International Review of Applied Linguistics, 2, 1964, pp. 205-217.

Sonderdruck aus IRAL Vol. II/3

LATENCY PHENOMENA IN PROLONGED LEARNING OF VISUAL REPRESENTATIONS OF RUSSIAN SOUNDS

Edward Crothers, Patrick Suppes, and Ruth Weir
Institute for Mathematical Studies in the Social Sciences
Stanford University

Julius Groos Verlag Heidelberg

LATENCY PHENOMENA IN PROLONGED LEARNING OF VISUAL REPRESENTATIONS OF RUSSIAN SOUNDS [1])

Edward Crothers, Patrick Suppes, and Ruth Weir
Institute for Mathematical Studies in the Social Sciences
Stanford University

Un aspect important de l'apprentissage d'une langue étrangère est la capacité de reconnaitre rapidement les représentations visuelles de mots présentés de vive voix dans la langue à apprendre. Pour étudier cet aspect, on a fait appel à un plan d'apprentissage par anticipation en utilisant la méthode des associations par paires. Pour chaque question ("item"), les stimuli étaient constitués par des mots russes enregistrés par un sujet dont la langue maternelle est le russe. Les réponses possibles, en caractères cyrilliques, étaient présentées sous forme de choix multiples. En entendant un mot, le sujet appuyait sur un bouton qui correspondait à un des choix écrits et apprenait aussitôt quelle était la réponse juste. Comme variable indépendante on a classé les questions russes selon la ressemblance de leur orthographe, d'une part entre elles, d'autre part par rapport aux caractères romains.

Un des buts principaux de cette recherche a été de déterminer la courbe des temps de réaction en fonction du nombre d'essais, étant donnè l'exactitude des réponses et le genre des questions. Les résultats de quinze séances d'apprentissage, avec des étudiants d'un niveau de propédeutique, furent les suivants:

1. Les temps de réaction aux questions en russe se stabilisaient progressivement à un chiffre plus élevé que pour les questions en anglais, bien qu'il n'y ait eu presque plus d'erreurs à partir de la cinquième séance.

2. La nature de la réponse et le degré d'entrainement déterminaient les temps de réaction. Ces temps étaient plus longs pour les erreurs que pour les choix corrects, et ils continuaient à se raccourcir après que l'apprentissage fut presque parfait.

3. Selon la catégorie auxquelles elles appartenaient, les questions déterminaient aussi les temps de réaction. Par exemple, les temps pour des questions utilisant des caractères cyrilliques semblables l'un à l'autre mais distincts des caractères anglais étaient plus longs que pour des questions comportant un caractère qui correspondait à des sons entièrement différents en russe et en anglais.

Ein wichtiger Aspekt des Fremdsprachenlernens ist die Fähigkeit, visuelle Darstellungen von Wörtern schnell zu erfassen, die mündlich in der zu lernenden Sprache geboten werden.

Um diesen Aspekt zu untersuchen, hatten Vfs. einen Plan vorwegnehmendes Lernen aufgestellt, der das Verfahren der Wortpaarassoziation benutzt. Für jede Frage ("item") bestanden die Stimuli aus russischen Wörtern, die von einem ge-

[1]) The research reported herein was conducted pursuant to contract SAE 9514 between Stanford University and the U.S. Office of Education, and was also partially supported by a grant from the Carnegie Corporation of New York. We wish to acknowledge the assistance of Elise Belenky in preparing and conducting the experiment.

bürtigen Russen aufs Tonband gesprochen worden waren. Nun wurden den Studenten, die alle ungefähr das Bildungsniveau von Abiturienten hatten, mehrere Antworten in kyrillischer Schrift vorgelegt. Sobald ein Student ein Wort hörte, löste er einen Knopf aus, der einer der zur Wahl gestellten schriftlichen Antworten entsprach. Dabei erfuhr der Student, welches die richtige Antwort war.

Eines der Hauptziele dieser Untersuchung bestand in der Bestimmung des Verhältnisses von Reaktionszeit zur Anzahl der Antwortversuche im Hinblick auf die Genauigkeiten der Antworten und die Kategorie der Fragen. Die Ergebnisse im Laufe von 15 Arbeitssitzungen können etwa folgendermaßen zusammengefaßt werden:

1. Die Reaktionszeiten auf Fragen in Russisch stabilisierten sich fortschreitend bei einer Dauer, die viel länger war als die entsprechende Dauer bei Fragen auf Englisch, obgleich vom Ende der fünften Sitzung an fast keine Fehler mehr gemacht worden waren.

2. Die Art der Antwort und die Vollständigkeit des Lernens bestimmten die Reaktionszeiten. Diese waren bei fehlerhaften Antworten viel länger als bei richtigen, und verkürzten sich noch, nachdem das Lernen fast fehlerfrei geworden war.

3. Auch die Kategorie der Fragen beeinflußte die Reaktionszeiten. Beispielsweise waren diese bei Fragen, die in einander ähnlichen, aber vom Englischen verschiedenen, russischen Buchstaben vorlagen, viel länger als bei Fragen, die aus Buchstaben zusammengesetzt waren, die völlig verschiedenen Lauten im Russischen und Englischen entsprachen.

Various writers (e.g., Luce, 1960; Woodworth and Schlosberg, 1956) have noted that much more attention is usually given to response probability than to choice latency*) as a measure of response strength during learning. In free-responding situations, Skinner and his associates have established laws describing response rate as a function of the reinforcement schedule. However, in choice-learning situations with discrete trials, few empirical investigations have been made concerning the ways in which latency depends on stimulus conditions and on choice probability. Also, in second-language learning, latency of response has been recognized as an important component of "fluency", yet with the exception of Lambert's (1955) work, little has been reported.

There are several fairly obvious reasons why latency warrants increased consideration as a measure of learning. First, it might permit detection of differences between stimuli in strength of association to assigned responses when no difference is manifest in the choice data. For example, Musgrave (1962) found latency

* Editor's note: "Response probability" refers to the likelihood of one or another response in the presence of a given stimulus. "Latency" refers to the delay between presentation of a stimulus and the onset of an observable response. In the next paragraph, learning to "criterion" and attaining "asymptote" in a learning task mean, respectively, reaching a pre-established level of correctness and consistency in response, and reaching a level of correct response that does not significantly change with subsequent trials. Such subsequent trials constitute "overlearning".

differences among stimulus classes in paired-associate learning (PAL) after the subjects had reached a learning criterion. Likewise, changes in latency of response to the same stimulus often continue for many trials after the proportion correct has reached asymptote (Brown and Huda, 1961; Millward, 1963).

It is also important to ascertain the shape of the mean latency curve over trials prior to criterion. The question is whether for a given experimental condition the curve declines monotonically, rises, remains flat, or exhibits some other form. If we conceive of the subject as passing from a guessing stage to an information-processing and decision-making stage, and finally to "automatic" responding, latencies might be expected to rise initially. Presumably, the phenomenon would be analogous to VTE behavior*) in rats (Tolman, 1938).

The present experiment sought to discover properties of the latency curves, both for correct and for incorrect choices, in a relatively long-term study of certain aspects of second-language acquisition. On each trial the subject heard a one-syllable Russian word and saw four alternative words typed in Cyrillic characters. His task was to choose the alternative which he thought was the visual representation of the auditorily presented word, and to indicate his choice by pushing the appropriate button. Another question asked was whether after prolonged overlearning, latencies would decline to times observed using English auditory and visual material. Accordingly, the subjects were trained daily for fifteen days and English items were included. As an aid in determining what features of the task critically influence latencies, several different classes of Russian items were presented. The classes differed in inter-item similarity (primarily visual and only secondarily auditory) and in similarity to the symbols of the Roman alphabet.

METHOD

Subjects. The subjects were seven students who had been secured through the Stanford University employment bureau. Each was paid a sum equivalent to $ 1.75 an hour. Only students whose native language was American English and who had no background in Slavic languages were eligible for participation.

Materials. The stimuli were Russian consonant-vowel-consonant (CVC) syllables printed in capital letters. Each item consisted of four CVC'S. There were four classes of items, which were distinguished by (a) the similarity, of the Cyrillic graphemes to Roman graphemes, and (b) intra-item similarity; i.e. presence vs. absence of identical or similar graphemes within the same item. A sample item from each class is shown in Table 1. The authors independently judged the pairs of graphemes as similar or dissimilar. Some arbitrary decisions were required,

*Editor's note : VTE stands for vicarious trial and error and refers to the tentative behavior of looking back and forth often observed in the behavior of rats at the choice points of mazes.

although the judgments almost always concurred. For example, it is clear that Ю and Ш are dissimilar and that Ш and Щ are similar to each other. Cyrillic characters which are mirror images of their Roman counterparts were rated as dissimilar to the latter.

Table 1
The Four CVC Syllables Shown Concurrently in a Typical Item of each Class

Class			
E	D	S	I
КОМ	ФАТ	ЩЕЩ	ФЁН
ТАМ	ЗОЛ	ЖАЛЬ	НЁС
ТЕМ	ТЮК	ЛАП	ВЬЮН
КЕМ	МЯЧ	ШЕПЬ	ЖУТЬ

The items in one class were composed entirely of graphemes which are common to the Roman and Cyrillic alphabets and correspond to roughly the same phonemes in the two systems. For convenience we refer to these as Class E items, emphasizing the similarity to English graphemes. Since there are only six such graphemes, there was substantial repetition of graphemes within an item. A second class of items included Cyrillic graphemes which are not in the Roman alphabet, but intra-item graphemic similarity was minimized. No grapheme appeared twice within an item, and no item contained two similar graphemes. We shall call these the Class D (distinctive) items. For a third item class the restriction on intra-item similarity was removed and there was a high degree of graphemic overlap within an item. These constituted the Class S (similar) items. The fourth and final class of Russian CVC'S included graphemes which appear in both Cyrillic and Roman, but represent different phonemes in the two systems. Of necessity, this class also included graphemes from Classes E, D and S. No attempt was made to maintain as much intra-item similarity as in Class S. These items will be denoted as Class I (interference). Each syllable consisted of three or four graphemes, depending on whether the "soft sign" was present. In addition, a fifth class of items was constructed using only CVC English words.

An item consisted of a correct alternative and three incorrect alternatives. For control purposes, (a) the incorrect alternatives accompanying a particular correct word varied from one list to another. (The lists are described in the next paragraph.) (b) The incorrect alternatives were words which were correct elsewhere in the list. Selection of incorrect alternatives was random with the restriction that they appear equally often.

Nine lists were prepared, differing only in which wrong alternatives were offered with a particular correct CVC. Since there were four classes of Cyrillic and one class of English items with twelve items per class, a list contained 60

items. Three of the nine lists were presented per session. Sessions were held once daily for fifteen days, so each list was used $\frac{3 \times 15}{9} = 5$ times.

Within a list, the items of Cyrillic Classes E, D, S and I were presented in random order. Of the 12 English items, six preceded and six followed the Cyrillic items. The English items were also presented in random order, and all randomizations differed from list to list. Depending on the subject's latencies, a session lasted from one-half hour to one hour.

For visual display the four members of an item were typed, one above another, in capital letters on the blank side of a standard white IBM card. The syllables were randomly ordered in the column. For auditory presentation of the Russian and of the English items, the correct alternatives were recorded on professional quality tape by a native Russian speaker and by a native American-English speaker, respectively.

Apparatus. A multi-unit automated system performed the stimulus-display and data-recording functions. The video unit consisted of a small modified television circuit, including a camera and screen. The camera served to compose and transmit the image of the typed card to the screen. The audio unit involved a console adapted to operate in synchrony with a standard card punch. A Beckman timer measured choice latencies. The response panel had four vertically aligned buttons, corresponding to the spatial arrangement on the screen of the four syllables. All equipment except the television screen and the response panel were housed in a control room adjacent to the subject's room. The rooms were treated to provide a high degree of sound insulation.

Procedure. The following instructions were read:

This is an experiment in learning to associate sounds and symbols. You will hear a word, and four words will appear on the TV screen, one above another. Press the button which corresponds to the word you heard. For example, if you heard the third word on the screen, press the third button on the panel in front of you. Next the button corresponding to the correct answer will light. After a pause, the procedure will be repeated with a new sound and set of words.

Try to select your answer and press the button as soon as possible, but do not go so fast that you make careless errors. If you have any questions, please ask them now. Press a button when you are ready to begin. Do you have any questions? (after the first six English items):

Now the procedure will be the same except that the sounds and symbols will be Russian. Decide which is the correct printed form of the sound you heard, and press the corresponding button. Do not be distressed because the symbols appear strange at first, but try to learn what sounds they represent. Again, answer as rapidly as you can without making careless errors. Later you will have more English words. If you have any questions, please ask them now. Press a button when you are ready to begin.

Each trial began with the playback of a single taped syllable. Then the four words of that item flashed on the screen, remaining until 1.50 sec. after the subject responded. Immediately after the subject pressed a button, the button corresponding to the correct choice was illuminated briefly so that the subject was informed of the correct answer. Following an interval of .77 sec., the next auditory presentation began.

The mechanics of the latency measurement required that two corrective factors be subtracted from the recorded latency in order to determine the true latency. The first factor subtracted was the interval from the onset of the latency timer to the offset of the taped stimulus word. Spectrographic measurement indicated that this interval averaged .57 sec. The second factor subtracted was the interval from the offset of the taped stimulus word to the onset of the alternatives on the TV screen. This interval was .50 sec.

RESULTS

To determine whether latencies had reached asymptote, the mean latency for a given subject and item class on Days 12 and 13 was compared with that quantity for Days 14 and 15. The Wilcoxon signed ranks test for matched pairs (Siegel, 1956) was performed. The calculated value of this statistic ranged from 6 to 14, depending on the class, and was evaluated with 7 degrees of freedom. None of the differences were significant, indicating that the latencies had stabilized.

Figure 1 shows the group proportion correct for each item class on each day. This proportion remained above .90 after Day 3, and all asymptotic proportions correct exceeded .95. On Days 1–4, the rank order from least to most difficult was: Class, E, D, I, and S. Sign tests among pairs of classes were run by pairing the frequencies correct in the two item classes on Days 1–4 by the same subject. In each case the difference between adjacently ranked classes was significant at the .05 level (two-tailed test) since the direction of the difference was the same for all seven subjects.

The corresponding mean latencies are indicated by Fig. 2. The latencies, like the choices, sharply distinguish one item class from another. In fact, the separation of the latency curves persisted throughout the course of the experiment, although the interclass differences in proportion correct were negligible after Day 4. Class S yielded the lowest proportions correct and the longest mean latencies. For items of this class, the proportions correct and mean latencies, respectively, were .60 and 5.70 sec. on Day 1, as against .97 and 2.10 sec. on Day 15. Two tests of the significance of the interclass latency differences were computed, one on the pooled data of Days 1–8 and a second on the Days 9–15 data. First the median latency over the indicated days was found for each subject and class. Then the medians were paired and tested in the same manner as were the frequencies of correct response. For both Days 1–8 and Days 9–15, the differences between pairs of item classes were in the same direction for all subjects, with the excep-

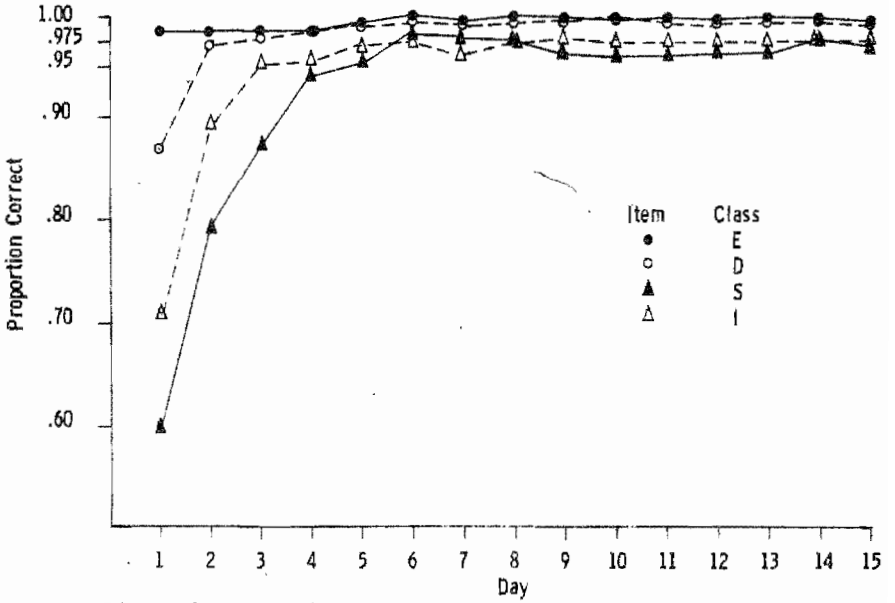

Figure 1. Observed proportion correct for items of each class. Each point is based on an N of 252 (7 subjects times 36 observations each). The mean latency over the responses represented by a particular data point is indicated by the corresponding point in Figure 2.

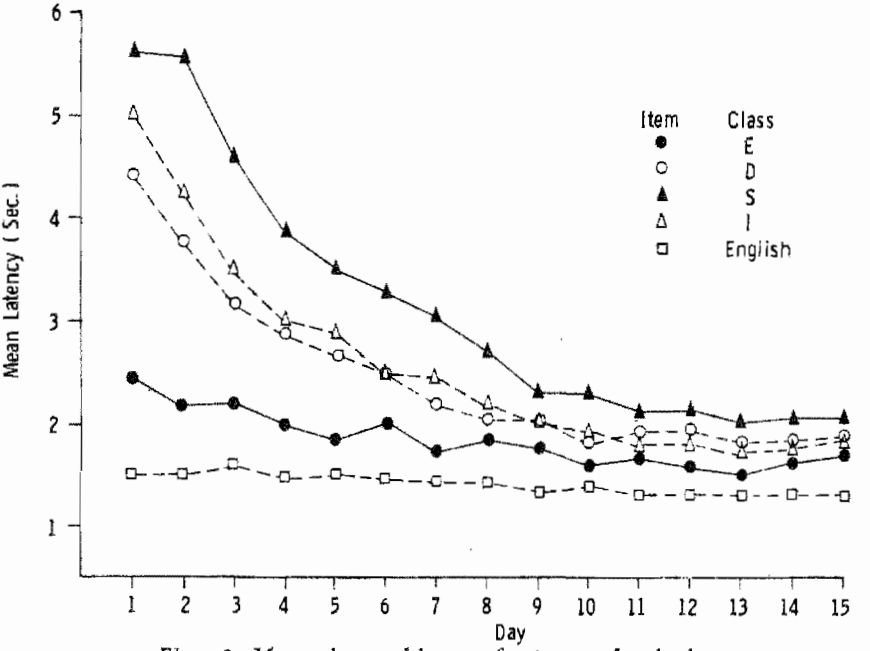

Figure 2. Mean observed latency for items of each class.

tion of the differences between Classes D and I. Hence, all differences are significant at the .05 level (two-tailed test), except for the Class D vs. Class I figure. The latter is not significant on either Days 1–8 or Days 9–15.

For the most part, the interclass differences in mean latency are not attributable to differences in proportion correct. Table 2 shows that even among correct responses the classes differed consistently in latency. Almost without exception in the individual subject data, the rank order of classes from short to long latencies was E, D, I, S. This agrees with the above mentioned ranking (cf. Fig. 1) which was based on proportion correct. In passing, we note that Table 2 reveals appreciable inter-subject differences in latency of correct response, even for the English items.

Table 2
Mean Latency (Sec.) of Correct Choices

	Russian Item Class				English Items
Subject	E	D	S	I	
1	1.75	2.20	2.94	2.49	1.27
2	1.72	2.32	2.92	2.27	1.30
3	1.26	1.49	1.93	1.58	0.95
4	1.90	2.47	2.74	2.33	1.52
5	2.10	2.72	3.25	2.80	1.49
6	2.00	2.90	3.84	3.05	1.53
7	2.07	2.79	3.38	2.80	1.68

The other major findings involve latency and choice trends independent of the item class. Inspection of Figs. 1 and 2 reveals that the decline in latencies continued for many trials after the proportion correct has stabilized above .95. The most conspicuous example of this tendency is that over Days 5–15 the Class S proportion correct increased about .015, while the latency dropped 43% from 3.50 to 2.00 sec. Secondly, it is worth noting that for all classes of Russian items the latencies were approaching asymptotes above the corresponding asymptotic English latencies.

The next analysis showed that for each subject the latencies were longer on incorrect than on correct choices. For the correct and incorrect choices, respectively, the Day 1 latency averaged over subjects was 3.44 and 5.74 sec. This difference in magnitude is a considerable fraction of the latency.

Except for one subject the latencies of correct choices decreased monotonically over Days 1–3. On the other hand, three subjects produced longer mean latencies of incorrect choices on Day 2 than on Day 1. To aid in resolving the problem of whether either latency curve displays an initial rise, a more detailed analysis

was performed. First, the trials by an individual subject on a particular item class were divided into those before and those after reaching a learning criterion. The rather strict criterion of twenty consecutive responses in the subject-item class sequence was adopted. Then the mean latency was found for each Vincent quartile of the trials prior to the error immediately preceding the start of the criterion run. For a discussion of this method of analysis, see Suppes and Ginsberg (1963). The treatment was extended to include trials after criterion in the following way. If there were 4N trials before the designated last error (N per quartile) the first 4N trials after this last error were analyzed in quartiles also.

To obtain sufficient observations, it was necessary to combine data, either from different subjects or from different item classes. The result of aggregating over subjects is shown in Fig. 3 for the item class which had the highest proportion of errors (Class S). The plot of the latencies of incorrect choices rises markedly. There is evidence of a moderate increase in latency of correct choices over

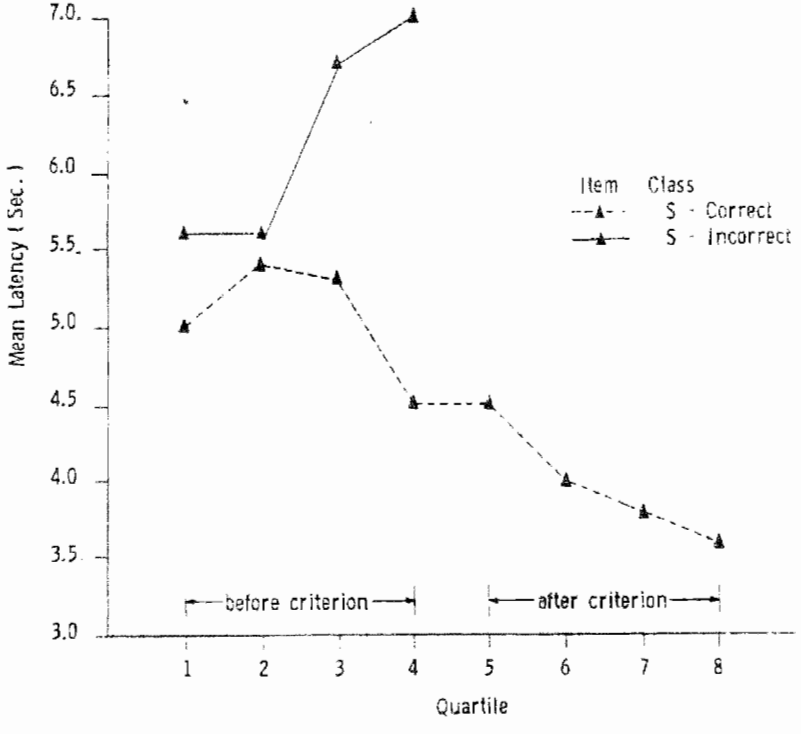

Figure 3. Mean over subjects of the latency per quartile for the item class (Class S) with the highest proportion of errors. There were 145 responses per quartile, of which 76, 52, 28, and 25, were errors in quartiles 1, 2, 3, and 4 respectively, before criterion. There were few errors after reaching criterion.

the first three quartiles. Subsequently, the latency exhibits a steady decline which continues well beyond the trial of last error.

On the other hand, if we combine data over item classes but keep the subjects separate, the initial increase in latency is not statistically significant. This conclusion was obtained by pairing each subject's mean latency on the first half of his precriterion trials with his mean latency on the second half of his precriterion trials.

Two Wilcoxon signed-ranks tests were computed, one using the latency data on correct responses and another using the error latencies. Neither test yielded a significant latency trend over trials. In summary, it appears that latencies did not decline steadily over precriterion trials. Whether the latencies actually increased during the early trials or remained fairly constant depends on how we aggregate data in performing the analysis.

Also, the latency was influenced by the position on the TV screen and response panel of the correct alternative. The mean latencies at the four positions were ranked separately for each subject, permitting a Friedman non-parametric analysis of variance. The χ^2 for summed ranks equalled 102.9, which is significant at the .001 level with 3 degrees of freedom. Inspection of the data indicated that latencies were slightly but consistently lower at the first (top) and third positions than at the alternate keys.

DISCUSSION

The latency measure proved sensitive to differences in item class, stage of learning, and correctness of choice. To probe more deeply the dependency of latency on these variables would require, among other things, an attempt to enumerate the behavioral events intervening between the stimulus and the measured response. The latency would be viewed as the sum of latencies of the subprocesses. We shall limit our remarks to factors which seem to have influenced the mean latencies in this experiment.

First, it should be noted that our experiment involved more stimulus discrimination than is usually required in PAL tasks. Discrimination of visual representations was necessary, along with some concomitant identification of auditory cues. Finally, the subject had to associate some subset of the auditory and visual cues of each item with each other. The terminal act in the stimulus-response chain was a key press rather than a verbal response. One would expect the reaction time component of the latency to be lower when the response is a simple motor act. Our experiment also differed from typical PAL experiments in that the task involved multiple-choice rather than recall. As a rough basis for comparison of our asymptotic latencies with those obtained in PAL recall, we cite a study using high-frequency word pairs and vocal responding (Williams, 1962). After nine correct responses, the latency had declined to about 1.25 sec., which approximates the 1.30 sec. asymptote we obtained with English items.

In our experiment the subject heard the syllable approximately .50 sec. before the syllables appeared on the screen. Hence, the visual scanning time, in addition to the auditory input, decision, and reaction times, contributed to the recorded latency. Presumably, the existence of a preferred scanning order was partly responsible for the significant relationship between the latency and the position of the correct alternative. An experiment is in progress in which the video onset precedes the auditory stimulus, thereby permitting a latency measurement with most of the scanning process discounted.

A major result was that latencies remained about .40 sec. higher on Russian than on English items. A slight practice effect is evident, since latencies decreased somewhat on English items. Nevertheless, latencies on Russian items never reached the Day 1 latency on English items. Whether or not this result can be attributed to knowledge of the English meanings remains to be determined.

Our finding of shorter latencies for correct than for incorrect choices is consistent with a recent result obtained in PAL (Eimas and Zeaman, 1963). Their design was a training trial followed by three test trials, and latencies for saying a number response to a consonant trigram were recorded. At present, it is not clear what conditions lead to an increasing curve on correct or on incorrect choices prior to last error. The observation that latencies decline appreciably in post-criterion responding extends the findings previously cited for PAL and demonstrates that the decline can continue for a considerable number of trials.

Although it is no small undertaking to discover the sources of variance in latency in these second language acquisition tasks, one should remember that other parameters of the latency distribution can reflect experimental manipulations. With rats in a 100% reward runway situation, it has been found (Bush and Mosteller, 1955; Weinstock, 1954) that the frequency distribution of running times changes in shape over trial blocks.

Our principal aims were to determine (a) the quantitative characteristics of the relations between latency, stage of learning, and correctness of response, and (b) whether or not after considerable overlearning the subject could respond as rapidly on the Cyrillic items as on the English items. The relevance of our data to these questions has been discussed. Now we shall indicate what inferences can be drawn from the observed dependency of difficulty on item class.

Caution is required in identifying the stimulus variables responsible for the inter-class differences. Since we worked with a natural language, individual stimulus dimensions were only partially controlled. For example, similarity of Russian phonemes to English phonemes, similarity of phoneme-grapheme correspondences, and intra-item similarity could all vary from one item class to another. Other examples can be readily cited.

One approach would entail constructing highly selected items and an impeccable experimental design for isolating and evaluating the individual stimulus factors and their interactions. However, in view of the magnitude of such a program and its peripheral connection with our present investigation, we followed a

more global approach. The method was to establish markedly different item classes, thereby allowing the assessment of overall stimulus effects. For example, to a first approximation, Classes E and S differed chiefly in that at the outset the subject was familiar with the Class E graphemes. As a second illustration, the comparison of Classes S and I indicates whether or not intra-item similarity is more important than interference from English in influencing item difficulty.

Now let us return to the data. An item analysis of the results clearly shows a very high percentage correct and the lowest latencies for Class E items. Moreover, except for the first day, the percentages correct for Class E and Class D items are about the same. That is to say, the subjects learned the new Cyrillic letter shapes which were quite dissimilar from Roman ones and from each other, during the first session. The latency, however, is significantly shorter for Class E items, yet is consistently higher than that of English items. This may be partly due to the difference in matching a familiar sound sequence with a familiar sequence of letters, as opposed to matching an unfamiliar sound sequence with familiar letters.

It is interesting to note that the same graphemic shapes representing different sounds in Roman and Cyrillic (Class I) proved less difficult than Cyrillic shapes bearing great resemblance to each other (Class S). The subjects apparently could free themselves more readily from a highly over-learned sound-symbol relationship than they could learn the distinctive graphemic features of a new system. As an additional possibility, we would also suggest that Russian phonemes either unfamiliar to American students (e. g., Ц,) or with unfamiliar distribution (e. g., initial Ж) may have contributed to the greater difficulty of Class S items. The long latencies of Class S items cannot be attributed solely to their high degree of intra-item graphemic similarity. This type of similarity was also present in the short latency Class E items. It appears that the combination of unfamiliar stimuli and intra-item similarity is responsible for the long latencies among Class S items.

Having determined that the Class S items yielded the longest latencies, several elementary tabulations were performed in an effort to localize the sources of difficulty within this class. The most striking finding was that the six items with longest mean latencies contained many more instances of the characters Ш, Щ, Ж, Ц, and Ц, than did the remaining six items. All of these graphemes, but not all phonemes, are unfamiliar to American subjects. Hence grapheme similarity seems to predominate over phoneme similarity in producing difficulty.

SUMMARY

An important aspect of second language acquisition is the ability to recognize rapidly the visual representations of aurally presented words in the new language. This problem was investigated by using a paired-associate anticipation learning design. For each item, the stimuli were Russian words recorded by a native speaker. The Cyrillic response alternatives were shown in multiple-choice fashion. When the subject heard a word, he pressed the key corresponding to one of the

alternatives, and was immediately informed of the correct answer. As an independent variable, the Russian items were classified according to the similarity of their constituent printed characters to one another and to English (Roman) symbols.

A major aim of the study was to determine the curve of latencies plotted against trials, as a function of whether the response was correct and of the item class. The results of fifteen learning sessions with college students as subjects were that:

(1) Latencies were asymptotically greater on Russian items than on English items, even though virtually no errors occurred after the fifth session.

(2) Latencies depended on the response and on the degree of training. Latencies were longer on errors than on correct choices and continued to decline after near-perfect learning had been attained.

(3) Latencies depended on the item class. For example, longer latencies were associated with items having Cyrillic characters similar to each other and distinct from English symbols than with items containing a single symbol which represented entirely different sounds in Russian and in English.

REFERENCES

Brown, J., and Huda, M. Response latencies produced by massed and spaced learning of a paired-associate list. *J. exp. Psychol.*, 1961, *61*, 360–364.

Bush, R., and Mosteller, F. *Stochastic models for learning.* New York: John Wiley and Sons, 1955, 321–327.

Eimas, P., and Zeaman, D. Response speed changes in an Estes paired-associate "miniature" experiment. *J. verb. Learn. and verb. Behav.*, 1963, *1*, 384–388.

Lambert, W. E. Measurement of the linguistic dominance of bilinguals. *J. abnor. soc. Psychol.*, 1955, *50*, 197–200.

Luce, R. Response latencies and probabilities. In Arrow, K., Karlin, S., and Suppes, P. (Eds.) *Mathematical methods in the social sciences.* Stanford: Stanford Univer. Press, 1959.

Millward, R. All-or-none latency results in a modified paired-associate learning experiment. Paper read at American Psychol. Assn., Philadelphia, 1963.

Musgrave, B. The effect of nonsense-syllable compound stimuli on latency in a verbal paired-associate task. *J. exp. Psychol.*, 1962, *63*, 499–505.

Siegel, S. *Nonparametric statistics.* New York: McGraw-Hill, 1956, 75–83.

Suppes, P., and Ginsberg, R. A fundamental property of all-or-none models. *Psychol. Rev.*, 1963, *70*, 139–161.

Tolman, E. The determiners of behavior at a choice point. *Psychol. Rev.*, 1938, *45*, 1–41.

Weinstock, S. Resistance to extinction of a running response following partial reinforcement under widely spaced trials. *J. comp. physiol. Psychol.*, 1954, *47*, 318–322.

Williams, J. P. A test of the all-or-none hypothesis for verbal learning. *J. exp. Psychol.*, 1962, *64*, 158–166.

Woodworth, R., and Schlosberg, H. *Experimental psychology.* New York: Holt, 1956, p. 41.