
Testing the Development of French Word Knowledge by Advanced Dutch- and English-Speaking Learners and Native Speakers

TINE GREIDANUS

Vrije Universiteit Amsterdam

Faculteit der Letteren

De Boelelaan 1105

1081 HV Amsterdam

The Netherlands

Email: t.greidanus@worldonline.nl

BIANCA BEKS

Vrije Universiteit Amsterdam

Faculteit der Letteren

De Boelelaan 1105

1081 HV Amsterdam

The Netherlands

Email: biancabeks@wanadoo.nl

RICHARD WAKELY

University of Edinburgh

DELCL-French

59/60 George Square

Edinburgh EH8 9JU

Scotland, United Kingdom

Email: Richard.Wakely@ed.ac.uk

The present study is a continuation of the work presented in the 2001 article by Greidanus and Nienhuis. In the current study, we also examine the quality of word knowledge among advanced learners of French as a second language (L2) by means of a word associates test. We studied the development of word knowledge among 6 groups of university-level participants, who were (a) native speakers of French and (b) learners of French as a foreign language with 2 different first languages (L1s), Dutch and English. The format of the test differed from that used in the 2001 Greidanus and Nienhuis study as follows: (a) The tested words were less frequently used French words; (b) the participants were native speakers of French in addition to two categories of advanced learners of French; (c) the number of associate words (fixed or not) was an independent variable. The findings showed that both native and nonnative speakers of French progressed in deep-word knowledge when the results of third- and fourth-year students were compared with those of first-year students. Although the test contained a considerable number of French-English cognates, the L1 English learners did not perform better than the Dutch learners. The words tested were not noticeably more difficult when chosen from the 10,000 word level rather than from the 5,000 word level.

IN RECENT YEARS, DEPTH OF LEXICAL knowledge has become a matter of growing interest. By *depth of lexical knowledge* we mean all the knowledge that a speaker can have about a given word. Nation (1990) distinguished four dimensions of lexical knowledge: form (oral, written), position (grammatical, collocations), function (frequency, appropriateness), and meaning (concept and associations). There are no tests that examine all these aspects of lexical knowledge concurrently. Tests aiming to measure the vocabulary knowledge of advanced learners concentrate on grammatical and semantic aspects. The Word Associates Test, as presented by Read (1993, 1998), examines in particular the lexical

network that a learner possesses about a word, that is, its paradigmatic, syntagmatic, and analytic characteristics. Until now, the test has been used in quite a number of studies, applied to a range of languages, and given to both first language (L1) and second/foreign language (L2) participants. Some of these studies examined technical aspects of the test format (for an overview, see Greidanus, Bogaards, Van der Linden, Nienhuis, & De Wolf, 2004). The results of these studies have shown clearly that the test format is useful and that it has "potential as a tool in research on the organization of the mental lexicon and the processes by which word knowledge is acquired" (Read, 1993, p. 359). One sample item of this test has the following structure.

The Modern Language Journal, 89, ii, (2005)

0026-7902/05/221-233 \$1.50/0

©2005 *The Modern Language Journal*

edit

arithmetic

revise

film

risk

pole

surface

publishing

text

For this item, the test taker has to indicate whether or not there is a relationship between the target word and the response words. Here, the association words are *film*, *publishing*, *revise*, and *text*.

Greidanus and Nienhuis (2001) used the Word Associates Test to assess deep-word knowledge of frequently used French words, and found the test a useful instrument. To date, the test has been used to measure knowledge of rather basic vocabulary, and the question arises whether it is also useful to measure depth of lexical knowledge among advanced students. Therefore, we decided to use the same test to examine the depth of knowledge of less frequently used French vocabulary items.

THE 2001 STUDY

Given that the present study is a continuation of the earlier work by Greidanus and Nienhuis (2001), we will review that study in some detail. In 2001, the authors examined the development of word knowledge of two groups of advanced learners of French as a L2 by means of a slightly revised version of the test format presented by Read (1993). The participants comprised 30 first-year and 24 third- and fourth-year Dutch university students of French language and literature. The test contained 50 words from the frequency list of Juilland, Brodin and Davidovitch (1970), which contains 5,083 words that are divided into five frequency ranges of 1,000 words each. From each of the five frequency ranges, there were 10 stimulus words. These words were both concrete and abstract and included nouns, verbs, and adjectives. The association words and the distractors either belonged to the same frequency range as the stimulus word or were more frequent. The purpose of the 2001 study was to examine the development of word knowledge as measured by a word associates test. Three aspects of lexical development in particular were studied. The first concerned a technical aspect of the word associates tests, namely which type of distractor was most suited to the participants: distractors that were semantically related to the stimulus word or semantically nonrelated to it. Most studies have used semantically nonrelated distractors, but this type of distractor seemed too easy for our participants. The second aspect concerned the types of word associations: paradigmatic (super- and subordinates, synonyms, and antonyms), syntagmatic (collocations and selection restrictions), and analytic (defining characteristics such as those used in dictionary definitions). Each test item contained one association of each type. It was predicted, on the basis of native language free associations, that

the scores for the paradigmatic associations would be higher than those for the syntagmatic associations. The final aspect of the study concerned the relation between the frequency of a word and the quality of knowledge about it. The hypothesis was that the more frequent the word, the more knowledge the learner would have about its different aspects.

The following is an example of the two versions of a test item, the one using semantically related distractors and the other using semantically non-related distractors.

<i>rive</i> (bank)	<i>rive</i> (bank)
<input type="checkbox"/> <i>artificiel</i> (artificial)	<input type="checkbox"/> <i>bord</i> (bank)
<input type="checkbox"/> <i>bord</i> (bank)	<input type="checkbox"/> <i>fleuve</i> (river)
<input type="checkbox"/> <i>côté</i> (side)	<input type="checkbox"/> <i>gauche</i> (left)
<input type="checkbox"/> <i>fleuve</i> (river)	<input type="checkbox"/> <i>paquet</i> (parcel)
<input type="checkbox"/> <i>gauche</i> (left)	<input type="checkbox"/> <i>prudent</i> (careful)
<input type="checkbox"/> <i>vague</i> (wave)	<input type="checkbox"/> <i>tombe</i> (grave)

The association words that figure in both versions are *bord* (paradigmatic), *gauche* (syntagmatic), and *fleuve* (analytic). The left-hand test item contains three semantically related distractors (*artificiel*, *côté*, *vague*); the right-hand version includes three semantically nonrelated distractors (*paquet*, *prudent*, *tombe*).

The test discriminated well between the two groups of participants. The mean scores were relatively high; consequently, the test was rather easy for the first-year students and too easy for the third- and fourth-year students.

The semantically related distractors were found to be more appropriate in assessing the quality of word knowledge of advanced learners than the semantically nonrelated distractors. Indeed, being more distracting than the semantically non-related distractors, they give a better idea of the word knowledge of the test taker.

The participants showed a preference for paradigmatic responses, as predicted. The mean scores for the analytic associations were also higher than those for the syntagmatic associations.

As to the third aspect of the study, there was a relationship between a word's frequency and the quality of learners' knowledge of it: the more frequent a word, the better the learner's knowledge of the aspects tested. Apparently, vocabulary grows in depth more slowly than it grows in breadth: The quality of knowledge of frequent, earlier acquired words is better than that of less frequent, more recently acquired words.

The overall conclusion of the 2001 study was that the format of the Word Associates Test appears to be appropriate for assessing the quality

of word knowledge of fairly advanced L2 learners. The test used in the 2001 study discriminated between the two groups of learners at different levels, but there was a ceiling effect; a more difficult test might have shown greater difference between the two levels.

THE PRESENT STUDY

In the present study, we set out to use the test format developed in the study by Greidanus and Nienhuis (2001) in order to examine the quality of word knowledge of less frequently used French words. Some of the questions considered in the present study were identical to those examined in the earlier study, whereas others were different. The types of associations (paradigmatic, syntagmatic, and analytic) were a constant factor in both studies. As for the question of semantically related distractors versus semantically nonrelated distractors, the results of the previous study showed that the semantically nonrelated distractors were less well suited to the advanced level of our participants; we decided, therefore, to eliminate this variable from the present study. The distractors were all semantically related to the stimulus word. The aspects that were new to this study were (a) the vocabulary tested, (b) the characteristics of the participants, and (c) the number of the associate words present in each item. The following sections will look at these areas in more detail.

Vocabulary Tested

The vocabulary tested in this study goes beyond the 5,000 most frequently used word level: We used words from frequency ranges 5,000 to 10,000. According to Nation (1990), "The 5,000-word level is on the boundary of high- and low-frequency vocabulary" (p. 79). The knowledge of individual learners rapidly diverges once the 5,000 level is passed. We predicted in this study that the differences between the three pairs of groups at different learning levels would be larger than the differences between the two groups at different learning levels in the Greidanus and Nienhuis (2001) study.

Participants: Native and Nonnative Speakers

The majority of studies using a word associates test format have examined the knowledge of either L2 or L1 speakers (see Greidanus et al., 2004, for an overview). The study by Verhallen, Özdemir, Yüksel, and Schoonen (2001) is an exception: The authors studied the deep-word

knowledge of L1 and L2 speakers of Dutch. The current study tested the knowledge of native speakers of French, and of learners of French as a L2 with two different native languages (Dutch and English). A deep-word knowledge test has the advantage that it is, in principle, independent of the native language of the participants. If such knowledge is indeed independent, then the results of the two categories of nonnative speakers of French, who had comparable backgrounds, should be more or less similar. The French native speakers served as the control group.

Fixed Number of Associate Words or Not?

The number of associate words is fixed in most studies using word associates tests and is always half the number of the possible responses. The study by Greidanus et al. (2004) was an exception because in it the number of associate words varied between two and four. Using a variable number of word associates has several advantages. Read (1993) found that participants tried to guess the meaning of an unknown stimulus word, and knowing that there were always three associates made guessing easier. A test with a variable number of associates should therefore be more difficult than a test with an unvarying number of associates. Using a variable number of associates also has a technical advantage because it makes items easier to construct. Greidanus et al. (2004) observed that "it is not always possible to find three correct association words for a given word. If one can make a test item with two or even one association word, this increases the range of testable words" (p. 203).

METHOD

The researchers carried out an experiment to examine the previously mentioned aspects of deep-word knowledge. The following sections describe the design of the experiment.

Participants

The test was administered to 149 participants, consisting of three pairs of groups: The first pair consisted of 24 first-year and 24 third- and fourth-year students in the French departments of the Free University Amsterdam and the University of Leiden, all of whom were native speakers of Dutch. The students had a secondary education with an average of 2 or 3 hours of French per week over a period of 6 years. Most of the third- and fourth-year students had stayed in France for more or less long periods, for holidays or studies

at French universities. They were thus advanced (first-year students) and very advanced (third- and fourth-year students) learners of French. These students are called D1 and D3/4. The second pair comprised 21 first-year and 20 fourth-year students of the French section of the University of Edinburgh, all native speakers of English. The students had all studied French during their secondary education and were therefore comparable to the Dutch-speaking students, though their exposure varied from 5 years (leading to Scottish Higher) to 7 years (leading to A Level); some students had started French in primary school or had spent fairly long periods in French-speaking countries. These students are called E1 and E4. The third pair, the control group, consisted of 30 first-year and 30 third-year students in the Department of Linguistics of the University of Paris X-Nanterre, all native speakers of French. These students are called F1 and F3.

Tests

Greidanus and Nienhuis (2001) used the frequency list found in Juilland et al. (1970). Because that list contains only the 5,083 most frequently used words, we could not use it for this study. The only suitable existing frequency list was by Imbs (1971). This list is based on a large corpus, consisting mainly of 19th- and early 20th-century texts, mostly of a literary nature. Because this list is somewhat biased and out-of-date, we used an additional list, by Verlinde and Selva (2001), which is based on a 50 million word corpus of contemporary newspaper texts. The words of this list form the entries in an electronic learners' dictionary for French (*DAFLES—Dictionnaire d'apprentissage du français langue étrangère ou seconde*).

In creating our list of words, we used a combination of the two lists. With two exceptions, each word selected had to be found in both earlier lists, and in the same thousand-word frequency range. In the 9,000–10,000 frequency range, there were no verb matches on the two lists. So we chose two verbs (*raccrocher* 'to hang up' and *éluder* 'to elude') which were in the 9,000–10,000 range on the Imbs list, and in the 8,000–9,000 range on the Verlinde and Selva list. This selection procedure yielded an initial list of 260 words. Next, we excluded the French-Dutch cognates and the words that appear in the *Français Fondamental* list (see Gougenheim, Michéa, Rivenc, & Sauvageot, 1967). These words are transparent or are taught in secondary school. Regarding the cognates, the test was created for native speakers of Dutch. It was subsequently used with English-speaking participants. Given that English has a greater num-

ber of cognates with French than does Dutch, the English group could be expected to have an advantage in this respect.

We chose the stimulus words for the test from each range of 1,000 words and included both concrete and abstract words, as well as nouns, verbs, and adjectives. It is well known (see, e.g., Juilland et al., 1970; Sciarone, 1979) that the proportion of words belonging to the different word classes changes as the frequency falls. The first 1,000 words include many grammatical words (prepositions, conjunctions, determiners, pronouns) and adverbs. In the subsequent frequency ranges, the proportion of nouns grows rapidly compared to that of verbs and adverbs. On the basis of Sciarone's work, we selected 6 nouns, 2 verbs, and 2 adjectives from each of the five frequency ranges, making 50 words in all. In addition, there were 10 filler words, required to test the question about the number of associates. We chose two words from each of the five frequency ranges but took no account of the grammatical category. For the list of the 60 stimulus words see Appendix A.

The association words were selected from three French dictionaries: *Lexis, dictionnaire de la langue française* (1975), *Le nouveau petit Robert* (1996), and *Thésaurus Larousse* (1992). Each association word had to be present in at least two of the three dictionaries because, in our view, if a relationship appears in more than one dictionary, it means that the link is sufficiently strong. In addition, the association words either belonged to the same frequency range as the stimulus word or were of a higher frequency. There were three types of associations: (a) paradigmatic, (b) syntagmatic, and (c) analytic, as in the Greidanus and Nienhuis (2001) study. It was not easy to find three associates of these types for each possible stimulus word; in particular, syntagmatic associates were often not to be found. Therefore, many words in the original list of 260 had to be discarded.

Each test item had three distractors. Greidanus and Nienhuis (2001) studied the question of the nature of the distractors. They found that semantically unrelated distractors were too easy for the participants tested. Therefore, in this study we opted for semantically related distractors. It was not easy to find distractors semantically related to each possible stimulus word because these distractors risk being too "attractive" to the participants who may consider them associate words too easily. After administering a preliminary version of the test, we replaced a certain number of distractors by "less attractive" distractors.

In order not to make the test too difficult, we chose to make the adjectives agree in gender and

number when the stimulus word was a feminine noun. The following is an example of such a case. (The associate words are preceded by a black square.)

stature (stature) *décision* (decision)
 haute (high)
 importance (importance)
 monument (monument)
 signer (to sign)
 taille (size; waist)

One of the questions concerned the possible effects of having a fixed number of associates or a variable number of them. In order to explore this question, we constructed two versions of the test. Test A had 60 items that each had three associates and three distractors: 50 of these items were real test items; the other 10 were fillers. Test B also had 60 items, 50 of which were identical to the 50 test items of test A; the remaining 10 items, the fillers, had a variable number of associates. In the instructions for test A, the participants were told that there were always three correct answers; for test B, they were told that the number of associate words varied. The items were presented in alphabetical order. The associates and the distractors were likewise presented in alphabetical order. There were two order variants of both tests: One half of the participants responded to the test items in the order from *a* to *z*; the other half responded to them in the reverse order, *z* to *a*. This procedure was used to reduce any effect attributable to fatigue. A preliminary version of the test was tried with five university teachers of French, one of whom was a native speaker, and three university students, one of whom was also a native speaker. On the basis of the responses of these participants, 18 items were modified.

The test was preceded by fairly extensive instructions (three-fourths of a page) explaining the different types of relationships between stimulus words and association words, followed by an example item with comments (see Appendix B).

The test was administered in class or individually. No time limit was imposed. When scoring, 1 point was awarded for each correct associate word selected and for each distractor not selected. The maximum was 6 points per item and 300 points for the whole test. The performance on the fillers was not part of a participant's score.

Hypotheses

Our hypotheses were as follows.

1. The third- and fourth-year students would give more correct answers than the first-year

students. This result would hold for nonnative speakers as well as for native speakers of French.

2. There would be no statistically significant differences between first-year Dutch students and first-year English-speaking students, nor between third- and fourth-year Dutch students and fourth-year English-speaking students.

3. The participants who were told that there were always three associates for each test item would give more correct answers than participants who were told that the number of associates varied.

4. There would be a relation between frequency and quality of knowledge; that is, the more frequent the word, the better the learner's knowledge of different aspects of the word.

5. The three types of associations (paradigmatic, syntagmatic, and analytic) would not be equally strong, the order of strength being: paradigmatic relationship > analytic relationship > syntagmatic relationship.

RESULTS

Results per Student Level

We had three different language groups and two levels of students. We first calculated the mean scores and the standard deviations for the six participant groups, treating the two versions, A and B, of the test together. These data appear in Table 1.

The distinct level groups behaved according to Hypothesis 1: The more advanced students gave more correct answers than the less advanced students, and this result held for both nonnative and native speakers.

As for Hypothesis 2, three one-way analyses of variance (ANOVAs) applied to Level 1 and Level 3/4 showed an effect of level (see Table 1). A Newman-Keuls analysis indicated that there was no difference between the D1 and E1 students, but they scored lower than any of the other groups. The other four groups all differed significantly from one another, with the E4 students being the lowest, followed by the D3/4 and F1 groups, whereas the F3 group had the highest scores ($p < .05$). Therefore, one may conclude that the test had good discriminating power at these advanced levels. The results of the Newman-Keuls analysis are shown in Table 2. The means for groups that did not differ significantly from one another (homogeneous subsets) are displayed in the same column.

In the Greidanus and Nienhuis (2001) study, the test discriminated between the two groups at

TABLE 1
Mean Scores and Standard Deviations of the Three Language Groups at Two Levels, with their *F*-value
(Maximum Score = 300)

Group	Level 1			Level 3/4			<i>F</i>
	Score (%)	<i>M</i>	<i>SD</i>	Score (%)	<i>M</i>	<i>SD</i>	
Dutch Speakers <i>n</i> = 24/24	60.9	182.7	15.4	72.2	216.5	22.5	<i>F</i> (1, 46) = 36.9*
English Speakers ^a <i>n</i> = 21/20	61.5	184.5	11.7	68.0	203.9	12.4	<i>F</i> (1, 39) = 26.6*
French Speakers <i>n</i> = 30/30	75.5	226.5	18.6	81.8	245.5	17.5	<i>F</i> (1, 58) = 16.6*

^aThe English speakers were studying in Scotland.

**p* < .001.

different learning levels, but there was a ceiling effect: The differences could have been greater with a more demanding test. Are the differences between the two Dutch level groups indeed greater in this test? The percentages of correct responses in the Greidanus and Nienhuis (2001) study were 82.9% for the first-year students and 86.8% for the third- and fourth-year students, a difference of 3.9%. In the present study, they were 61% and 72% respectively, a difference of 11%. The standard deviations were also much greater in the present experiment: 15.4 for the first-year students and 22.5 for the third- and fourth-year students, compared to 5.8 and 5.4, respectively, in the Greidanus and Nienhuis (2001) study. So this test of less frequently used words discriminates better than the test of frequently used words. In addition, the knowledge within each group is more varied.

Our second hypothesis was partially disproved: The E4 students scored significantly lower than did the D3/4 students. This was a rather intriguing finding. Why did the E4 group score lower than their Dutch counterparts? They had an advantage in that some of the item words were cognates. The results of the E1 students did not differ

significantly from those of the D1 students, but they should have been better, given their advantage. This finding led us to look into the question of the cognates, which was not originally planned as a part of our study. The test contained 23 French-English cognate stimulus words¹ and 27 noncognate stimulus words. The mean of the E1 students on these cognates was 4.0 per item (maximum = 6); their mean on the noncognates was 3.4 per item (maximum = 6). The mean of the E4 students on the cognates was 4.4 per item; their mean on the noncognates was 3.8 per item. The difference between cognates and noncognates for both groups was 0.6 per item. The progress of the E4 students was identical for noncognates and cognates (a mean of 0.4 per item). A *t* test revealed that the differences between the mean scores for cognates and noncognates were significant (*p* < .001). If the English-speaking participants were better on French-English cognates than on noncognates, this implied that the Dutch participants outscored them on the noncognates. The Dutch participants had been exposed to English in secondary school and in the course of their studies, but they certainly did not have the same knowledge of the French-English cognates as the English-speaking participants.

TABLE 2
Newman-Keuls Analysis of the Scores of the Six Groups

Group	<i>N</i>	1	2	3	4	5
English 1 ^a	21	184.5				
Dutch 1	24	182.7				
English 4 ^a	20		203.9			
Dutch 3/4	24			216.5		
French 1	30				226.5	
French 3	30					245.5

^aThe English speakers were studying in Scotland.

Results per Test Version

In order to test Hypothesis 3, we had two versions of the test (see section *Tests*). The participants taking test A were told that there were always three associates for each test item, and those taking test B were told that the number of associates varied. Table 3 shows the mean scores and the standard deviations for the six participant groups. The reliability of the test (Cronbach's alpha) was .91 for version A and .88 for version B.

TABLE 3

Mean Scores and Standard Deviations of the Six Level Groups for the Two Test Versions, with their *F*-Value (Maximum Score = 300)

Group	Test Version A			Test Version B			
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>F</i>
Dutch 1	12	180.8	15.6	12	184.5	15.6	$F(1, 22) = 0.33, ns$
Dutch 3/4	12	216.0	18.3	12	217.0	27.0	$F(1, 22) = 0.01, ns$
English 1 ^a	11	185.4	11.9	10	183.6	12.1	$F(1, 19) = 0.12, ns$
English 4 ^a	10	206.3	8.0	10	201.6	15.7	$F(1, 18) = 0.72, ns$
French 1	15	227.1	19.9	15	225.9	17.8	$F(1, 28) = 0.03, ns$
French 3	15	254.4	12.4	15	236.5	17.5	$F(1, 28) = 10.4^*$

Note. Version A = fixed number of associates; Version B = variable number of associates.

^aThe English speakers were studying in Scotland.

* $p < .005$.

We observed small differences between the two versions for each of the three language pairings. These differences were not significant, except for the French native speakers at the highest level. Hypothesis 3 was therefore disproved for the non-native groups and for the first-year French native speakers, but not for the third-year French native speakers: The mean score for test A participants was higher than that for test B participants. This result is rather curious. We stated in a previous section (*Fixed Number of Associate Words or Not?*) that “knowing that there were always three associates made guessing easier. A test with a variable number of associates should therefore be more difficult than a test with an unvarying number of associates.” Apparently, this statement applies only to the highest-level group of native French speakers. We are not very sure how to interpret this finding. Did the third-year French native speakers rely more on guessing than the other five groups?

Because the French native speakers functioned as control groups, and given that the results for the nonnative groups are unambiguous, we will in the following sections treat the two test versions together.

Results per Frequency Range

Greidanus and Nienhuis (2001) found that there was a relationship between quality and quantity of lexical knowledge, that is, that deep knowledge about frequently used words was better established than the same type of knowledge about less frequently used words. Does this result also hold for words above the 5,000 word level? To answer this question, that is, to test Hypothesis 4, we distinguished five levels of frequency, as in the 2001 study. For each frequency level of 1,000 words which was represented by 10 test items, we calcu-

lated the scores of the students of the three pairs of groups of different learning levels. Figure 1 shows the mean scores per group, for the five frequency ranges.

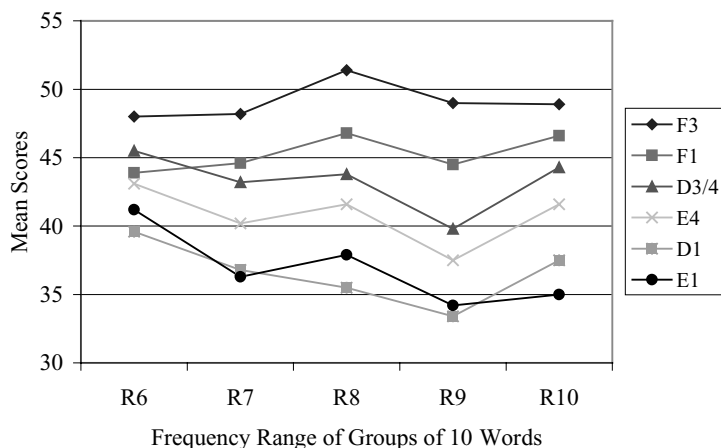
The mean scores of the six participant groups diminished from one frequency range to the next in 12 cases, but also increased in 12 cases. There are few general conclusions possible. For all six groups the means for Frequency Range 9 were higher than for Frequency Range 8. The means for Frequency Range 10 were higher than those for Frequency Range 9 for five out of six groups. One-factor within-subjects analyses for each of the six groups showed an effect of frequency for all groups ($p < .001$), but there was no tendency along the lines of Hypothesis 4.

Results per Type of Association

As noted previously, the association words used in the test represented different types of associations: paradigmatic, syntagmatic, and analytic. Each test item contained one paradigmatic association, one syntagmatic association, and one analytic association. In the Greidanus and Nienhuis (2001) study, we found that the means for the paradigmatic associations were higher than those for the syntagmatic associations; the same was true when comparing the analytic associations with the syntagmatic associations. The data for the present study appear in Table 4.

In order to test Hypothesis 5 that the order of strength of the three types of associations would be paradigmatic > analytic > syntagmatic, we carried out within-subject analyses on association (paradigmatic, syntagmatic, analytic), followed by Bonferroni pairwise analyses. Association had an effect in all 6 groups. The pairwise analyses in each group between paradigmatic, syntagmatic,

FIGURE 1
Participant Groups' Mean Scores for Knowledge of Less Frequently Used Words (Maximum Score = 60)



Note. Frequency Range 6 = 6,000; 7 = 7,000; 8 = 8,000; 9 = 9,000; 10 = 10,000. D1 = first-year Dutch-speaking students; D3/4 = third- and fourth-year Dutch-speaking students; E1 = first-year English-speaking students; E4 = fourth-year English-speaking students; F1 = first-year French-speaking students; F3 = third-year French-speaking students.

and analytic associations showed significant differences ($p < .001$) between all three in the expected direction for the D3/4, F1, and F3 students. In the D1 group, there was a difference between paradigmatic and syntagmatic, and syntagmatic and analytic associations, and the same applied to the E1 and E4 students. Thus, our hypothesis held for the three most advanced groups; for the three other groups, paradigmatic and analytic associations are both better acquired than syntagmatic associations, but there was no statis-

tical difference between paradigmatic and analytic associations. This finding corresponds with the results in the Greidanus and Nienhuis (2001) study.

A two-factor mixed ANOVA showed that the factor Association was significant, $F(2, 286) = 451.1$, $p < .001$, as was the interaction between Group and Association, $F(10, 286) = 6.2$, $p < .001$. Post hoc Newman-Keuls analyses showed that, for the paradigmatic association, the E1 and D1 students did not differ from one another. Both groups had

TABLE 4
Mean Scores and Standard Deviations for the Three Different Types of Associations: Paradigmatic, Syntagmatic, Analytic (Maximum Score = 50)

Group	Paradigmatic		Syntagmatic		Analytic	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Dutch 1 <i>n</i> = 24	31.3	6.4	19.4	7.2	30.2	6.7
Dutch 3/4 <i>n</i> = 24	41.4	7.2	27.0	8.7	37.7	6.5
English 1 ^a <i>n</i> = 21	31.2	5.2	19.6	5.2	31.0	6.6
English 4 ^a <i>n</i> = 20	36.9	3.5	24.4	5.1	36.0	4.9
French 1 <i>n</i> = 30	45.5	4.2	30.8	4.9	40.1	4.2
French 3 <i>n</i> = 30	45.8	5.9	35.9	7.4	39.5	5.9

^aThe English speakers were studying in Scotland.

TABLE 5
Newman-Keuls Analysis of the Scores for the
Paradigmatic Relations

Group	<i>n</i>	1	2	3	4
English 1 ^a	21	31.2			
Dutch 1	24	31.3			
English 4 ^a	20		36.9		
Dutch 3/4	24			41.4	
French 1	30				45.5
French 3	30				45.8

^aThe English speakers were studying in Scotland.

lower scores than the E4 group, whose score was lower than that of the D3/4 group, with the two French groups having the highest scores. There was no difference between the two French groups. Similar outcomes were found for the syntagmatic association, but here there was a difference between the two French groups. As for the analytic association, there were just two groups with the first-year L2 students in the first group and all the others in the second group (see Tables 5, 6, and 7).

DISCUSSION AND CONCLUSION

The results when applied to our five hypotheses led to the following observations. For Hypothesis 1, third- and fourth-year university students did indeed give more correct answers than their first-year colleagues, and this finding was true for both native French groups and nonnative groups, and for both versions of the test (except for the F3 students, who did significantly better in the fixed number of associates condition). These results may appear to be what one would expect, but it must be stated that we are dealing with an autonomous learning process: There had been no explicit teaching of the lower frequency vocabulary. The F1 speakers outscored all the English and Dutch speakers, but still scored significantly lower than the F3 speakers. The test had good

TABLE 6
Newman-Keuls Analysis of the Scores for the
Syntagmatic Relations

Group	<i>n</i>	1	2	3	4
English 1 ^a	24	20.4			
Dutch 1	21	19.6			
English 4 ^a	20		24.4		
Dutch 3/4	24		26.9		
French 1	30			30.8	
French 3	30				35.9

^aThe English speakers were studying in Scotland.

TABLE 7
Newman-Keuls Analysis of the Scores for the
Analytic Relations

Group	<i>n</i>	1	2
English 1 ^a	24	30.2	
Dutch 1	21	31.0	
English 4 ^a	20		35.9
Dutch 3/4	24		37.7
French 1	30		39.5
French 3	30		40.1

^aThe English speakers were studying in Scotland.

discriminating power in this study. We did not encounter the ceiling effect found in the Greidanus and Nienhuis (2001) study.

The nonnative speakers in the present study did not score as well as the nonnative speakers in the Greidanus and Nienhuis (2001) study, but this finding was to be expected given the less frequently used vocabulary tested in the present study. With 61% to 72% correct answers, they scored rather well. If we had not eliminated the *Français Fondamental* words, their scores would even have been better.

As for the native speakers, we see that lexical development is an ongoing, continuous process; the third-year students performed better than the first-year students, with room for improvement. The difference between the two learning level groups is due almost entirely to their relative mastery of the syntagmatic associations.

Hypothesis 2 stated that there would be no statistically significant differences between nonnative speakers with different native language backgrounds, but of approximately the same learning level. Although there were no statistically significant differences between first-year Dutch students and first-year English students, the third- and fourth-year Dutch students outscored the fourth-year English students. How can we explain this finding? We have already pointed out that English has a greater number of cognates with French than does Dutch. It is possible that the E1 group appeared better than it really was, with students able to guess correctly that French forms meant (roughly) the same as words with similar forms in English. This possibility does not explain the low level of progress of the E4 group, unless one assumes that English speakers adopt, early on, a strategy of assuming that the “density” of French-English cognates permits them to devote their energies to other areas, such as grammar and pronunciation. This hypothesis would need further investigation.

We saw that the E4 group progressed as much on cognates as on noncognates. Recognizing a French-English cognate does not necessarily imply that one masters its whole lexical network. If we look at a concrete example, the *stature* item given previously, we see that the associates provided were *haute*, *importance* (cognate), and *taille*. The distractors (*décision*, *monument*, and *signer*) were all cognates. The major progress from the E1 group to the E4 group was on the word *taille* (for which the E1 group had 8 correct answers out of 21 and the E4 group had 15 correct answers out of 20). This progress is probably due to the fact that *taille* is a polysemous word, and therefore difficult. This is just one example showing the complexity of the cognate question. Other item words showed different patterns. In tests of lexical depth, it seems likely that cognates play a role that is distinct from their role in tests of lexical breadth. Because the issue of cognates was not a part of our research design, we can only speculate on this question.

Hypothesis 3—participants who were told there were always three associates for each test item would give more correct answers than participants who were told that the number of associates varied—was not confirmed amongst the L2 groups or the first-year French speakers, so this does not seem to be a significant factor among foreign language learners.

Hypothesis 4 concerned the relationship between the frequency of words and the quality of knowledge. There was no clear pattern here, in that, as frequency diminished from 5,000 to 10,000, we did not find the expected drop in quality of knowledge. Whether this finding is due to the fact, observed by Nation (1990), that the 5,000 word level is a sort of boundary beyond which knowledge, even among apparently homogenous groups, can be very varied, is uncertain.

Our idea (Hypothesis 5) that the three types of associations (paradigmatic, syntagmatic, and analytic) would not be equally strong, with the order of strength being paradigmatic relationship > analytic relationship > syntagmatic relationship, was largely proved correct, though there was no significant difference between paradigmatic and analytic associations for the two first-year L2 groups or for the E4 group.

When we compare the development of knowledge within the language pairings, we see that the more advanced L2 groups did better in all three categories than the first-year groups. However, although the French third-year group outperformed the first-year group on the syntagmatic associations, it did not outperform the less ad-

vanced F1 group on the paradigmatic and analytic associations. The latter two categories seem well established for native speakers at the age of 18 approximately, whereas improvement is still possible in syntagmatic relationships.

We conclude with some remarks about the Word Associates Test as such.

1. The test was found to have good discriminating power between the different participant groups. The psychometric data resulting from this study and other studies are good.

2. The test is independent of the native language of the participants: The corresponding L2 groups, D1 and E1, and D3/4 and E4, respectively, have similar results, despite the fact that the test contained many French-English cognates. In this study, the participants all had a European language as their native language. It remains to be seen if the same results will be obtained with participants with totally different native language backgrounds.

3. A word associates test is efficient to use, compared with other methods of deep-word knowledge testing. It is not always possible to find sufficient associates for a given word, but the fact that it is not always necessary to have three (or four) associates facilitates test construction.

4. The test can be used to assess deep-word knowledge, but it can also be used for studies on the vocabulary acquisition process, as has been illustrated in this study.

ACKNOWLEDGMENTS

We would like to thank a number of people without whose help and encouragement our work and this article could not have reached completion. We owe a great deal to Cor Koster (Vrije Universiteit Amsterdam), who carried out the statistical analyses and gave us statistical advice. We thank Danièle Flament-Boistrancourt (Université de Paris X-Nanterre), who administered the test to the French participants. We are also very grateful to Serge Verlinde and Thierry Selva (both of the Katholieke Universiteit Leuven), who put their frequency list at our disposal. We finally thank the anonymous reviewers, whose remarks have certainly helped to make the revised version of this article an improvement on the first version.

NOTE

¹ We distinguished two categories of cognates: (a) full cognates (same form, same meaning) and (b) near-cognate words (similar form, same meaning). Four words fell into category (a): *aptitude* (aptitude), *mêlée*

(mêlée, scrum), *profitable* (profitable), and *stature* (stature), and 19 words fit into category (b): *assigner* (to assign), *avantageux* (advantageous), *barrer* (to bar, to block), *brèche* (breach), *conjonction* (conjunction), *croisade* (crusade), *éluder* (to evade), *embarrassant* (embarrassing), *encombrer* (to encumber, to obstruct), *équité* (equity, fairness), *funérailles* (funeral), *impartialité* (impartiality), *irréprochable* (irreproachable), *lucidité* (lucidity), *oignon* (onion), *périlleux* (perilous), *rareté* (rarity), *recouvrer* (to recover), and *sordide* (sordid).

REFERENCES

- Dubois, J. (Ed.). (1975). *Lexis. Dictionnaire de la langue française* [Lexis. Dictionary of the French language]. Paris: Larousse.
- Gougenheim, G., Michéa, R., Rivenc, P., & Sauvageot, A. (1967). *L'élaboration du français fondamental (1er degré): Etude sur l'établissement d'un vocabulaire et d'une grammaire de base* [The elaboration of fundamental French (first grade): Study on the establishment of a basic vocabulary and grammar]. Paris: Didier.
- Greidanus, T., Bogaards, P., Van der Linden, E., Nienhuis, L., & De Wolf, T. (2004). The construction and validation of a deep word knowledge test for advanced learners of French. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition and testing* (pp. 191–208). Amsterdam: John Benjamins.
- Greidanus, T., & Nienhuis, L. (2001). Testing the quality of word knowledge in a second language by means of word associations: Types of distractors and types of associations. *Modern Language Journal*, 85, 467–477.
- Imbs, P. (1971). *Etudes statistiques sur le vocabulaire français. Dictionnaire des fréquences. Vocabulaire littéraire des XIXe et XXe siècles* [Statistical studies on French vocabulary. Dictionary of frequencies. Literary vocabulary of the 19th and 20th centuries]. Nancy, France: C. N. R. S.–T. L. F.
- Juilland, A., Brodin, D., & Davidovitch, C. (1970). *Frequency dictionary of French words*. The Hague, The Netherlands: Mouton.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.
- Péchoin, D. (Ed.). (1992). *Thésaurus Larousse. Des idées aux mots, des mots aux idées* [Larousse Thesaurus. From notions to words, from words to notions]. Paris: Larousse.
- Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, 10, 355–371.
- Read, J. (1998). Validating a test to measure depth of vocabulary knowledge. In A. Kunnan (Ed.), *Validation in language assessment* (pp. 41–57). Mahwah, NJ: Erlbaum.
- Rey-Debove, J., & Rey, A. (Eds.). (1996). *Le nouveau petit Robert*. Paris: Dictionnaires Le Robert.
- Sciarone, A. G. (1979). *Woordjes leren in het vreemdetalenonderwijs* [Vocabulary learning in foreign language instruction]. Muiderberg, The Netherlands: Coutinho.
- Verhallen, M., Özdemir, L., Yüksel, E., & Schoonen, R. (1999). Woordkennis van Turkse kinderen in de bovenbouw van het basisonderwijs [Lexical knowledge of Turkish children in the upper grades of primary education]. *Toegepaste Taalwetenschap in Artikelen*, 61, 21–33.
- Verlinde, S., & Selva, T. (2001). Nomenclature de dictionnaire et analyse de corpus [Dictionary nomenclature and corpus analysis]. *Cahiers de Lexicologie*, 79, 113–139.

APPENDIX A

Test Words with their Ranges in the Verlinde and Selva List and the Imbs List

Range	Word	Verlinde and Selva	Imbs
5,000–5,999	aptitude	5828	5362
	assigner	5001	5431
	avantageux	5684	5322
	barrer	5513	5219
	brèche	5199	5797
	cascade	5960	5050
	croisade	5904	5454
	défaillance	5341	5152
	lucidité	5722	5994
	périlleux	5867	5868
	savouurer*	5425	5692
	volée*	5704	5632

(Continued)

Range	Word	Verlinde and Selva	Imbs
6,000–6,999	baguette	6057	6467
	démenti	6301	6389
	déroute	6592	6910
	émoi	6491	6662
	encombrer	6086	6283
	engloutir	6572	6146
	s'évader*	6733	6545
	funérailles	6772	6624
	gîte	6548	6310
	grief*	6431	6016
	irréprochable	6943	6676
	ultérieur	6084	6574
	7,000–7,999	augure	7989
clouer*		7773	7377
coutumier*		7098	7899
crisper		7973	7772
émerveiller		7885	7667
équité		7163	7246
fût [n.]		7113	7466
infime		7941	7589
monceau		7558	7453
oignon		7413	7334
renversement		7185	7612
sordide		7823	7177
8,000–8,999		affût	8045
	bille	8073	8802
	effréné	8989	8454
	fougue	8715	8454
	fronde*	8422	8705
	mêlée	8771	8101
	proférer	8704	8101
	profitable	8824	8491
	rareté	8567	8558
	recouvrer	8748	8905
	riposter*	8428	8149
	zinc	8498	8705
	9,000–9,999	cigale	9162
conjonction		9907	9762
cuisant		9239	9013
éluder		8594	9380
embarrassant		9455	9980
étagère		9988	9639
impartialité		9717	9314
ligament*		9801	9438
parjure		9247	9639
printanier*		9779	9863
raccrocher		8179	9096
stature		9025	9419

Note. The words marked with * were used as fillers.

APPENDIX B

Deep Knowledge of French Vocabulary Test: Version A

[CIRCLE ONE ANSWER]

Enrolled in French 1 OR French 4 Matric. No.

Mother tongue/main language: English OR Bilingual English and [specify]

Sex: M F

The purpose of this test is to evaluate your deep knowledge of French vocabulary. By “deep knowledge” we mean NOT the number of words you know, but WHAT you know about certain words. In other words, we are looking at the QUALITY rather than the AMOUNT of your knowledge of French vocabulary.

All the items in this test follow this format:

jeune *âge* (age)
 blanc (white)
 fille (daughter; girl)
 livre (book)
 oiseau (bird)
 vieux (old)

Please fill in the squares preceding those words which have a CLEAR CONNECTION with the stimulus word (*jeune* in the example). There are different types of connection between pairs of words:

1. The words may be synonyms: *vieux-âgé* (old- aged)
2. The words may be antonyms (opposites): *haut-bas* (high- low)
3. One word may have a more general meaning than the other: *langue-français* (language- French)
4. One word may have a more specific meaning than the other: *studio-habitation* (studio, apartment- dwelling)
5. One word may often appear in the same context as the other: *coucher de soleil, bain de soleil, coup de soleil, soleil + briller, soleil + se coucher* (sunset, sunbathing, sunstroke, sun + to shine, sun + to set)
6. One word may be used to define the other: *intituler + titre* (*intituler* = *donner un titre* à ‘to entitle + title,’ since *to entitle* is ‘to give a title to’)
7. There may be a link in function or use between one word and an other: *stylo-écrire* (pen- to write)
8. One word may refer to a part, whereas the other word refers to the whole: *toit-maison* (roof- house).

So the test asks you to select the words which have a clear connection with the stimulus word. In the case of *jeune* (young), you should reply as follows:

jeune *âge* (*jeune* = *peu avancé en âge*/young = not far advanced in age)
 blanc (white)
 fille (*jeune fille*/young girl—fixed expression)
 livre (book)
 oiseau (bird)
 vieux (old—the words are antonyms/opposites)

The other words *blanc*, *livre*, *oiseau* have no clear connection with *jeune*.

The test contains 60 items. If you think that there is a clear, close link between the stimulus word and any one of the other words given, fill in the relevant square.

Please answer every item. Please note that the correct number of squares to be filled in is always three, so always fill in three squares.

Note. For the Deep Knowledge of French Vocabulary Test: Version B, the instructions are identical, with the exception of the last two lines:

Please answer every item. Please note that the correct number of squares to be filled in varies. (It is not always three as in the example.)