]That concrete words appear to have denser representations than abstract words (De Groot, 1989; Kieras, 1978) could also explain the finding that concrete words are easier to learn than abstract words (e.g., Ellis & Beaton, 1993a; Van Hell & Candia Mahn, 1997).

References

Baddeley, A. D., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language, 27*, 586–596.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10,* 340–357.

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language, 24*, 89–106.

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

Chen, H.-C. (1992). Lexical processing in bilingual or multilingual speakers. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 253–264). Amsterdam: North-Holland.

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

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.

Cristoffanini, P. M., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of Spanish-English cognates. *Quarterly Journal of Experimental Psychology, 38(A)*, 367–393.

De Groot, A. M. B. (1989). Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 15*, 824–845.

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

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., 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.

De Groot, A. M. B., & Poot, R. (1997). Word translation at three levels of proficiency in a second language: The ubiquitous involvement of conceptual memory. *Language Learning, 47*, 215–264.

Ellis, N. C., & Beaton, A. (1993a). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning, 43*, 559–617.

Ellis, N., & Beaton, A. (1993b). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *Quarterly Journal of Experimental Psychology, 46A*, 533–558.

Ervin, S. M., & Osgood, C. E. (1954). Second language learning and bilingualism. *Journal of Abnormal Social Psychology 49,* Supplement, 139–146.

Gregg, V. (1976). Word frequency, recognition and recall. In J. Brown. (Ed.), *Recall and recognition* (pp. 183–216). London: John Wiley & Sons.

Istituto di Linguistica Computazionale (in press). *Corpus di italiano contemporaneo.* Pisa.

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

Kieras, D. (1978). Beyond pictures and words: Alternative information-processing models for imagery effects in verbal memory. *Psychological Bulletin, 85*, 532–554.

Kirsner, K., Lalor, E., & Hird, K. (1993). The bilingual lexicon: Exercise, meaning and morphology. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 215–248). Amsterdam/Philadelphia: John Benjamins.

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

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, pp. 389–395). London: John Wiley & Sons.

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.

Meara, P. (1993). The bilingual lexicon and the teaching of vocabulary. In R. Schreuder & B. Weltens (Eds.). *The bilingual lexicon* (pp. 279–297). Amsterdam/Philadelphia: John Benjamins.

Nation, P. (1993). Vocabulary size, growth, and use. In R. Schreuder & B. Weltens (Eds.). *The bilingual lexicon* (pp. 115–134). Amsterdam/Philadelphia: John Benjamins.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language, 30*, 331–347.

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.

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

Thomas, M. H., & Wang, A. Y. (1996). Learning by the keyword mnemonic: Looking for long-term benefits. *Journal of Experimental Psychology: Applied, 2*, 330–342.

Uit den Boogaart, P. C. (1975). *Woordfrequenties in geschreven en gesproken Nederlands* [Word frequencies in written and spoken Dutch]. Utrecht: Oosthoek, Scheltema, & Holkema.

Van Hell, J. G., & Candia Mahn, A. (1997). Keyword mnemonics versus rote rehearsal: Learning concrete and abstract foreign words by experienced and inexperienced learners. *Language Learning, 47,* 507–546.

Van Hell, J. G., & De Groot, A. M. B. (1997). *Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association.* Paper submitted for publication.

Weinreich, U. (1953/1974). *Languages in contact: Findings and problems.* The Hague: Mouton. (Originally published in 1953 by Linguistic Circle of New York, New York).

Wimer, C. C., & Lambert, W. E. (1959). The differential effects of word and object stimuli on the learning of paired associates. *Journal of Experimental Psychology, 57*, 31–36.

Appendix

The Italian-Dutch Stimuli and Their Word Frequency Values, Cognate Ratings and Picture-Agreement Scores

High-Frequency Cognates

Italian	Dutch	English translation	Word frequency	Cognate rating	Picture agreement (%)
autobus	bus	bus	49	6.28	100
barba	baard	beard	13	5.28	80
castello	kasteel	castle	24	5.94	75
lampada	lamp	lamp	18	6.00	100
leone	leeuw	lion	12	5.17	100
mulino	molen	mill	15	5.17	100
naso	neus	nose	52	5.50	100
occhio	oog	eye	330	4.44	95
orecchio	oor	ear	48	4.00	100
orologio	horloge	watch	12	5.28	100
penna	pen	pen	11	6.28	100
pipa	pijp	pipe	22	6.00	100
sigaretta	sigaret	cigarette	35	6.83	80
sole	zon	sun	49	4.61	95
stella	ster	star	19	4.06	100
tavolo	tafel	table	106	5.17	100
telefono	telefoon	telephone	25	6.94	100
televisore	televisie	television	52	6.28	100
torre	toren	tower	18	5.78	100
treno	trein	train	44	5.78	95

Low-Frequency Cognates

Italian	Dutch	English translation	Word frequency	Cognate rating	Picture agreement (%)
ancora	anker	anchor	4	5.28	95
arpa	harp	harp	2	4.56	95
banana	banaan	banana	7	6.83	100
canguro	kangoeroe	kangaroo	0	6.00	100
cannone	kanon	cannon	2	6.50	100
chitarra	gitaar	guitar	2	5.67	95
coccodrillo	krokodil	crocodile	0	6.28	75
corona	kroon	crown	10	5.33	100
elefante	olifant	elephant	9	6.00	100
elicottero	helikopter	helicopter	7	4.89	95
forchetta	vork	fork	8	4.22	90
giraffa	giraf	giraffe	0	6.94	95
pera	peer	pear	1	6.06	100
pinguino	pinguin	penguin	0	6.83	85
racchetta	racket	racket	0	6.33	95
slitta	slee	sledge	3	4.22	85
tigre	tijger	tiger	2	6.06	85
torta	taart	cake	6	5.83	95
tromba	trompet	trumpet	1	5.11	85
vaso	vaas	vase	0	6.22	100

High-Frequency Noncognates

Italian	Dutch	English translation	Word frequency	Cognate rating	Picture agreement (%)
aereo	vliegtuig	plane	44	2.44	85
bandiera	vlag	flag	19	1.89	95
cane	hond	dog	57	2.11	100
cappello	hoed	hat	39	1.72	80
cavallo	paard	horse	54	2.44	95
chiave	sleutel	key	25	1.72	100
chiesa	kerk	church	291	2.89	100
coltello	mes	knife	11	1.50	100
farfalla	vlinder	butterfly	12	2.00	95
fiore	bloem	flower	64	2.11	100
fucile	geweer	rifle	22	1.83	90

High-Frequency Noncognates (cont.)

Italian	Dutch	English translation	Word frequency	Cognate rating	Picture agreement (%)
gonna	rok	skirt	22	1.44	95
letto	bed	bed	115	2.11	100
mela	appel	apple	11	1.50	95
mosca	vlieg	fly	19	1.50	90
nuvola	wolk	cloud	21	2.39	80
pesce	vis	fish	42	2.22	85
scarpa	schoen	shoe	34	2.00	95
scatola	doos	box	12	1.67	100
zattera	vlot	raft	15	1.56	90

Low-Frequency Noncognates

Italian	Dutch	English translation	Word frequency	Cognate rating	Picture agreement (%)
aquilone	vlieger	kite	0	1.22	85
ascia	bijl	axe	2	1.50	95
bilancia	weegschaal	scale	2	2.44	90
cancello	hek	gate	7	1.56	85
candela	kaars	candle	9	3.78	100
carriola	kruiwagen	wheelbarrow	0	2.44	100
cervo	hert	deer	0	2.17	90
cigno	zwaan	swan	0	1.50	90
ciliegia	kers	cherry	2	2.06	95
fiocco	strik	ribbon	3	1.44	100
fragola	aardbei	strawberry	2	1.44	95
guanti	handschoenen	gloves	8	1.83	100
martello	hamer	hammer	2	1.89	90
orso	beer	bear	6	1.83	100
rubinetto	kraan	tap	7	1.72	90
scopa	bezem	broom	0	1.28	100
spazzola	borstel	brush	2	1.56	100
tartaruga	schildpad	turtle	2	1.67	100
trapano	boor	drill	0	1.22	100
vestito	jurk	dress	10	1.78	95