

UC Office of the President

Reprints

Title

Teaching a Large Russian Teaching a Large Russian Language
Vocabulary by the Mnemonic Keyword Method

Permalink

<https://escholarship.org/uc/item/3bf6w6g4>

Authors

Raugh, Michael R.
Schupback, Richar D.
Atkinson, Richard C.

Publication Date

1977

Peer reviewed

TEACHING A LARGE RUSSIAN LANGUAGE VOCABULARY BY THE MNEMONIC KEYWORD METHOD

MICHAEL R. RAUGH, RICHARD D. SCHUPBACH, and RICHARD C. ATKINSON

Stanford University

ABSTRACT

This study evaluates the effectiveness of a mnemonic procedure, called the keyword method, for teaching a large Russian language vocabulary to college students. The method divides the study of a vocabulary item into two stages. The first stage requires the student to associate the spoken Russian word with an English word (the keyword) that sounds like some part of the foreign word; the second stage requires the student to form a mental image of the keyword “interacting” with the English translation. Thus, the keyword method can be described as a chain of two links connecting a foreign word to its English translation through the mediation of a keyword: the foreign word is linked to a keyword by a similarity in sound (acoustic link), and the keyword is linked to the English translation by a mental image (imagery link). A computer controlled curriculum using the keyword method served as a supplement to the second-year Russian language course at Stanford University. Students studied a large basic vocabulary over an 8 to 10-week period. Data obtained during the study and student reports indicate that the keyword method was highly effective.

There are many obstacles to the mastery of a foreign language. One that has received little experimental study is vocabulary acquisition (Holley, 1971; Hughes, 1968). For the past three years we have been experimenting with foreign-language vocabulary acquisition through the use of a mnemonic procedure called the *keyword method*. This method is related to the classical technique used by Cicero and other Roman orators for memorizing long speeches and other information (Yates, 1972). Our previous studies have shown the keyword method to be a remarkably efficient means of teaching a foreign language vocabulary under the special conditions of the psychological laboratory (Atkinson, 1975; Atkinson and Raugh, 1975; Raugh and Atkinson, 1975). The study reported here goes beyond the psychological laboratory to determine whether the keyword method can be used as a supplement to the Russian language curriculum offered by the Department of Slavic Languages and Literatures at Stanford University.

The keyword method is a mnemonic procedure for associating a foreign word with its English translation. The method divides the study of a word into two stages. The first stage involves associating the spoken foreign word

with an English word that sounds approximately like some part of the foreign word. As an example from Spanish the word *caballo* (pronounced somewhat like “cob-eye-yo”), contains a sound that resembles the spoken English word “eye”; we call such a similar sounding word a *keyword*. In general, the keyword will have no relationship to the foreign word except similarity in sound. The second stage of the keyword method requires the subject to form a mental image of the keyword “interacting” with the English translation; this stage is comparable to a paired-associate procedure involving the learning of unrelated English words. In the case of *caballo* (translation: “horse”) one could form a mental image of something like a cyclopean eye winking in the forehead of a horse or a horse kicking a giant eye.

As an example from Russian, consider the word *zдание* (translation: “building”) [1]. It is pronounced roughly as “zdon-yeh,” with emphasis on the first syllable, and it contains a sound that resembles the English word “dawn.” Using “dawn” as the keyword, one could imagine the pink light of dawn reflected in the windows of a tall building.

The keyword method can be described as a chain of two links connecting a foreign word to its English translation through the mediation of a keyword. The foreign word is linked to the keyword by a similarity in sound (the *acoustic link*); in turn the keyword is linked to the English translation by a learner-generated mental image (the *mnemonic* or *imagery link*) [2]. One procedure we have used for applying the keyword method is to present the subjects with a series of foreign words. As each foreign word is pronounced its keyword and the English translation are displayed. During the presentation of each item the subject must associate the sound of the foreign word with the given keyword and, at the same time, generate a mental image relating the keyword to the English translation.

The preselection of keywords is an important aspect of the keyword method. Atkinson and Raugh (1975) have obtained independent measures of the effectiveness of a keyword, and have used these to predict learning by the keyword method; their results suggest that effectiveness of the keyword method depends upon a careful selection procedure. Accordingly, we have found it useful to employ a panel of individuals familiar with the keyword method to make keyword selections. In preparing a study vocabulary a keyword is considered eligible if it satisfies the following criteria:

1. The keyword sounds as much as possible like a part (not necessarily all) of the foreign word.
2. It is easy to form a memorable imagery link connecting the keyword and its English translation.
3. The keyword is unique (different from other keywords used in the vocabulary).

Criterion 1 allows for flexibility in the choice of keywords, since any part of a foreign word could be used as the key sound. What this means for a polysyllabic foreign word is that anything from a monosyllable to a longer word (or even a short phrase that “spans” the whole foreign word) might be used as a keyword. As examples of the two extremes, “truce” could be used as a keyword for Russian *truslivyj* (translation: “cowardly”), and the keyword phrase “Pierre is sick” could be used for *persik* (translation: “peach”). Criterion 2 attempts to make the imagery link as simple and memorable as possible. Concrete nouns often are good keywords, because they are easy to image; abstract nouns for which symbolic imagery springs to mind also are effective keywords. A good keyword is easily imaged in isolation; however, it must also be imageable in relationship to its paired English translation. Criterion 3 is used to avoid the ambiguities that could arise if a given keyword were associated with several foreign words. The selection of unique keywords is not a serious constraint even for a very large vocabulary. In the present study 675 words were used, and the selection of keywords presented no problem.

An example of the kind of laboratory studies that have encouraged us to pursue the keyword method is reported in Atkinson and Raugh (1975). Subjects learned a vocabulary of 120 Russian words; the vocabulary was divided into three 40-word subvocabularies for presentation on separate days. The experiment was run under computer control and involved two independent groups of subjects – a keyword group and a control group. The computer presented prerecorded Russian words through headphones, keywords and English translations were presented on a CRT display, and the subject entered his responses into the computer by means of a typewriter keyboard. The experiment began with an introductory session during which subjects were familiarized with the equipment and given some instruction in the phonetics of Russian; subjects in the keyword group were also given instructions on the keyword method. On each of the following three days one of the subvocabularies was presented for a cycle of three study/test trials. The study part of a trial consisted of a run through the subvocabulary; each Russian word was pronounced three times and simultaneously its English translation was displayed on the CRT. For the keyword subjects the keyword was also displayed on the CRT, set off in brackets. The test phase of a trial was exactly the same for both groups; a Russian word was pronounced and the subject had up to 15 seconds to type the translation. No feedback was given and no keywords were presented on test trials. A comprehensive test covering the entire vocabulary of 120 items was given on the fifth day of the experiment. Without warning subjects were called back six weeks later for a second comprehensive test.

On all daily test trials the keyword group obtained superior scores; each day the keyword group learned more words in two study trials than the

control group did in three trials. The results of the Comprehensive Tests were also striking; on the first Comprehensive Test the keyword group recalled 72% of the total vocabulary whereas the control group recalled only 46%. Six weeks later the keyword group recalled 43% of the words and the control group recalled 28%. These are indeed large differences and highly significant statistically.

This study was one in a series of laboratory experiments that demonstrated the effectiveness of the keyword method [3]. The most dramatic demonstration involved a similar experimental design using a Spanish vocabulary. The principal difference was that the control group was told to use a rote rehearsal procedure when studying items. None of the control subjects objected to the rehearsal procedure or found it unnatural, but on a comprehensive test they recalled only 28% of the words. The keyword group recalled 88%. In the Russian experiment described above, the control subjects were highly motivated to do well and were encouraged to use whatever strategies they thought would be most effective. The observed difference between the keyword and control subjects was not a matter of motivation; both groups were highly motivated and attentive to the task.

These results encouraged us to study the keyword method in the less controlled and more complicated setting of the classroom. The first quarter of Stanford's second-year Russian course appeared to be ideal for a variety of reasons. First of all, Russian is a particularly difficult language. The beginning student must learn a grammatical structure that differs radically from English. In addition, vocabulary acquisition is complicated by the fact that there are few cognates in the basic vocabularies of English and Russian. These two problems combine to force a "budgetary crisis" with regard to commitment of the students' time and attention: under normal classroom circumstances the student cannot be expected to master Russian grammar in one year and at the same time develop a broad vocabulary. As a result a compromise is struck in which the student is introduced to as much grammar as possible during the first year, but the range of vocabulary is comparatively small. At Stanford, as elsewhere, the acquisition of a wide-ranging vocabulary is put off until the second year after the student has acquired a sufficient knowledge of grammar.

There were other reasons for testing the keyword method in second-year Russian. We knew the extent of the students' vocabularies fairly well; moreover, we had access to the classroom word lists used in the second-year course. Knowing the "classroom vocabulary," we could construct an additional vocabulary (a "trace vocabulary") that would be (a) unfamiliar to the student, (b) not taught in the regular course, and (c) similar in frequency of occurrence to the vocabulary being learned in the classroom. The computer curriculum involved both the classroom and trace vocabularies.

In the study reported here a variant of the keyword method was used.

This variant, called the *free-choice procedure*, permits a student to request a keyword only when desired. The student sits before a computer console, hears a Russian word through headphones, and simultaneously studies the English translation on a display scope. If the student wishes to see a keyword, he presses an appropriate key on the computer console, and a keyword appears on the scope alongside the translation.

A vocabulary of 675 words was used in the present study, divided into twenty-seven 25-word subvocabularies for presentation over a nine-week period. The experiment used the same computer apparatus described in the Russian experiment mentioned above. Each week involved four sessions with the computer. The first three sessions were *study sessions*; on each study session a completely new list of 25 words was presented for study and test. The fourth session was a *review session* (also called a *weekly review*); this session involved a review of the 75 words presented on the preceding three study sessions. The cluster of three study sessions followed by a review session made up a *study week*. There were nine such study weeks, each week involving a new vocabulary of 75 words.

METHOD

Subjects and equipment.

Thirteen Stanford University students participated (7 males and 6 females). Each student spoke English as the native language, and had attended the first quarter and was currently enrolled in the second quarter of the second-year Russian course at Stanford University. A detailed account of the computer system, visual display devices, and the audio setup is given in Atkinson and Raugh (1975).

Stimulus material

A vocabulary of 675 Russian nouns, verbs, adjectives, adverbs, and other parts of speech, with associated keywords was selected for the programmed vocabulary. Eighty percent of the items (540 words) were derived from the second-year Russian classroom word lists. Words were taken directly in the order of their occurrence on the word lists; only the perfective form of certain verbs was not used [4]. The 540 words are referred to as the *classroom vocabulary*. The classroom vocabulary was divided into 27 sublists of 20 words each and named in order: "classroom sublist 1" through "classroom sublist 27." Thus, classroom sublist 1 contained words selected from the first classroom word lists, and classroom sublist 27 contained words taken from the last lists.

The remaining 20% of the vocabulary (135 words) are referred to as the *trace vocabulary*, and special constraints were imposed on their selection. The trace vocabulary was composed of middle frequency Russian words that

are not introduced during the first- and second-year 'course in Russian at Stanford; student exposure to these items was limited to the experiment. The trace vocabulary contained 60% nouns, 20% verbs, and 20% adjectives. Like the classroom vocabulary, it was also divided into 27 sublists matched (according to the judgment of the experimenters) in abstractness and imageability; items were distributed so that each sublist of the trace vocabulary contained 3 nouns, 1 verb, and 1 adjective.

The keywords for all vocabulary items were selected by a three-person panel whose members were familiar with Russian and the keyword method. Table I presents a sample of Russian words and the corresponding keywords.

Procedures

During the first session (Session 0) the proctor showed each student how to start the computer program that conducted the curriculum. The program itself explained all of the remaining procedures. After giving instructions on the use of the keyboard and audio headset, the program introduced

TABLE 1

Some Trace Vocabulary Items with Related Keywords

| Russian | Keyword | Translation |
|----------------|---------------|-------------------|
| 1. RYT' | [rich] | TO DIG |
| 2. ZHELUDOK | [low duck] | STOMACH |
| 3. OVJOS | [adios] | OATS |
| 4. STUPIT' | [stupid] | TO STEP |
| 5. TAPOCHKI | [top] | SANDALS, SLIPPERS |
| 6. MNITEL'NYJ | [miniature] | PARANOID |
| 7. VALENKI | [vile inky] | FELT BOOTS |
| 8. STYDLIVYJ | [stud levi] | BASHFUL |
| 9. MSTITEL'NYJ | [a bit steep] | VENGEFUL |
| 10. VALIT' | [vile leech] | TO DUMP, DROP |
| 11. TARAKAN | [tar a can] | COCKROACH |
| 12. TRESKA | [police car] | COD |
| 13. BARAN | [Ron] | RAM |
| 14. PETUX | [pick tooth] | ROOSTER |
| 15. BOBR | [pauper] | BEAVER |
| 16. LIFCHIK | [lift cheek] | BRA |
| 17. JOZH | [gauche] | HEDGEHOG |
| 18. KLOP | [whop] | BEDBUG |
| 19. SOSNA | [so small] | PINE |
| 20. DERZKIJ | [dares't you] | IMPERTINENT |
| 21. MAZAT' | [Ma's itch] | TO RUB |
| 22. ZHADNYJ | [shot me] | AVARICIOUS |
| 23. TERPET' | [tear page] | TO ENDURE |
| 24. NAGLYJ | [an ugly] | IMPUDENT |

keywords as a means of comparing and contrasting the sounds of English and Russian words; see Atkinson and Raugh (1975) for details of this procedure.

After the keyword practice, written instructions (reproduced in Appendix A) were given on learning methods. They explained that while a Russian word was being pronounced, a keyword (or keyword phrase) would be displayed in brackets at the left-hand margin of the screen and the English translation would appear to the right. Students were instructed to learn the keyword first and then picture an imaginary interaction between the keyword and the English translation; the instructions also stated that if no such image came to mind, the student could generate a phrase or sentence incorporating the keyword and translation in some meaningful way.

The presentation of instructions was followed by a practice series of ten Russian words; each Russian word was spoken while the English translation and appropriate keyword were displayed. Following the practice series a test trial occurred on which each Russian word was spoken and the student attempted to type the English translation. A second study trial was given, followed by a second test trial; this concluded the introductory session (Session 0). The students were told that the next four sessions would be similar to the practice session but that beginning with Session 5 the keyword would appear only when requested. Students were asked to schedule their next four computer sessions in advance. They were informed that each session would require 30 to 50 minutes. They were told not to take more than one session per day, and to complete the four days within a week. Except for these two constraints, students were permitted to set their schedules as they wished.

On Session 1, the computer program composed a 25-word study list by adding to the first (20-word) classroom sublist a (5-word) trace sublist selected randomly for each subject. Session 1 consisted of five successive study/test trials. A study trial consisted of randomized presentation of the 25-word study list; each Russian word was pronounced three times while the keyword and translation were presented on the display scope. For the first and second study pass the presentation was timed for 10 seconds per item; for the third, fourth, and fifth study passes the presentation was timed for 7 seconds. A test trial consisted of a randomized presentation of the 25-word study list: each word was pronounced three times without any visual display. The student was allowed 7 seconds to respond. If a single letter was typed within 7 seconds, the time period was extended to 9 seconds; if the student typed the first two letters of the translation correctly, the program automatically completed the word on the display scope, but if the first two letters were incorrect the program erased the scope and advanced to the next test item without feedback. Throughout the vocabulary curriculum, the same timing, two-letter response convention, and randomized presentation procedures were followed on a test trial.

Sessions 1, 2, and 3 followed identical formats. The only difference was that each day involved a completely new set of vocabulary items: the classroom vocabulary progressed in the same manner for each student (sublist 2 in Session 2, sublist 3 in Session 3), however, the presentation order of trace sublists was randomized for each student, so that, for example, in Session 2 one student might receive trace sublist 8 while another student might receive trace sublist 20. The random selection algorithm insured that each student received each trace sublist on exactly two days (first on some study session, and next on the associated weekly review session). The selection algorithm also insured that each trace sublist appeared equally often (for different students) during each week of the quarter.

Session 4 was a weekly review. The review began with a test trial that covered the 75 words presented in the preceding three sessions. The test was followed by a study trial over the 75 words, with presentation timed at 10 seconds per item. The first test pass and study pass were followed by randomized repeats of the test pass, the study pass, and, finally once again, the test pass.

Sessions 5 through 8 were identical to Sessions 1 through 4 except that keywords were no longer displayed automatically. When a word was pronounced for study, only the English translation was displayed. If the student wished to see the associated keyword, he could press an appropriate key on the keyboard, and the keyword appeared in brackets to the left of the English translation. Students were told that they were not obliged to use the keyword method and that the possibility of requesting a keyword existed only for their convenience.

The procedures established during the second week (Sessions 5 – 8) were maintained throughout the remaining weeks. A subject could complete the curriculum in 8 to 10 weeks, depending upon individual scheduling.

Within one week after completing the experiment, students were tested for recall of the 135 trace vocabulary items. The test was conducted on-line using the same testing procedures that had been used throughout the vocabulary curriculum. Upon completing the test, the same 135-words test was immediately tested again with a new randomized presentation order.

RESULTS

Much of the analysis is based upon the probability of a keyword request on a study trial and the probability of a correct response on a test trial. Table 2 introduces a nomenclature for discussing trials and the associated probabilities. When an item is first presented on a study session it receives five study-test cycles denoted as S1, T1, S2, T2, . . . S5, T5. The item is again presented in the review session for three additional tests and two studies in the order T6, S6, T7, S7, T8. Finally, all trace items are tested

TABLE II

Diagram Giving Nomenclature for Discussing Study and Test Trials, and Associated Probabilities

| | Study Session | | | Review Session | | | Comprehensive Test | | | |
|-------------------|---------------|----|----|----------------|----|----|--------------------|----|----|-----|
| Trial No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Trial Type | S1 | S2 | S3 | S4 | S5 | T6 | T7 | T8 | T9 | T10 |
| Observed Variable | k1 | k2 | k3 | k4 | k5 | c6 | c7 | c8 | c9 | c10 |

twice at the end of the 8- to 10-week period, and these presentations are denoted as T9 and T10 (also referred to as the Comprehensive Test).

The entry c_i of Table II refers to the estimated probability of a correct response on the i -th test trial. Similarly k_i refers to the estimated probability of a keyword request on the i -th study trial. Week 1 (Sessions 1, 2, 3, and 4) was a practice week with special instructions, and consequently those data are disregarded in the computation of c_i and k_i .

An important constraint of the vocabulary curriculum was that the classroom vocabulary was presented synchronously with the ongoing classwork. The trace vocabulary was selected to provide a small core of items that could be presented in randomized order and varied over students. Accordingly, the results reported in this study are primarily based upon analyses of trace vocabulary data [5].

Figure 1 presents results for the trace vocabulary averaged over weeks (with Week 1 excluded). The curve labelled $p(k)$ gives the probability of a keyword request as a function of study trials and the curve labelled $p(c)$ gives the probability of a correct response as a function of test trials. The curve $p(c)$ shows how performance improves with practice, falls off with disuse, and refreshes with review. The increase in $p(c)$ from T9 to T10 is explained by the fact that on T9 students were being tested on items most of which they had not rehearsed for several weeks; many items that were “on

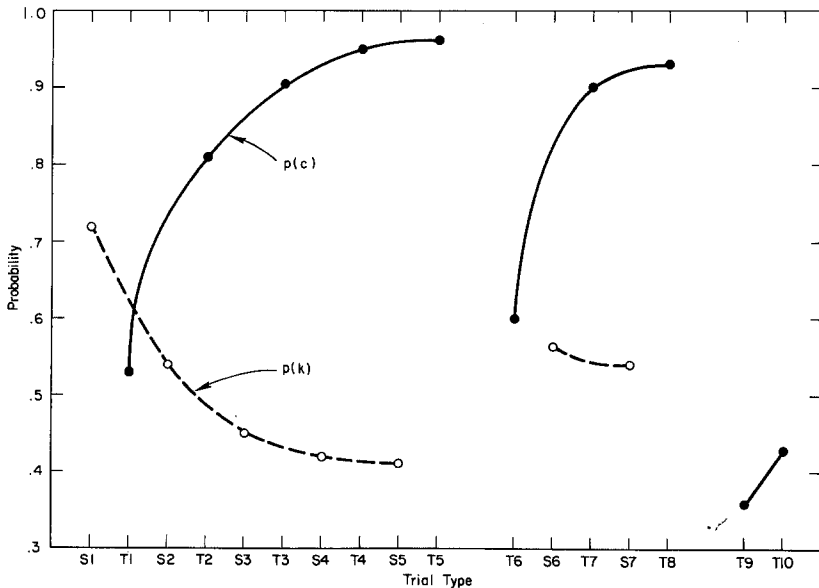


Figure 1. Probability of a keyword request and probability of a correct response as a function of trial types. Data are averaged over trace vocabulary items.

TABLE III

Probability That a Subject Requested a Keyword for an Item on the n^{th} Study Trial Given That on the Preceding Test He Either Missed the Item (\bar{c}_{n-1}) or Was Correct (c_{n-1})

| | Trial Number (n) | | |
|------------------------|------------------|-----|-----|
| | 2 | 3 | 4 |
| $p(k_n \bar{c}_{n-1})$ | .59 | .49 | .42 |
| $p(k_n c_{n-1})$ | .47 | .44 | .42 |

the tip of the tongue" but not recalled on T9 were recalled on T10. Data for individual subjects comparable to the averages in Figure 1 are presented in Appendix B.

A comparison of $p(k)$ and $p(c)$ in Figure 1 shows an inverse relationship between test proficiency and probability of a keyword request, indicating that subjects are most likely to request a keyword when studying an item they do not know. An item analysis reveals that a keyword request was more probable if the student had missed the word on the preceding test trial than if he had been correct. Table III gives the results for S2 through S4. For example, if a student responded incorrectly to a word on T1, then with probability .59 he requested a keyword for that item on S2; however, if he was correct, then the probability of requesting a keyword was only .47.

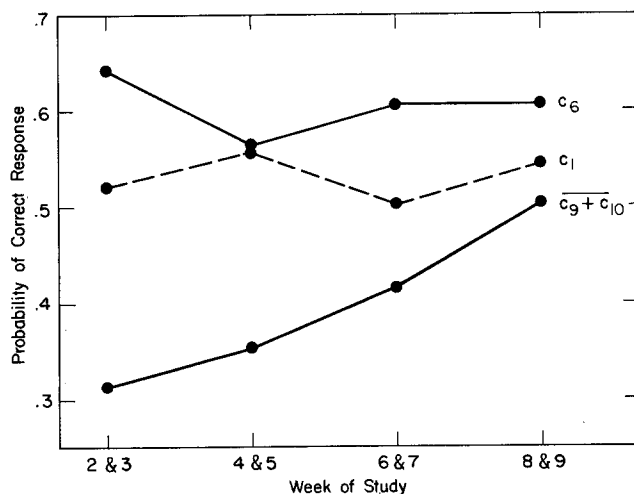


Figure 2. Probability of a correct response on test trials T1, T6, and (T9+T10) as a function of the study week for a vocabulary item. Data are averaged over trace vocabulary items.

Raugh and Atkinson (1975) report similar results in an experiment involving free-choice.

Figure 2 presents the trace vocabulary data as a function of study week. For example, the curve labelled c_1 gives the probability that a student responded correctly to an item on the first test trial as a function of study week (week-pair). Similarly, curve c_6 gives the probability that a student responded correctly to an item on the first weekly review test as a function of the week when the item was studied. Neither curve c_1 nor c_6 varies significantly from week to week, indicating that performance does not change over weeks. The curve labelled $\overline{c_9+c_{10}}$ gives Comprehensive Test results averaged over T9 and T10, categorized by study week. For example, performance on words learned during Weeks 2 and 3 is .31, whereas performance on words learned during Weeks 8 and 9 is .50. A positive recency effect is expected; words learned in the later sessions should be better recalled on a comprehensive test than those learned in the early sessions (Atkinson and Shiffrin, 1968).

Figure 3 shows the probability of a keyword request for trace vocabulary items as a function of study trial and the week of study. The data presented in Figure 3 are similar to those in Figure 2. Thus, for example, the

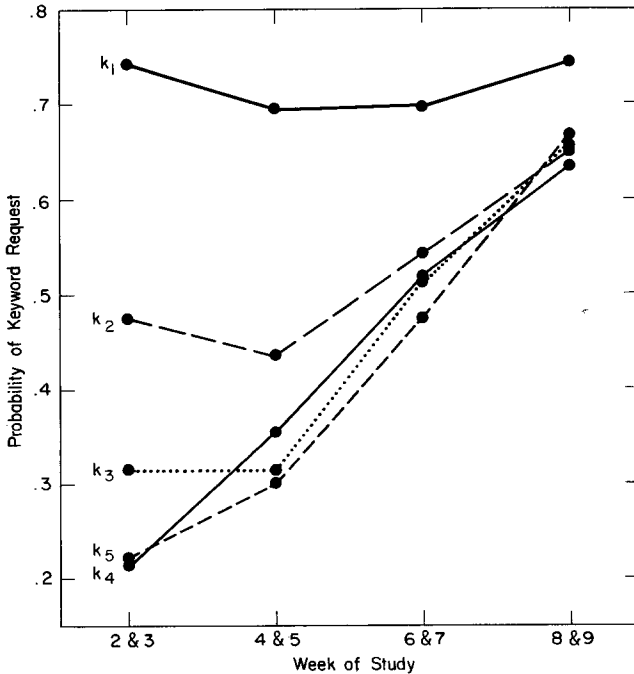


Figure 3. Probability of a keyword request on the first five study trials as a function of the study week for a vocabulary item. Data are averaged over trace vocabulary items.

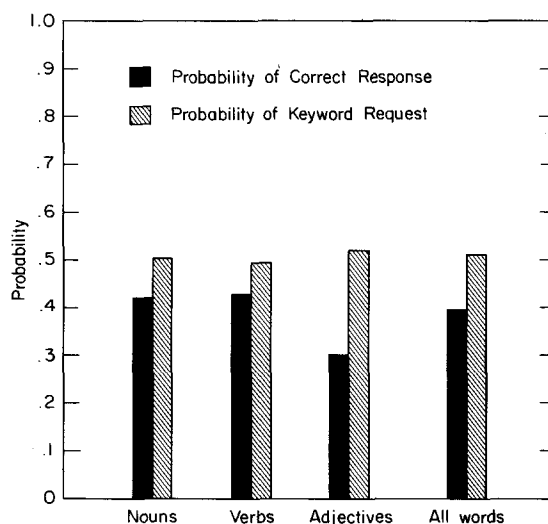


Figure 4. Probability of a correct response averaged over test trials T9 and T10 and the probability of a keyword request averaged over study trials S1 through S5 both as a function of grammatical type. Data are averaged over trace vocabulary items.

curve labelled k_1 gives the probability of a keyword request on the first study trial (S1), categorized by week of study. Curve k_5 gives the probability of a keyword request on S5 as a function of the week of study. Note that k_1 remains high throughout the quarter (average $k_1 = .72$, see Figure 1), with no significant variation from week to week. A surprising result is that on subsequent study trials, keyword requests increased dramatically over weeks. For example, consider k_5 ; during Weeks 2 – 3, the probability that a subject requested a keyword on the fifth study trial was .21, whereas during Weeks 8 – 9 the probability was .63.

Although the trace vocabulary was composed of different numbers of nouns, verbs, and adjectives, these three grammatical classes were distributed evenly throughout the trace sublists. Thus, each student on any given study day received 3 nouns, 1 verb, and 1 adjective, as trace vocabulary items. Figure 4 presents the average of c_9 and c_{10} and the average of k_1 through k_5 as a function of grammatical class; note that data are averaged over all trace vocabulary items presented during Weeks 2 through 9. Keyword requests did not vary significantly as a function of class. Although test results for nouns and verbs were comparable, performance on adjectives was somewhat poorer. A possible explanation may be related to the fact that the adjectives in the trace vocabulary were substantially longer in terms of syllable count than nouns and verbs (nouns averaged 2.20 syllables per word, verbs averaged 1.75, and adjectives 3.33). During interviews students often stated that the keyword method is most easily applied to words of concrete meaning. Since the adjectives chosen for the trace vocabulary were mostly

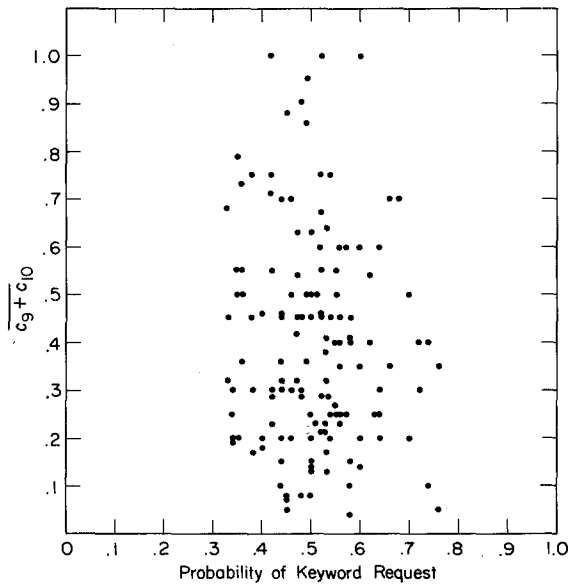


Figure 5. Scatter plot of performance on the Comprehensive Test (average of c_9 and c_{10}) versus probability of a keyword request during initial study (average of k_1 through k_5). Each point corresponds to a specific trace vocabulary item and represents an average over subjects.

qualitative (not relative) it could be argued that they were on the average more abstract than the nouns and verbs. Thus, the disparity may be an effect of either abstractness or word-length; we cannot say whether the poorer performance on adjectives reflects an inherent problem with adjectives.

Figure 5 presents a scatter plot in which each point represents performance on a trace word averaged over all students. The abscissa gives the probability of a keyword request (averaged over study trials 1 – 5) and the ordinate gives the probability of a correct response to the item on the Comprehensive Tests (averaged over T9 and T10). Note that there is considerably more variation in test performance than in keyword requests. The trace words are presented in different orders for different subjects; computing an average over subjects for a given trace word means that different subjects saw that item on different study weeks. Thus, variation in the scores presented in the scatter plot does not reflect variations due to week of study. The correlation coefficient for the scatter plot is .13.

Table IV presents the conditional probability that a student responded correctly to an item on a test trial given that he responded incorrectly to the item on the preceding test trial, for T2 through T5. The probabilities do not differ significantly from one another over trials and lend support to the hypothesis that vocabulary learning satisfies an all-or-none process (Atkinson, Bower, Crothers, 1965, p. 105). This finding supports earlier

TABLE IV

The Probability a Subject Responded Correctly to an Item on Test Trial n Given That He Responded Incorrectly to the Same Item on Test Trial $n-1$, Namely, $p(c_n|\bar{c}_{n-1})$

| Test Trial n | |
|--------------------|----------|
| Probability | Estimate |
| $p(c_2 \bar{c}_1)$ | .69 |
| $p(c_3 \bar{c}_2)$ | .66 |
| $p(c_4 \bar{c}_3)$ | .68 |
| $p(c_5 \bar{c}_4)$ | .66 |

studies on vocabulary learning reported by Atkinson and Paulson (1972) and Atkinson (1972).

DISCUSSION

Previous laboratory studies have shown that the keyword method is a highly effective procedure for learning a foreign language vocabulary. The concern of the present study was to determine how Russian language students would behave in the context of a free-choice vocabulary curriculum. Part of the answer is reflected in Figure 1; students, given a free choice, elect to use the keyword method frequently. It is apparent from Figure 3 that the effect does not depend upon novelty, since the probability of a keyword request on the first study contact remains near an average of .72 throughout the nine weeks of the experiment. In fact, as can be seen from Figure 3, total keyword requests increase over the period of the study.

Why do subjects request keywords? In a carefully controlled experiment, Raugh and Atkinson (1975) analyzed each word with respect to (a) the number of keyword requests made, (b) the subject's recall of the word on a delayed test, and (c) the "difficulty" of the word. Difficulty was defined as the probability of an error for the word, using an independent group of subjects who did not use the keyword method. They found that keyword requests were positively related to word difficulty, but negatively related to a subject's recall of the item; also, a keyword request was more probable when a subject missed the same word on the preceding test trial than if he was correct. A small number of items were easily learned without the keyword method using cognates and other special features, but in most cases the keyword method was employed.

In the study reported here students used keywords in a similar fashion. On early study trials they were more likely to request a keyword for an item they had missed on the preceding test trial than if they had been correct.

Thus, keywords are requested selectively as a learning aid.

Keyword requests may be one part of a learning process that includes many other strategies. In self-reports, [6] students have described the main alternatives: rote rehearsal, recognition of a cognate, and identification of a familiar root. When a student encounters an item for which there is no obvious alternative to the keyword method, he requests the keyword. If the word is not mastered immediately, the keyword is requested on subsequent trials until either the word is learned by the keyword method, or by the discovery of a non-obvious cognate or root; failing these, the word may finally be learned by rote rehearsal. Thus, a keyword request can be regarded as an exploratory action indicating that the subject is seeking a way to learn an item; the subject tries various approaches and frequently succeeds by using the keyword method.

Students differ in the emphases they give to various learning strategies. One of our students rarely used the keyword method. Another student always used the keyword method. Most students, however, first sought a familiar root, and if not successful then reverted to the keyword method. The high incidence of keyword requests reflects the extent to which students failed to learn an item by an alternative method; in a sense, the keyword method succeeds by default.

In deciding whether to use the keyword method, several problems need to be considered. One problem is that keywords might interfere with correct pronunciation. Our experiments do not deal with the issue, but we have discussed it with a number of experts on language instruction. Although opinions vary, most believed that the keyword might well facilitate, rather than interfere, with pronunciation. The keyword method has features in common with the method of contrasting minimal pairs — a common technique for teaching phonetics by contrasting words that differ slightly in pronunciation. Further, if the practical use of a language is the principal goal, then effective vocabulary-learning methods should be used even if they do interfere with pronunciation. Another problem to be considered in using the keyword method is whether items learned in this way will be retrieved more slowly, particularly once the item has been thoroughly mastered. We have little direct evidence on this point, but our experience with the method suggests that it should not be a problem. Once an item has been learned thoroughly, it comes to mind immediately; rarely is the learner aware of the related keyword unless he makes a conscientious effort to recall it. More experiments need to be done, but introspective reports suggest that the keyword method will not interfere with retrieval once an item has been mastered.

Some evidence suggests that students use mediating strategies similar to the keyword method when learning a vocabulary, even if not instructed to do so. Ott, Butler, Blake, and Ball (1973) in a paper on the use of mental

imagery in vocabulary learning, report that subjects not given special instructions when asked to learn a foreign vocabulary often resort to using English mediating words combined with imagery or other mnemonic aids. Their observation suggests that the keyword method is not essentially different from techniques commonly employed by students. The major difference, apart from the fact that the experimenter supplies the keyword, is the extent to which the method is applied.

Our previous experiments and the demonstration study reported here convince us of the usefulness of a computer-based vocabulary drill employing the keyword method. In all of our studies the majority of subjects have been highly favorable to the keyword method, and have appreciated the drill experience.

In designing a vocabulary drill program two limitations of the present program should be kept in mind. First, our study was conducted without using the written form of Russian words. The reason is that the computer display device could not present Cyrillic script, and we felt that a transliteration system would interfere with the Cyrillic form presented in the classroom. Many students remarked that they would like to have seen the Cyrillic form of the word displayed each time the spoken word was presented for study and test. The second limitation of our program was that many items were presented for repeated study and test long after a student had learned them. Such prolonged practice on items already mastered is inefficient as well as boring and distracting. A more effective application of the keyword method would permit the computer system to monitor the student's performance on each item and systematically drop from further study those items that have been responded to correctly on prior tests. A number of schemes of this type have been examined experimentally and some have proved to be highly effective. For a review of optimal sequencing procedures in vocabulary learning see Atkinson (1976) and Atkinson and Paulson (1972).

Appendix A

Instructions on the Keyword Method Presented During Session 0

A large amount of the time you spend in learning a foreign language is devoted to the learning of the vocabulary. This is especially true of Russian. For this reason it would be worth your time and effort to develop efficient strategies for learning new vocabulary. In experiments conducted over the past two years at Stanford's Institute for Mathematical Studies in the Social Sciences, the Keyword Method, which will be described below, has proved to be a highly effective means of learning foreign language vocabulary. You will probably find it a fascinating and pleasant way to learn, because you create the "tools of the trade" for yourself.

Normally, in the weeks ahead, you will not be constrained to learn by any particular method – you will be free to use whatever method you prefer. But for the first week we want you to employ the Keyword Method exclusively, until you are thoroughly practiced in its use. We think that you will find this initial training period to be valuable.

In all of the study that follows, you will have Russian words presented to you, one at a time. Each word will be pronounced three times while its English translation is displayed on the screen. For the first week only, the keyword will be automatically displayed between brackets to the left of the English translation. (Remember that keywords are derived from the *sounds* of Russian words and have nothing to do with their meanings.) After a Russian word has been pronounced, the visual display will continue for a short time, then the program will advance to the next item. Beginning the second week, keywords will not be displayed unless you request them.

Here is how the Keyword Method works. When a keyword is displayed with the English translation, the computer will pronounce the appropriate Russian word three times (this period of time is called the “pronunciation phase”), then allow a brief pause for quiet study (the “quiet phase”):

DURING THE PRONUNCIATION PHASE, CONCENTRATE ON LEARNING THE KEYWORD

DURING THE QUIET PHASE, ASSOCIATE THE KEYWORD WITH THE ENGLISH TRANSLATION BY USING “MENTAL IMAGERY.” Do this by visualizing an imaginary situation in which the keyword and the translation “interact” in some graphic way. The image can be as wild and absurd as you like, in fact some people say the wilder the better. The point is to make the image vivid and memorable.

For example, suppose the following keyword and translation appeared on your screen:

[OAK] BELL

The computer would first pronounce the Russian word (which sounds somewhat like “zvahn-oak,” accent on the last syllable), then allow a pause for quiet study. During the quiet phase, you should imagine an interaction between an oak and a bell. Following are some examples of what you might imagine:

1. An oak tree in a belfry.
2. An oak tree with little brass bells for acorns,
3. An oak tree growing beneath a giant bell jar.

Any of these images could help you to recall that OAK was paired with BELL. These images are simply suggestions; it would be better for you to create your own image to suit your own taste. You will find that after some practice it is usually easy to create such images. And no matter how silly they may seem at first, images are powerful memory aids. So take advantage of your innate ability to recall imagery!

In brief, for your first week of study the strategy you should employ for learning a translation is to

FIRST, during the pronunciation phase, learn the keyword.

SECOND, during the quiet phase, create a distinctive mental image in which the keyword and the translation interact in a graphic way. For this interaction, stick to one good “picture” — do not confuse yourself by imagining more than one interaction. Then, later, when you hear the Russian word, you will think of the keyword and the image, which will in turn remind you of the English translation.

As a second example, consider the Russian word for building; it sounds somewhat like “zdawn-yeh” (accent on the first syllable). Suppose the following appeared on your screen:

[DAWN] BUILDING

While the computer is pronouncing “zdawn-yeh” three times, you should concentrate primarily on learning the keyword. After the computer has completed the pronunciation, you should then create an image relating DAWN to BUILDING. For example, imagine dawn, when the city skyline is tinged with pink, with the early morning sun reflected in the windows of a building; or picture dawn in the desert with a single incongruous building (such as a skyscraper) standing in the cool morning air. Whatever you choose to visualize, make the scene as distinct and vivid as possible.

Most of the words you will study in this curriculum are nouns. However, many will be verbs and adjectives. You shouldn’t have trouble keeping things straight. Most verbs connote action, and you can easily picture an action. Likewise, most adjectives connote a quality that is easily visualized. To make the matter clear, when a Russian verb is presented by the computer, the English translation will be displayed with a “V” placed in parentheses to its right, so you will never have trouble distinguishing noun forms from verbs.

For example, if the Russian word for “to crawl” were spoken into the earphones, you would see the following displayed on your screen:

[PULSE] CRAWL (V)

and you would not have to wonder whether the noun or verb form were being presented. You could immediately begin to imagine, say, the exaggerated pulse of an earthworm in the act of crawling.

As another example, if the Russian word for “courteous” were being pronounced, you might see the following display:

[USE THE AIRPLANE] COURTEOUS

There is no problem because the word is clearly an adjective. To transform this into imagery, you might imagine an extremely courteous porter beckoning you towards the on-ramp, saying “use the airplane.”

You will find that some translations or keywords are abstract and consequently not easy to picture directly. It is often possible in such cases to think of symbolic imagery; in the example of “courteous” given above, you might picture the porter bowing deeply as he directs you to the airplane. The easily visualized act of bowing then becomes a symbolic reminder of the fact that the porter is courteous. As another example, to visualize “thought,” you might picture some thoughtful person you know scratching his head.

As the example of “use the airplane” implies, you will meet examples of keyword “phrases.” You might think at first that a keyword phrase would be more difficult to visualize than just a single keyword. But, in fact, a keyword phrase or an exclamation, such as GEE WHIZZ or USE THE AIRPLANE, can be highly effective. When you are confronted with a keyword phrase, all you need to do is to imagine a situation in which the phrase or exclamation is appropriate, then exaggerate the situation to make it memorable.

If you have trouble thinking of imagery to relate a particular keyword and translation, you could resort to a phrase or sentence that connects the keyword to the translation. In fact, many persons who are not used to using mental imagery as a memory aid start out by thinking that phrases are “more natural” than imagery. For example, suppose the keyword were GOD and the translation were INCH; in this case you might well have difficulty thinking of a suitable image, whereas you might easily think of a phrase like “pull GOD an INCH” or a sentence like “GOD doesn’t budge an INCH.” If you stall over the imagery, then such a phrase or sentence can be useful as an alternative memory aid. But remember: there is a certain amount of skill involved in making up imagery, and while it may seem “unnatural” or difficult at first, it gets easier with practice. As mentioned before, mental imagery is a powerful memory aid, and it would probably be worth your time to develop the skill.

As a final bit of advice on keywords, note that if an occasional keyword sounds a little “out of key” to your ear, and a better keyword occurs to you, then use your own. But keep in mind that the keyword you choose must be easy to remember and easy to visualize.

Finally, you must realize that the Keyword Method is not a total solution to the problem of learning foreign language vocabulary. The fact is that fluent usage comes only as a result of much practice in the idiom. The

Keyword Method is merely a means of assisting you to develop your RECOGNITION of a lot of words quickly; it is simply your first contact with the new vocabulary. To get the full benefit of the method, you must constantly deepen your exposure to the vocabulary in various grammatical and conversational exercises, and in reading. For it is only by hearing, seeing and using the words in context that you gain full control of them.

After you are satisfied that you understand this introduction you may return to the computer console, and start up the program the way the proctor showed you (you will always begin in this way). You will then receive a short practice session on 10 Russian words to get a feel for the procedures for the next four sessions. As mentioned at the outset, keywords will be presented automatically only during the first week. Thereafter, they will be displayed only when you request them by a simple procedure that will be explained later.

Appendix B

Individual Subject Data on Several Indicators Averaged Over Trace Vocabulary Items and Weeks 2 Through 9

| Subject | k_1 | $\overline{k_1+\dots+k_5}$ | c_1 | $\overline{c_1+\dots+c_5}$ | c_6 | c_9 | c_{10} |
|---------|-------|----------------------------|-------|----------------------------|-------|-------|----------|
| 1 | .10 | .21 | .50 | .81 | .59 | .50 | .56 |
| 2 | .92 | .22 | .59 | .87 | .73 | .37 | .39 |
| 3 | .81 | .67 | .84 | .94 | .71 | .48 | .53 |
| 4 | .94 | .70 | .32 | .61 | .29 | .21 | .28 |
| 5 | .86 | .66 | .38 | .76 | .60 | .33 | .34 |
| 6 | 1.00 | .64 | .44 | .76 | .39 | .21 | .32 |
| 7 | .33 | .24 | .47 | .82 | .61 | .26 | .32 |
| 8 | .90 | .64 | .47 | .78 | .69 | .43 | .50 |
| 9 | .96 | .79 | .62 | .89 | .57 | .49 | .57 |
| 10 | .95 | .80 | .43 | .85 | .63 | .37 | .45 |
| 11 | .20 | .29 | .69 | .90 | .71 | .43 | .58 |
| 12 | .76 | .22 | .54 | .83 | .63 | .26 | .36 |
| 13 | .67 | .54 | .57 | .81 | .67 | — | —* |

* Note: Student 13 did not take the Comprehensive Test.

Acknowledgements

This research was supported by the Office of Naval Research, Contract no. N00014-67-A-0012-0054, and by Grant MH-21747 from the National Institute of Mental Health. The authors wish to thank Professor Joseph A. Van Campen of the Department of Slavic Languages and Literatures at

Stanford University for advice on problems of vocabulary acquisition in second-language learning.

Notes

- 1 Printed Russian words are presented in a standard transliteration of the Cyrillic alphabet into the Roman alphabet.
- 2 This method can be modified to produce a variety of related learning strategies by changing the ways in which the two links are formed. For example, instead of using an acoustic link, one could use an *orthographic link* by selecting the keyword based on a similarity of spelling rather than of sound (thus, "ball" might be used as a keyword for *caballo*).
Or the mnemonic link could be based upon a verbal construct (rather than a mental image) involving a sentence whose subject is the keyword and whose object is the English translation.
- 3 For a review of these studies see Atkinson (1975).
- 4 At least during the first two years of study of Russian the student must learn two separate vocabulary items for each single verb lexeme, i.e., the forms of the imperfective and perfective aspects. In most cases the two forms differ only in the presence of a prefix and/or suffix, e.g., *sdelat'* (perf)/*delat'* (imp) "to do, make," *napisat'* (perf)/*pisat'* (imp) "to write," etc. The teaching of such minor differences through the keyword method is difficult because it involves the teaching of two separate keywords for forms that are similar. For this reason it was decided to teach only the imperfective form of the verb in those cases where the two aspects differ in terms of an affix alone. Where an aspectual pair differs substantially in form both aspects are taught as individual vocabulary items, e.g., *skazat'* (perf)/*govorit'* (imp) "to speak, say," *vzjat'* (perf)/*brat'* (imp) "to take."
- 5 Test performance on classroom items was slightly higher than on trace items (some of the students worked ahead of their classroom assignments and were familiar with a portion of the classroom words); correspondingly, the probability of a keyword request on classroom items was slightly lower than for trace vocabulary items. For example, overall $k_1 = .64$, trace $k_1 = .72$; and, overall $c_1 = .59$, trace $c_1 = .53$.
- 6 For a detailed analysis of self-reports in an earlier study, see Raugh and Atkinson, 1975.

References

- Atkinson, R. C. "Ingredients for a theory of instruction." *American Psychologist*, 1972, 27, 921 – 931.
- Atkinson, R. C. Mnemotechnics in second-language learning. *American Psychologist*, 1975, 30:821–828.
- Atkinson, R. C. "Adaptive instructional systems: Some attempts to optimize the learning process." In D. Klahr (Ed.), *Cognition and instruction*. Hillsdale, N. J.: Erlbaum, 1976.
- Atkinson, R. C., Bower, G. H. and Crothers, E. J. *An Introduction to Mathematical Learning Theory*. New York: Wiley, 1965.
- Atkinson, R. C. and Paulson, J. A. "An approach to the psychology of instruction." *Psychological Bulletin*, 1972, 78, 49 – 61.

- Atkinson, R. C. and Raugh, M. R. "An application of the mnemonic keyword method to the acquisition of a Russian vocabulary." *Journal of Experimental Psychology*, 1975, 104, 126 – 133.
- Atkinson, R. C. and Shiffrin, R. M. "Human memory: A proposed system and its control processes." In K. W. Spence and J. T. Spence (Eds.), *The psychology of learning and motivation*. Vol. II. New York: Academic Press, 1968.
- Holley, F. M. "A study of vocabulary in context: The effect of new-word density in German reading materials." *Foreign Language Annals*, 1973, 6, 339 – 347.
- Hughes, J. P. *Linguistics and Language Teaching*. New York: Random House, 1968.
- Ott, C. E., Butler, D. C., Blake, R. S. and Ball, J. P. "The effect of interactive-image elaboration on the acquisition of foreign language vocabulary." *Language Learning*, 1973, 23, 197 – 206.
- Raugh, M. R. and Atkinson, R. C. "A mnemonic method for learning a second-language vocabulary." *Journal of Educational Psychology*, 1975, 67, 1 – 16.
- Yates, F. *The art of memory*. Chicago: University of Chicago Press, 1972.