Reading Rates and the Information Rate of a Human Channel

By J. R. PIERCE and J. E. KARLIN

(Manuscript received August 31, 1956)

The limitation on the rate at which information can be transmitted over an ordinary telephone channel is a human one. In this study people read words as fast as they were able to; from these results some deductions are made about the capacity of a human being as an information channel. The discrepancy between human channel capacity measured thus (40–50 bits/sec) and telephone and television channel capacity (about 50,000 bits/sec and 50,000,000 bits/sec respectively) is provocative.

INTRODUCTION

In communication over an ordinary telephone channel, the limitation on the rate at which information can be transmitted appears to be a human one. For instance, by use of a vocoder, the required channel capacity can be reduced greatly with only a moderate reduction in the quality of the reproduced speech.[1]

It would be of great interest to measure the information rate necessary to provide a satisfactory sensory input to a human being. It is not clear how this could be done. Something which may be related and for which a lower bound can be measured is the capacity of a human being as an information channel.

An evaluation of and understanding of the limitations on the information rate of the human channel might ultimately be of practical importance for two reasons. First, it might help to tell us what sort of task to set a human being when he is necessarily a part of a system involving information transmission. Thus, a man can transmit information faster by reading than by tracking. Secondly, the understanding might somewhat illuminate the problem of the channel capacity necessary to provide a satisfactory sensory input, and so might help to reduce the channel capacity required in electrical communication between human beings.

Previous investigations indicate[2, 3] that reading aloud attains the fastest rate at which a human being can be demonstrated to transmit information, as contrasted with, for example, typing, playing the piano, or tracking.*

The work presented here, while undertaken independently, is in general similar to and in agreement with that reported for reading rate experiments by Licklider, Stevens and Hayes,[2] and by Quastler and Wulff.[3] However, we have considered some factors in more detail than these workers, and also, contrary to the former group, we find that, under optimal conditions, reading with tracking has a lower information rate than reading alone.

The chief problem investigated was:

(1) Taking people as they are, with no additional training, how fast, in bits per second, can they transmit information by reading?

(2) What principal factors control this limiting rate?

The experimental procedure consisted simply of people reading aloud as rapidly as they could typed lists of words. Each list was composed of a single vertical row of 12 groups of 5 words, giving a total of 60 words per page. In each instance, the words were chosen at random from a given vocabulary of words. If n is the number of words in the vocabulary and if the words are chosen with equal probabilities, and if all words are read correctly,† the amount of information which is conveyed or transmitted through the human being measured in bits is[4]

$$\log_2 n \text{ bits/word}$$

When the vocabulary for a particular experiment has much fewer than 60 words, certain words must necessarily be repeated several times within a list. When the vocabulary is much greater than 60 words, repetitions are necessarily few and differences in reading rate among different vocabularies would be expected only if the vocabularies differed in nature, as in syllable length or familiarity of words.

Unless otherwise specified, each result quoted below is the average reading speed for two lists for each of three readers, chosen as representing fast, medium and slow readers for people with at least a high school education. The results on these three readers are substantially similar to those on ten similar readers used in preliminary experiments. The chief experiments performed, and some interpretations of them, follow under numbered headings. Some supplementary experiments are then described briefly and the over-all results are commented on.

---

* Here tracking means successively pointing to a series of marks.

† In preliminary experiments the reader's voice was recorded, and it was found that errors in reading aloud occur very seldom if ever.

PRINCIPAL EXPERIMENTS

Experiment 1: Effect of Vocabulary Size

The larger the vocabulary size the higher the information rate conveyed by a given word reading rate. However, one might think that it would be possible to read randomized lists of, say, 4 words substantially faster than lists of 8, or 16 or more words.* How is the word rate affected as the vocabulary size increases?

To investigate vocabulary size as such, it is necessary as far as possible to avoid the influence of differences in word length or familiarity. To this end, words in each vocabulary were chosen at random from the 500 most common words in the language;[5] a few words were then changed so as to keep an average of 1.5 syllables/word for each vocabulary. Figs. 1(a) and 1(b) show parts of typical lists for vocabulary sizes of 2 and 256 words respectively. The order of reading the different size vocabularies was randomized.

Fig. 2 shows that *reading rate is essentially independent of vocabulary sizes from 4 to 256 words when familiarity and word length are kept fairly constant.* The reading rates for the three readers for the 256-word vocabulary are 3.8, 3.7 and 3.0 words/sec, giving information rates of 30, 30 and 24 bits/sec respectively.

The word rate for a 2-word vocabulary is systematically a little greater than for larger vocabularies. This effect, which is statistically significant, is best seen in Fig. 2 in the average curve (dashed). The writers feel on the basis of subjective impressions that this may result from a tendency to group words in pairs in recognizing and speaking them. Among 2 words there are only 4 ordered pairs. It is apparent from the data that no such effect is noted among the 16 ordered pairs occurring with the 4-word vocabulary.

The last point on the curves in Fig. 2 illustrates the importance of familiarity and word length. When words are taken at random from a 5,000-word dictionary (12.3 bits/word), the reading rates drop to 2.8, 2.7 and 2.1 words/sec, yielding information rates of 34, 33 and 26 bits/sec respectively, which are very close to the rates 30, 30, 24 for the 256-word vocabulary.

However, these dictionary lists involve some unfamiliar words and average 2.2 syllables/word.

---

* When the light is very dim, the reading rate is slowed, and is faster for small vocabularies than for large vocabularies. Reading tests were done at normal light levels, which are very much brighter than those at which a slowing due to inadequate illumination is observed.

Experiment 2: Effect of Word Length and Familiarity

It was not clear from Experiment 1 how much of the drop in word rate for the dictionary list was affected by decreased familiarity and how much by increased word length.

These two variables were then untangled in a separate experiment. Word lists were prepared which kept both length and familiarity relatively constant for a given list. The words were chosen from a list of the 20,000 most frequently encountered words in the language.[6] Reading rates were measured for the thousand most familiar words, for the ninth to tenth thousand most familiar, and for the nineteenth to twentieth thousand most familiar words.

The results are shown in Figs. 3(a), (b), and (c). There is considerable consistency among readers as to the relative effect of length and familiarity. The most familiar trisyllable words, for example, are read about as rapidly as the least familiar monosyllables.

A confirmatory demonstration of the effect of familiarity upon reading rate is shown in Fig. 4. This shows reading rates for randomized lists of eight nonsense words averaging 1.5 syllables/word (e.g., jevhin, tosp) which are necessarily totally unfamiliar when the reader first encounters them. As the reader becomes more familiar with the words on successive readings, his word rate increases until he approaches the rates of familiar words in Fig. 2.

Experiment 3: Preferred Vocabulary for Increasing Transmission Rate

The transmission rate is the product of the reading rate and the logarithm to the base 2 of the vocabulary size. To maximize the rate we

Fig. 1 — Parts of typical lists for vocabulary sizes.

grew
action
grew
grew
action

action
grew
grew
action
action

grew
action
grew
action
grew
(a)

foot
tomorrow
count
issue
rain

month
earth
cook
build
corner

yard
history
forest
pleasant
wrong
(b)

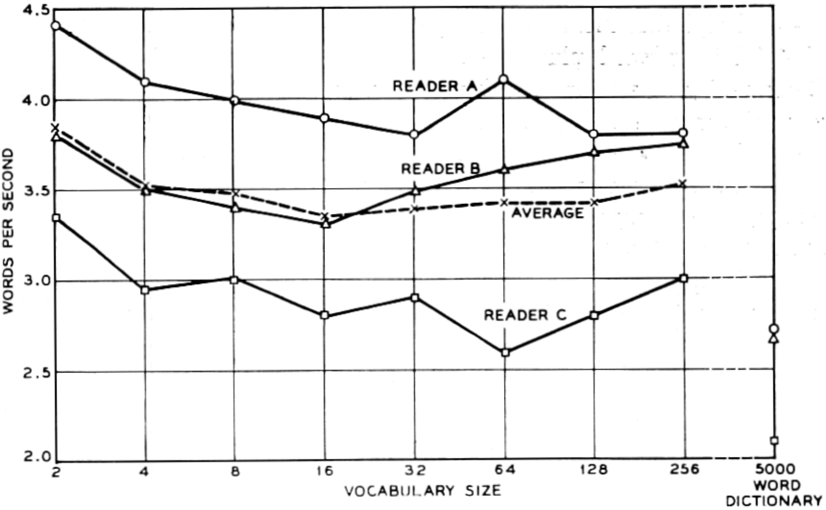

Fig. 2 — Reading rate is essentially independent of vocabulary sizes under certain conditions.

must make each of these factors large. As Fig. 3 indicates, reading speed tends to decrease as vocabulary size increases. From the data in Fig. 3, a rationale (shown in Appendix I) was developed for use in searching for an improved vocabulary which would maximize transmission rate. This indicated that the 2,500 most familiar monosyllables chosen with equal probability should form a very good vocabulary and one which is simple to construct and use. For a 2,500-word list we have 11.3 bits/word.

Reading speeds for such preferred lists were 3.7, 3.4 and 3.0 words/sec, giving information transmission rates of 42, 39 and 34 bits/sec. Some data on the distribution of this rate found among Bell Telephone Laboratories employees is given in Fig. 5.

Experiment 4: Prose and Scrambled Prose

The experiments above were all with discrete words. Reading rates for non-technical prose* are appreciably higher — 4.8, 4.7 and 3.9 words/sec for the three readers. However, such prose has a good deal of redundancy. Shannon[7] arrives at a figure of around 1 bit/letter for a

---

* Extracts were taken from New York Herald Tribune, the novel "East River" by Sholem Asch, "Vermont Tradition" by Dorothy Canfield Fisher and the *Scientific American*. Such material was chosen as being of the same sort of prose as was used by Dewey[5] in his word counts from which Shannon[7] made his estimate of information content of printed English.

27-word alphabet including the space, or 5.5 bits/word for the average of 4.5 letters/word plus one space following a word. Newman and Gerstman[8] give a figure of 2 bits/letter. It is quite uncertain, however, what the true value may be. Table I compares the information rate for the preferred list with that for prose assuming 5 and 10 bits/word.

When words were taken at random from the same prose sources, the reading rates dropped to 3.7, 3.3 and 2.7 words/sec. These rates are about the same as for the preferred list.

The information content of scrambled prose can be estimated much more accurately than that for prose, since the correlations associated

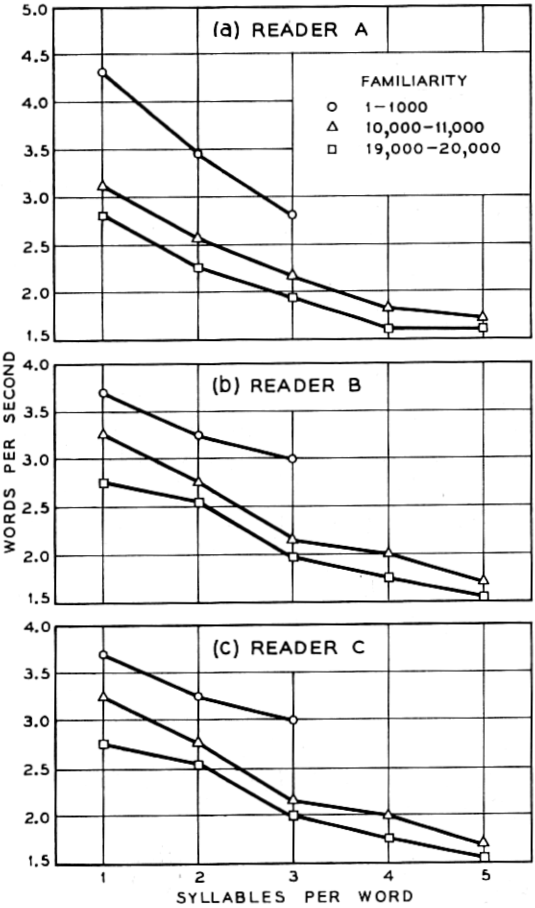

Fig. 3 — Effect of word length and familiarity.

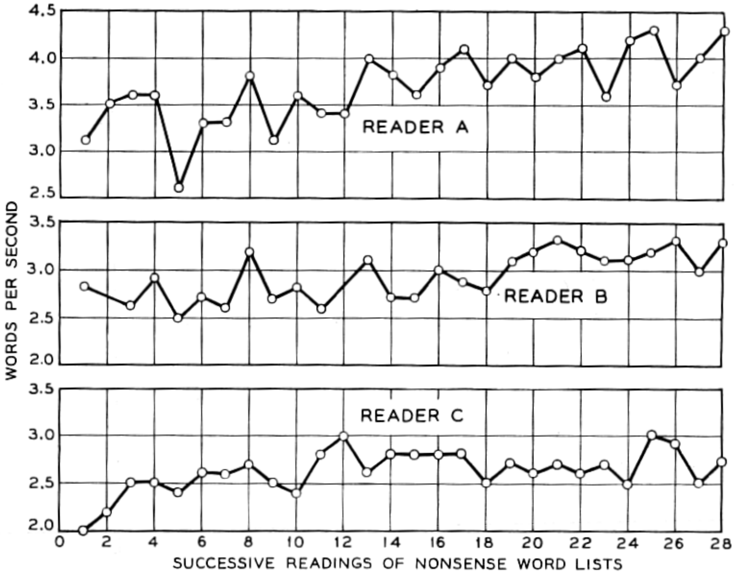

Fig. 4 — Confirmatory demonstration of the effect of familiarity upon reading rate.

with word order have been removed and the bits/word depend only on the frequency of occurrence of words in prose, which is known. Thus, Shannon[7] gives a figure of 11.82 bits/word which applies to scrambled prose, provided the prose has the same word frequencies as that from which the statistics were derived. The information rates for words from a 5,000-word dictionary (Experiment 1) for the preferred lists, and for scrambled prose are given in Table II.

The information rate for scrambled prose is less reliable than the others, because we are not sure that the word frequencies used by Shannon apply to the prose used by us, but we used the type of material cited by the reference he quotes. It is clear that the information rate for scrambled prose is high as compared with most other lists.

Table II shows the gain which may be made by fitting the task to the human being — in this case, by choosing a suitable word list. We may note that the gain appears greater in the case of reader A than in the case of reader B. This need not be experimental error. One would suppose that there are optimal lists for individuals. Indeed, if we compare Figs. 3(a) and 3(b) we see that for reader A the word rate for monosyllables drops by a factor 0.72 in going from the first thousand to the tenth thousand, while for reader B the drop is only a factor 0.88. This

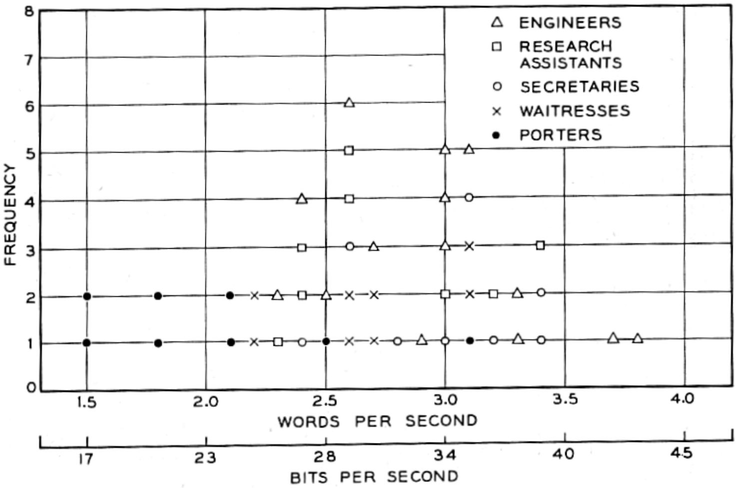

Fig. 5 — Distribution of reading rates for preferred vocabulary.

indicates that the optimal list would be somewhat different for reader A than for reader B. More extensive data would, however, be required to confirm this hypothesis.

Experiments not described here in detail showed that reading rates for digrams (successive pairs of words related as in English text) are intermediate between those for prose and discrete words.

Experiment 5: Effect of Multiple Channels

Licklider[1] has found that when the reader attempts simultaneously to perform a tracking operation while he is reading, his reading rate remains almost unimpaired, and the tracking information is added to that of reading alone. This two-channel transmission gave him his highest rate of transmission. We obtained the reverse finding. Reading the preferred list gave us our highest transmission rate. Simultaneous

Table I

Material	Information Rate (bits/sec)		
	A	B	C
Preferred list	42	39	34
Prose (5 bits/word)	24	23	19
Prose (10 bits/word)	48	47	39

TABLE II

Material	Information Rate (bits/sec)		
	A	B	C
5,000-word dictionary	33	33	26
Preferred list	42	39	34
Scrambled prose	43	39	32

reading and tracking gave a lower total transmission rate. However, Licklider and we agree on the magnitude of this maximum — between 40 and 45 bits/sec for facile test subjects.

Measurements on combined reading and tracking rates were made in Experiment 5 using words from the preferred lists. Whereas Licklider's readers made a dot within a box next to the word read, our readers placed a dot as close as possible to a vertical line next to the word read (e.g. dog |·). The computation of transmission rate is shown in Appendix II. The reading-while-tracking rates were 2.4, 2.0 and 1.4 words/sec. The computed information rates are given in Table III.

It may be seen that the reading rate during tracking dropped so much that the two channels together give a total information rate less than those for reading the preferred list alone. Licklider's reading lists were words chosen randomly from a dictionary and are presumably not chosen optimally for maximum information rate — his information rates for reading alone were 30–35 bits/sec, as compared with the 32–43 bits/sec found here for the scrambled prose and preferred lists. However, if we assume that our reading-while-tracking rate, which is much slower than the reading rate for scrambled prose or for the preferred lists, is limited largely by tracking, we might have obtained a slightly higher information rate in reading-while-tracking by using a larger list of words. This is suggested by the fact that Licklider's and our experiments obtain about the same reading-while-tracking speeds.

TABLE III

	Information Rate (bits/sec)		
	A	B	C
Reading (while tracking)	26.6	22.1	15.4
Tracking (while reading)	10.7	11.0	11.7
Reading and Tracking	37.3	33.1	27.1
(Rates for same word list from Experiment 3 — reading only)	(42)	(39)	(34)

Experiment 6: Effect of Physiological Utterance Limitations

One of our best indications that the maximum reading rate of a subject is determined by mental rather than by physical limitations is that discrete word lists were read no faster silently than aloud. This may appear contrary to very high silent reading rates widely quoted. This can be explained by the fact that in reading much prose we do not and need not recognize every word in order to get the sense. Presumably, if an author made every word say something, his prose could not be read with understanding at such high rates.

We can also show in another way that the mere uttering of the words does not determine the reading speeds observed. A memorized prose phrase ("This is the time for all good men to come to the aid of their country") was repeated several times at rates of 7.5, 9.1 and 8.4 words/sec for the three readers.

Fig. 6 compares word rates for repeating a phrase with the word rates previously discussed. The radically faster rate for repeating a phrase is not the only feature to be observed in this figure; the three readers are not in the same order of speed as is preserved through the reading experiments. This would suggest that it is word recognition rather than speaking speed which accounts for differences among the reading rates of different people.

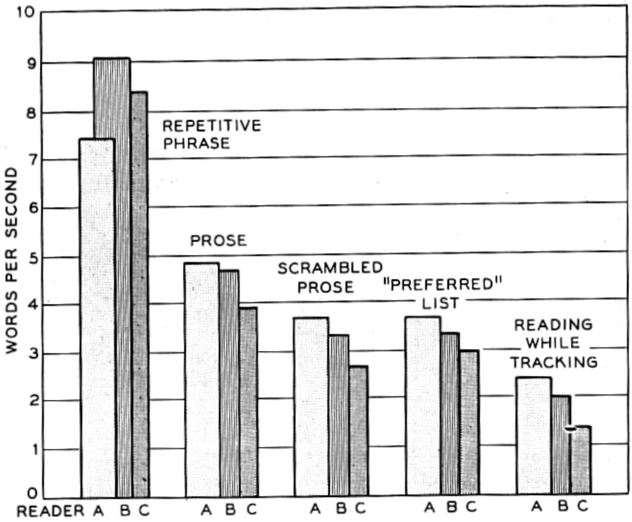

Fig. 6 — Effect of physiological utterance limitations.

DISCUSSION AND SUPPLEMENTARY EXPERIMENTS

Conclusions from Principal Experiments

The conclusions which can be reached with reasonable assurance from these experiments are rather narrow. They might be stated:

1. Information is best transmitted through a human channel by means of well-chosen acts (reading well chosen words in this case) involving many bits per act, that is, much choice per act. Cutting down drastically the bits per act does not substantially increase the speed at which the individual act is accomplished.
2. The lower bound of information transmission through the human channel of rapid readers seems to be about 43 bits/sec. This estimate is a little higher than that found by Licklider,[2] and may be close to a limiting rate.
3. This limiting rate can be achieved by the simple act of reading either randomized lists from suitably selected words or scrambled prose.
4. Both familiarity and length of words are important in determining reading speed. The relative effect of these two variables on reading speed is rather complex.

Beyond these narrow conclusions, there is much understanding yet to be achieved in the general field of the speed of human mental and physical responses and operations. Thus, it seems worth while to mention other experiments which were done in the course of the present investigation and experiments carried out by other workers, and to speculate somewhat concerning the whole of this experimental work.

Multiple Tasks

The reading-while-tracking experiments touch on an important problem. We have all heard of wireless operators who can receive and subsequently type out a message while carrying on a conversation or playing chess. There is nothing in this feat to indicate an information rate greater than that we have found. Actually, the rate of receiving prose by International Morse Code by ear is around 0.58 word/sec;[9] this is slow compared with the rates we have considered.

Our experiments with tracking followed experiments in which words in the lists were randomly printed in red or black, and in which the subject spoke red words in a louder tone of voice than black words, or pressed one key for red words and another for black words. In these cases, the added information, one bit per word, was so small as to make no clearly discernible difference in information rate for the large vocabularies. The speed for reading loud and soft was less than for reading-while-keying. This may imply something about the relative efficiency of human beings performing two tasks by using two sets of muscles as against using one set in two different ways.

It is common experience that we can walk about and carry out other simple tasks while talking or thinking. It is possible though not obvious that some sort of automatic, almost purely reflexive response — as, moving the left hand when the right hand is touched — could with practice be carried out quite independently of a task such as reading. The information rate for such responses would be small, the experimental error would make it difficult to settle the question, and the interpretation of such an experiment would not be entirely clear.

The Patterns Which Govern Reading Time

Early in the experiments the question was raised whether readers may not read letter by letter or syllable by syllable. Several findings bear on this.

Fig. 3 shows clearly that the reading time for a two-syllable word is much less than twice the reading time for a one-syllable word.

One of us knows a negligible amount of German. German syllables are, however, reasonably familiar. It was found that in reading German aloud he had the same reading rate in syllables per second as a man whose native language was German had in words per second. The two readers had substantially the same reading speed in English. Presumably in reading German one man recognized syllables and the other recognized words. This also reinforces the conclusion that reading rate is not limited by the time taken to utter words.

Some experiments were done using lists of common Chinese characters and lists of the corresponding English words. Average word rates over three lists for two readers who could read both languages are given in Table IV. The slightly lower rate for English is plausibly explained by the fact that Chinese was the reader's native language. All words were

TABLE IV

Reader	Words/sec	
	Chinese Words	English Words
E	2.7	2.3
F	3.3	3.2

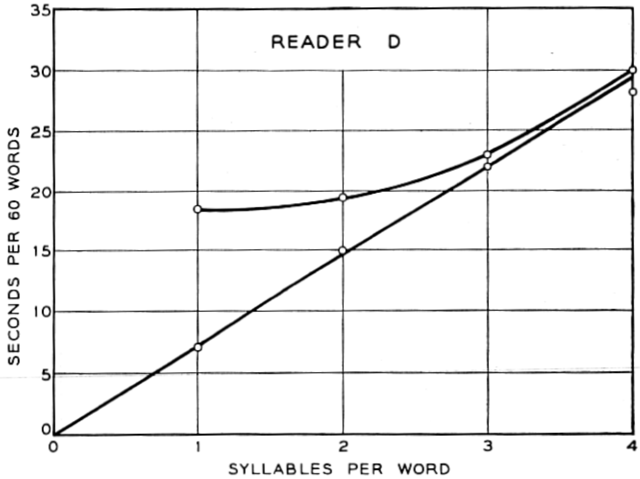

Fig. 7 — Patterns governing reading time.

necessarily monosyllables in Chinese and happened to be monosyllables in English.

In one case a word is made up of a sequence of letters, each standing for a sound, and in the other it is made up of a number of strokes which are meaningless individually, yet in each case a word is taken as a unit or pattern requiring nearly the same time for reading.

We found that the rate for reading arabic numerals is substantially the same as for reading familiar words. Each numeral is an individual pattern to be recognized.

In a first effort to find the effect of syllable length on reading rate, a subject read several lists made up respectively from vocabularies of 16 single-syllable, 16 two-syllable, 16 three-syllable, and 16 four-syllable words. None of the words was very unfamiliar to start with, and all were presumably very familiar after the subject had read several randomized lists composed of the same words.

The outcome of the experiment is shown in Fig. 7. For comparison, the points associated with the lower straight line are *time for 60 words* for repeating, as rapidly as possible, a one-, a two-, a three- and a four-syllable word. The points on the upper curve are reading time for 60 words for the randomized lists of the familiar one-, two-, three- and four-syllable words.

In dealing with such groups of highly and uniformly familiar words, it appears that, roughly, a certain time is required to recognize the word regardless of length, and this time governs the reading rate up to the

TABLE V

Reader	Words/sec		
	Scrambled Prose	Scrambled Paragraph	Prose
A	3.7	4.0	4.8
B	3.3	3.7	4.7
C	2.7	2.9	3.9

point at which the reader is uttering words continuously as fast as he can. This is consistent with a strong subjective feeling that what limits the rate is the difficulty of "recognizing" the word as one looks at it, and that once the word is recognized one can utter it while recognizing the next word.

It would of course be wrong to conclude from this experiment that multisyllable words are in general recognized as quickly as single syllable words, for it would be possible to recognize one among a known group of 16 multisyllable words without looking at the whole word. Indeed, Fig. 3 indicates a substantial difference of reading rate between one- and two-syllable words of like frequency of occurrence. This was not observed in reading the specially familiar lists of one- and two-syllable words.

Why is prose read faster than scrambled prose? It might be that some short phrases are recognized as individual patterns. However, there is another factor at work. A scrambled paragraph of prose is read slower than the same paragraph in its natural word order but faster than scrambled prose from a book or a long stretch of prose, as can be seen from Table V.

It should be noted that reading speed differs for different prose, and that when comparisons among prose, scrambled paragraphs and scrambled prose are made, similar material should be used.

The fact that a scrambled paragraph is read faster than scrambled prose might be explained by saying that we expect, we are more ready to recognize, words which are repetitions of earlier words or words which are closely related in sense to earlier words than we are unrelated words. Thus, the greater reading speed for prose than for scrambled prose seems to be due only in part if at all to the recognition of phrases rather than words as individual patterns.

Rate of Mental Processes

The rate at which information passes through a human channel in reading experiments is indisputable. Quastler[3] has attempted to go beyond this and estimate information processing rates in the brain from the performance of lightning calculators, by dividing the performance of the calculation into a sequence of tasks equivalent to consulting memorized multiplication tables and performing additions. It is hard to interpret such a study clearly, for it is quite possible that there are many sorts of mental acts which take different times to perform, just as multiplication and addition take different times in an electronic computer. A tentative experiment we performed indicated something of the sort.

Randomized lists were made up from vocabularies (a) of names of common animals and vegetables in equal numbers, and (b) of common men's and women's names in equal numbers. In reading these, a subject was asked, not to read the word aloud, but merely to press one key with his right and another key with his left hand; in (a) left-animal, right-vegetable; in (b) left-man, right-woman. The same subject later read the lists aloud. Pressing keys took 40 per cent longer than reading aloud. (The additional time is not related to the keying operation itself; for a 2 word list, for example, keying speed is much faster than reading speed.) Presumably an additional mental operation was involved, but it was not one for which the time was equal to that for reading. This experiment was not pursued further, partly because no clear conclusion could be drawn from it. Had it been pursued and randomized lists of the same words used repeatedly, the rate might have gone up. Conceivably, cow and horse could become for a subject merely different ways of spelling left, and lettuce and carrot variant spellings of right. In this case we would end with a two-word reading experiment.

Reading Rate as a Psychometric Datum

It is interesting to speculate on the possible relationship of reading rate to general intelligence or some other aptitude. Certainly, Fig. 5 indicates some such relationship. We might also ask in connection with Fig. 3, does a rate which falls less rapidly with frequency of occurrence indicate a larger vocabulary, and can we measure vocabulary by reading speed tests? Certainly, measuring the speed of reading aloud is very simple, and such tests might have some psychometric utility.

The Channel Capacity Required for Satisfactory Communication

In conclusion, we cannot help but wonder that the highest information rate noted — 43 bits/second — is so much lower than the channel capacity* of a telephone or a television circuit (around 50 thousand

---

* This is the limiting channel capacity given by (2.1) of Appendix II. The practical rate at which binary digits can be sent over a telephone circuit with simple equipment is less than 1000 bits/second.

bits/second for telephone and 50 million bits/second for TV). This would not be surprising if the limitation we observe had been one of the speeds at which words can be uttered, but it appears rather to be a mental one, one of recognizing what is before the eyes. To the authors, it seems reasonable that this mental limitation may apply to a human being's ability to absorb information, that is, to the information rate needed to present a satisfactory sensory input to a human being. If it does, then why do we need so much channel capacity to convey to him an acceptable sound or picture?

This can be explained in part by the inefficiency of our present communication methods. Despite its present imperfections, the vocoder makes it clear that clearly understandable speech can be transmitted using far less channel capacity than that required in ordinary telephony.[1]

However, it is quite likely that even with the most efficient of encoding means we will have to use far more than 43 bits/second for a picture transmission channel. While only a portion of the image of the transmitted picture falls on the fovea at any instant, we can cast our eyes on any portion of the received picture. If the pick-up camera device and the received picture followed eye movements, a much less detailed picture would serve. Even with our eyes fixed, we can concentrate our attention on a particular part of our field of vision, and this is something that the pick-up camera cannot track. There may be similar effects in our apprehension of sounds.

In the light of present knowledge it is impossible to estimate the minimum channel capacity required to transmit sound and pictures in a satisfactory manner. It will take work far beyond the measurement of reading rates to enable us to make such an estimate.

ACKNOWLEDGMENTS

The writers are indebted to D. L. Letham, who carried out some work preliminary to that reported here, to J. L. Kelly for helpful comments, to D. Slepian for help in putting the appendices in a mathematically more acceptable form, to Miss Renee Hipkins who carried out most of the experiments, and to three indefatigable readers, Mrs. Mary Lutz, S. E. Michaels and A. P. Winnicky.

Appendix I

ON OBTAINING GOOD VOCABULARIES

The experiments in the body of the paper indicate that the choice of a good vocabulary is important in attaining a high information rate in reading lists of words.

In these experiments the randomized lists may be regarded as an information source consisting of a sequence of code elements or symbols (words) in which there is no correlation between successive symbols. If in making up the lists the sth word of the vocabulary is used with a normalized probability p_s, the entropy in H in bits per word, and hence the amount of information per word, is

$$H = -\sum_s p_s \log_2 p_s \text{ bits} \tag{1.1}$$

If all words appear with an equal probability $1/m$ where m is the number of words in the vocabulary, as in the case of experiments 1–3, p_s is $1/m$ for each of the m words and in this special case

$$H = \log_2 m \tag{1.2}$$

In the case of scrambled prose, for instance, the probabilities are different for different words. This will be true also if in making up word lists we choose words randomly from a box containing different numbers of different words.

Let t_s be the time taken to read the sth word of the vocabulary. *Let us assume that t_s is the same for the sth word no matter what context that word appears in in the randomized list.* If this is so, the average reading time per word, t, will be

$$\bar{t} = \sum_s p_s t_s \tag{1.3}$$

the word rate will be $1/\bar{t}$, and the information rate R will be

$$R = \frac{H}{\bar{t}} = \frac{\sum_s p_s \log_2 p_s}{\sum_s p_s t_s} \tag{1.4}$$

Suppose we have available a vocabulary of words and know the reading time t_s for each word. The problem is to choose p_s in terms of t_s as to maximize R. This is easily done; however the result can also be obtained as a special case of the problem treated in Appendix 4 of "The Mathematical Theory of Communication."[4] In Shannon's $\ell_{ij}^{(s)}$, the subscripts i, j refer to passing from state i to state j. In our case there is only one state, and $\ell_{ij}^{(s)}$ should be identified with t_s for all i and j. Similarly, we identify $p_{ij}^{(s)}$ with p_s. C is the maximum rate, so $\log_2 W = C$. Shannon's equation

$$p_{ij}^{(s)} = \frac{B_j}{B_i} W^{-\ell_{ij}^{(s)}}$$

becomes

$$p_s = 2^{-Rt_s} \tag{1.5}$$

since there is only one B.

Shannon's determinantal equation

$$\left| \sum_s W^{-\ell_{ij}^{(s)}} - S_{ij} \right| = 0$$

becomes

$$\sum 2^{-Rt_s} = 1 \tag{1.6}$$

In (1.5) and (1.6) we have a means of evaluating p_s in order to attain the maximum information rate R.

The data we actually have concerning words is that for some class s of words, say, the monosyllables in the 8,000–9,000 words in order of familiarity, the reading time has some value t_s, presumed to be the same for all words in the class, and that there are N_s words in this class. In this case we must assign to each word in the sth class the same probability p_s given by

$$p_s = 2^{-Rt_s} \tag{1.7}$$

and we must have

$$\sum_s N_s p_s = \sum_s N_s 2^{-Rt_s} = 1 \tag{1.8}$$

Using the same amount of data given in Fig. 3, for the 20,000 most common words, but for a different reader, estimates were made of N_s and t_s for all the classes consisting of words of each number of syllables in each range of occurrence of 1,000 words. Then the optimum values of p_s for words in each class and the maximum rate R were computed.

Using (1.4), rates were also computed for choosing words with equal probability from among the first m thousand words and from among the first m thousand monosyllables, as functions of m. These rates had

TABLE VI

Nature	Computed Rate, bits/sec
Maximum Rate	33.5
Maximum for equi-probability monosyllables (from first 8,000 words)	32.4
Maximum for equi-probability among words of all lengths (first 5,000 words)	30.2

maxima for vocabularies of optimal sizes. Table VI compares the various rates computed.

As it is much easier to make up lists from the 2,500 monosyllables among the first 8,000 words with equal probabilities than it is to make up lists from among all words with a different probability for each class, and as the information rates computed were close together, the former alternative was chosen.

The use of scrambled prose provided an easy way to make up good lists.

Appendix II

TRACKING EXPERIMENT

A well-known formula for channel capacity R in bits/sec is[4]

$$R = B \log_2 \left(1 + \frac{P_s}{P_n}\right) \tag{2.1}$$

This gives the limiting rate at which information can be transmitted over a channel with a bandwidth B by a signal of power P_s, in the presence of a gaussian noise of power P_n, with an error rate smaller than any assignable number.

In most cases, the actual rate is much smaller than this limiting rate. In general, the rate is the entropy of the received signal minus the entropy of the noise. In the particular case of a gaussian signal source as well as a gaussian noise, each represented by 2B samples a second, the calculation based on entropies gives exactly (2.1). Let us then apply (2.1) to the tracking experiment.

Suppose that a large number N of samples do have a gaussian distribution of mean square amplitude $\overline{x^2}$. Suppose that we make an error d_n in reproducing the nth sample, that these errors are gaussian, and that the mean square error is $\overline{d^2}$

$$\overline{d^2} = \frac{1}{N} \sum_n \overline{d_n{}^2}$$

We see from (2.1) that ideally we can use these reproduced samples to transmit M bits of information where

$$M = \frac{N}{2} \log_2 \left(1 + \frac{\overline{x^2}}{\overline{d^2}}\right) \tag{2.2}$$

In the reading and tracking experiments, randomized words from the 2,500 commonest monosyllables were arranged with equal vertical spacings but with various horizontal positions. To the right of each word was a short vertical line. The distances x_n of these lines from the vertical centerline of the paper were obtained from a list of random numbers with a gaussian distribution such that for the list $\overline{x^2} = 1$ inch. Of course, $\overline{x^2}$ for each list would depart from this value. As the words were read, the reader used a pencil to make a dot as near as possible to the corresponding vertical line. For each sheet, the departures d_n from the vertical lines in inches were measured and $\overline{d^2}$ was computed. The number of bits M for pointing for that sheet were then taken as

$$M = \frac{N}{2} \log_2 \left(1 + \frac{1}{\overline{d^2}}\right) \tag{2.3}$$

REFERENCES

1. E. E. David, Naturalness and Distortion in Speech Processing Devices, Jour. Acoustical Soc. Am., July, 1956.
2. J. C. R. Licklider, K. N. Stevens, J. R. M. Hayes, Studies in Speech, Hearing and Communication, Technical Report, Acoustics Laboratory, MIT, Sept. 30, 1954.
3. H. Quastler et al, Human Performance in Information Transmission, Report No. R-62, Control Systems Laboratory, University of Illinois, March, 1955.
4. C. E. Shannon and W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, 1949.
5. G. Dewey, Relative Frequency of English Speech Sounds, Harvard University Press, 1923.
6. E. L. Thorndike, A Teacher's Word Book of the Twenty Thousand Words Found Most Frequently, Teachers College, Columbia University, 1932.
7. C. E. Shannon, Prediction and Entropy of Printed English, B.S.T.J., **30,** pp. 50–64, Jan., 1951.
8. E. B. Newman and C. J. Gerstman, A New Method for Analyzing Printed English, J. Exp. Psychol., **44,** pp. 114–125, 1952.
9. Keith Henney, *Radio Engineers Handbook*, 3rd Ed., p. 568, McGraw-Hill, 1941.