DOCUMENT RESUME

ED 067 640 24 CS 000 176

AUTHOR	Gibson, Eleanor J.
TITLE	The Relationship Between Perceptual Development and the Acquisition of Reading Skill. Final Report.
INSTITUTION	Cornell Univ., Ithaca, N.Y. Dept. of Psychology.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-9-0046
PUB DATE	Oct 71
GRANT	OEG-2-9-420446-1071(010)
NOTE	104p.
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	Auditory Perception; *Cognitive Development; *Cognitive Processes; Morphology (Languages); Orthographic Symbols; *Perceptual Development; *Reading Processes; *Reading Research; Semantics; Syntax; Visual Perception

ABSTRACT

The work described in this report is aimed at understanding the role of cognitive development, especially perceptual development, in the reading process and its acquisition. The papers included describe: (1) a theory of perceptual learning, (2) an investigation of the perception of morphological information, (3) the role of categorical semantic information in a visual search task, (4) an investigation of orthographic structure in a visual search task, (5) the role of both syntactic and semantic information in an experiment involving anagram solution, (6) a comparative study of auditory and visual temporal presentations of Morse-code-like patterns, and (7) the design for an investigation of perceptual ordering strategies in relation to categorization in recall. These studies shift the emphasis in reading research from decoding to realizing the syntactic and semantic information; incorporating the rule systems that differentiate reading from the rote learning; and developing economical, adaptive ways of processing the encoded message. (See CS 000 177.) (Author/TO)

FINAL REPORT

Project No. 90046

Grant No. OEG-2-9-420446-1071(010)

THE RELATIONSHIP BETWEEN PERCEPTUAL DEVELOPMENT AND THE ACQUISITION OF READING SKILL

Eleanor J. Gibson

Cornell University

Ithaca, New York

October, 1971

The research reported herein was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

PREFACE AND ACKNOWLEDGMENTS

The theoretical paper and the experiments reported here are largely the outcome of an attempt by the principal investigator to apply a theory of perceptual learning and development to the reading process and its acquisition. The viewpoint is that the reading process in all its complexity--what is learned, what processes are involved in learning (especially what reinforces or selects what is learned), how the good reader skillfully shifts his strategies of information pick-up with the demands of the task--must be understood before we can give advice for instructional methods that is more than piece-meal or opportunistic. A theoretical paper thus leads the papers included in the report, aimed at describing the complexity of the task and the shifting of the reader's attention to different informational features according to their task utility. Several of the following papers seek to provide evidence for this theme in the form of experiments.

It should be noted, however, that a number of people besides the principal investigator were deeply involved in this work, and that they took a highly creative part in it. Some of them were research assistants who received stipends from the grant, but some were students who worked in a very independent way without remuneration from the grant as part of their graduate training. The dissertation at the end is one example. It is included in the form of an abstract, since its length precluded copying it in entirety.

All of us wish to acknowledge the invaluable help of Mrs. Carol Kannus, who managed, despite our clumsy copy, to get everything we asked for--data sheets, letters to parents, manuscripts--in handsome form. We particularly wish to acknowledge our debt to the Ithaca Schools, especially to the administration, faculty, and students of the Northeast School, without whose cooperation and participation most of this work could not have been done.

TABLE OF CONTENTS

LIST OF TABLES

SUMMARY

The work described in this report is aimed at understanding the role of cognitive development, especially perceptual development, in the reading process and its acquisition. The first paper describes briefly the principal investigator's theory of perceptual learning, and goes on to describe the information in the written word that is available for pick-up (graphological, phonological, semantic, and syntactic features). It suggests that these classes of information may be picked up independently, that there is a developmental sequence in learning to do this, and that the reader, as he becomes skilled, learns to assign priorities for pick-up to those features having greatest utility for the task. It emphasizes that reading involves many different tasks and many different processes.

The first experimental paper is an investigation of the perception of morphological information (inflections) in brief visual presentations of words in three age groups (third grade, fifth grade, and adults). Results suggest that this information is a feature of words that is independently and increasingly picked up as reading skill develops. There follows a paper on the role of categorical semantic information in a visual search task. The third paper is an investigation of orthographic structure in a visual search task. It turns out that both semantic and orthographic information have low utility for visual search of word and letter lists and are ignored, with practice, in favor of pick-up of purely graphic information. This is true, at least, for readers who have progressed to fifth grade or beyond.

The role of both syntactic and semantic information is studied in an experiment involving anagram solution by third and fourth grade children. These children make use of syntactic structure, when it is there, only if given a strong hint. Semantic structure (category membership) is picked up, utilized, and sought for in further problems, suggesting that discovery of this economical structure facilitates linguistic problem solving (which many reading tasks can be considered to be) and is reinforcing in the sense that it leads to a strategy of search for further similar structure.

How the ability to perceive and make use of economical structure develops was studied in another experiment with second and third graders. The structure in this case was a redundant spelling pattern that could act as a collative principle in a discrimination task. Only 50% of the children found and used the economical strategy, so this experiment has led to others, currently in progress, to investigate with improved procedures the roles of age, salience of spelling patterns, and the ability to use redundancy.

Other experiments include a comparative study of auditory and visual temporal presentations of Morse-code-like patterns to first-grade children, and a criticism of this paradigm as an appropriate

analogue of decoding in reading. Another is concerned with the development of immediacy of pick-up of word-meaning from the printed word, using a Stroop-like interference task especially designed to be appropriate for first and second grade children. Finally, a thesis designed to investigate perceptual ordering strategies in relation to categorization in recall is presented in abstract form. This study included three age groups (first, third, and fifth grade) and two types of training (active search and verbal tuition), to determine their effects on ability to use order. Pictorial and written material were also compared in the study.

In general, these studies seem to the authors to imply the importance of changing our main emphasis from the study of simply learning to decode in reading to the more complex processes that determine whether a reader is getting the information that he wants from the printed message. Translating graphic symbols to a phonological representation is a minor aspect of reading and not an end in itself. Realizing the syntactic and semantic information; incorporating the rule systems that differentiate reading from rote learning; and developing economical, adaptive ways of processing the encoded message should be targets for research, for it has become obvious that skills of these kinds do not come automatically with the ability to decode.

Perceptual Learning and the Theory of Word Perception*

Eleanor J. Gibson

Abstract

Perceptual learning involves the learning of distinctive features and higher order invariants, learning progressing actively toward the most economical features and structure. Features of words are classified as phonological, graphic, semantic, and syntactic. Features of these classes are processed independently and sequentially. Ordering of priorities changes with development, and shifts strategically with the demands of the task. Evidence is presented for priority differences for each class of feature depending on task differences.

This paper is the outcome of two long-time endeavors of the author--the development of a theory of perceptual learning, and a program of research on reading. The aim is to try to show how the two are related. First, the theory of perceptual learning will be described, as briefly as possible. It attempts to answer three questions: First, what is learned? Second, How? What are the processes involved? Third, what is the motivation and reinforcement for perceptual learning?

What is Learned?

I believe that what is learned in perceptual learning are distinctive features of things, of representations of things, and of symbolic entities like words; also the invariants of events that occur over time; and finally the economical structuring of both. I think the information for learning these is potentially present in stimulation, to be picked up by the observer given the proper conditions for it.

Consider some examples. Sets of distinctive features characterize objects and entities both natural and artifactual--the furnishings of the world, such as people, dwelling places, things to eat; and, particularly relevent for the present topic, symbols written on pieces of paper, like letters and words. The set of letters of our alphabet is characterized by a set of distinctive features, which in different combinations permit a unique characterization of each one. My students and I have spent much time trying to describe the set of distinctive features that are shared by letters of the Roman alphabet. We have had some success, since confusion matrices obtained experimentally reveal, via cluster analysis, some contrasting features that can be diagrammed in a tree structure (Gibson, Schapiro, & Yonas, 1968).

What about invariants of events? These occur over time.

*Invited address for Division 3, American Psychological Association.

The nicest examplex of learning to detect them occur in perceptual development. The constancies can be understood as invariants under transformations that occur in an event like the approaching or receding of an object, or rotation of it. Perceived existence of an object despite temporary occlusion by a screen of some sort is another (the event of "going in front of" of "going behind"). What has this to do with words? Events like appearance, disappearance, and reappearance have meaning, and these very meanings are expressed spontaneously in the child's earliest two-word utterances (Hi Daddy; all-gone ball; more milk; ball again). These utterances appear to be invoked by the event, quite literally mapped to it (Bloom, 1970; Brown, 1970). Detection of invariance, in other words, is prior to symbolic referential meanings and is reflected in them. Semantic features of words generally indicate perceivable features of the world, both events and things. Verbs for instance can be classified as action or state, an important semantic distinction arising directly from differentiating invariants in the environment. Nouns can be classified as count or mass, a distinction that depends on differentiating things that have fixed borders or shapes from those that are fluid, like sand or water.

Finally, what are some structures discovered by perceptual learning? Entities in the world, both natural and artifactual, have structure; that is, relations between features. These relations can be subordinate and superordinate. Both superordinate and subordinate structure are progressively discovered in development for both objects and events--the structure of events was referred to by Brunswik as the "causal texture of the environment." Examples of the subordinate and superordinate structure of words are obvious. Words come in sentences; sentences are parts of paragraphs; phrases are parts of sentences.

Processes of Perceptual Learning

How are these things learned? Not, I believe, by associating a response of any sort, or an image or a word, to a "stimulus." Distinctive features and invariants must be discovered--extracted from the multiplicity of information in the flowing array of stimulation. We have always accepted the notion of abstraction to explain the genesis of concepts. I think a process akin to abstraction--the dissociation of an invariant from transforming or variable context--happens also at a perceptual level. The phoneme is segregated from the flow of speech heard by the pre-verbal infant and its invariant distinctive features are abstracted from many varying samples and over many transformations produced by different speakers. The process has to be one of abstraction--there is no response to be associated, nor is there an identically repeated unvarying stimulus.

What happens to the variable, irrelevant components of stimulation? When something invariant is abstracted and selected out for attention, what happens to the rest? I think that the process of abstraction is accompanied by a filtering process that attenuates and suppresses the irrelevant--this is what happens to the words that <u>aren't</u> heard in the dichotic listening experiments, for instance.

Finally, a very important process in perceptual learning is the operation of mechanisms of external attention. The sensory systems are all active and exploratory. "Looking," "listening," and "feeling" are terms that describe the search for information in stimulation. They also underline the fact that perception and perceptual learning are <u>active</u> processes. There is improvement and flexibility in attentive strategies depending on age and on the business in hand.

<u>Motivation and Reinforcement</u>

What selects the good strategy, the economical feature, the structure that most effectively orders the information? This is the question of motivation and reinforcement in perceptual learning. I do not think the motivation is to be found in drives like tissue needs, nor the reinforcement in reduction of a metabolic drive, nor in cessation of punishment. I think there is a built-in need to get information about one's environment. One could call this, as a number of psychologists do, "intrinsic cognitive motivation." I can call it the "search for invariance." This motivation is always related to the task in hand, for different information is needed for different tasks. But active looking for information about the solid, permanent safe places of the world, and the invariant aspects of events (like the swift approach of an object) is essential for behaviors like locomotion and avoidance of predators. A desire to understand what others are saying seems equally basic for learning to comprehend language.

Reinforcement of perceptual learning (indeed of all cognitive learning) is, I would contend, the reduction of uncertainty. Discovery of structure, or discovery of the economical distinctive feature, or of the rule describing an invariant reduces the "information load," and leads to permanent perceptual reorganization for the viewer of the world so viewed. This kind of discovery also leads to repetition of the successful strategy when a similar occasion occurs. The evidence for this is slight at present, but I have been conducting experiments to see whether discovery of structure is indeed reinforcing. I think the experiments I am going to cite as illustrations of changing strategies in selection of word features will show the self-regulating, adaptive character of the process.

How are Words Perceived?

Now we are ready for the question of how words are perceived. I shall concentrate on how they are perceived in reading, since that is where my research has been centered. The answer sounds simple, and is not new. <u>Words, like other entities in our</u>

environment, are perceived by detecting their distinctive features.* But what are a word's distinctive features?

Features of Words

A word is part of a vast system of information. The way to identify this information is to refer to what is learned when we learn to hear and speak, and to read and write. The information dealt with in hearing-speaking is traditionally divided into three aspects or classes, phonological, semantic, and syntactic information. There are three parallel aspects for reading-writing; graphological, semantic, and syntactic information.

These classes of information tell us what kinds of features a word can have. When written, it can have graphological features of many kinds. For example, it has a characteristic shape and length (referred to as "word shape" by reading teachers). Within the word there are letter shapes, themselves differentiated by distinctive features. And then the word has orthographic structure; letters are combined into words according to a rule system so that given combinations or clusters are permissible only in certain locations and contexts. Q must be followed by U, for instance, Qu can begin a word or a syllable, but it cannot end them (Venezky, 1970).

How do we know whether a potential feature of a word is actually being detected as a feature? A useful method has been to set up tasks that produce some errors and then to study what is confused in the errors. We can infer some of the features of letter shapes that are being noticed when a child confuses E and F but not E and O. We infer that a larger graphic structure is perceived when the child confuses palindromes, like "saw" and "was." He is generalizing something relational. We can also study features by looking at accurate generalizations, like discovering a rule about spelling patterns and transferring it to new cases (Gibson, 1970).

A word has potential phonological features, even when it is perceived by reading instead of being perceived by hearing. When we read poetry, we are keenly aware of acoustic similarity in the rhymes, for instance. Homophones, albeit they are spelled differently, are sometimes confused. The fact that a string of letters presented visually is pronounceable vastly affects its readability. This was demonstrated by Gibson, Pick, Osser, & Hammond (1962). We

*The notion that a word is essentially a complex of features has been suggested and explored previously by Anisfeld & Knapp (1968), Bower (1967), Fillenbaum (1969), Katz & Fodor (1963), and Wickens (1970). The present paper shares this notion with these authors, but differs from a number of them in conceiving the perception of words, like that of things and events, to be a process of selective detection of information rather than one of encoding from bits of sense-data.

made up pseudo-words that could be pronounced, though they were not "real" words. We constructed them so that they began with initial consonant clusters that could not end a word, and were terminated with clusters which could not begin one. Then unpronounceable strings were constructed from the pronounceable ones by exchanging the initial and final consonant clusters. An example is the construction of CKURGL from GLURCK.

When these letter-strings were presented tachistoscopically in a mixed order, subjects invariably read the pronounceable version more accurately than the unpronounceable version. The advantage might have been due to an easier pick-up of acoustic structure; or easier pick-up of articulatory structure. But it could also have been influenced by good orthographic structure, quite apart from phonological features, for we found later that profoundly deaf Ss showed the same relative advantage (Gibson et al. in Levin & Williams, 1971). These are all potential features of a word, but from different feature classes.

A word has many semantic features. It may have markers of various kinds indicating classes or properties, like "animate" nouns, "proper" nouns, "count" nouns, "mass" nouns. It may stand for objects belonging to taxonomic categories like edible things, or for events like looming or disappearance. It has similarity and contrast relations--that is, synonyms and antonyms. We know these features are picked up because one can show experimentally that they are confusible within classes,* and they tend to "cluster" in recall. Semantic features of a word include values, as well. Words can be rated for pleasantness-unpleasantness, or ranked on Osgood's semantic-differential scales.

A word's syntactic features are equally obvious. It is a part of speech, like noun or verb or adjective. It has a role in a sentence, like subject or object. And it may possess a morphological marker, like pluralization or tense. This last feature, morphological information, was investigated in an experiment by Gibson & Guinet (1971). We wondered whether the length of a word correctly perceived tachistoscopically could be increased by adding a well-known inflected ending to a base word, as compared to an uninflected word of equal length. Are the endings themselves a unitary structure?

We added inflected verb endings to three types of base words; real familiar words, pronounceable pseudo-words that were anagrams of them, and unpronounceable pseudo-words, and provided comparisons of uninflected words of different lengths. These words were shown tachistoscopically to subjects from third grade, fifth grade, and the elementary psychology course. The results were not

*See Wickens (1970) for a summary of many relevant experiments.

what we had anticipated, but they were very illuminating. Subjects were not able to perceive a longer word correctly when a base word was expanded by adding an inflected ending. But the endings gave evidence of unitariness, for there were significantly fewer errors in inflected endings than in endings of base words of equivalent length (especially when these were not meaningful or pronounceable). Furthermore, when there were errors in the inflected endings, the errors tended to be substitutions of other inflected endings. There was a confusion within a morphological feature-class. These latter two tendencies increased from third to fifth grade, showing progressive pick-up of syntactic features.

Why wasn't a longer word perceptible when the syntactic marker was added? Because the subject had to process an extra feature. It is features that must be processed, not elements like letters or syllables, and the process takes time. No matter how long the base word, its features must be recognized and so must the morphological tag.

Three Hypotheses

I would like to suggest three hypotheses about how these features are perceived in reading. First, I think the four general classes of features, phonological, graphological, semantic, and syntactic are processed independently and sequentially, in a kind of hierarchy. If presentation time is cut short, a feature low in the hierarchy may be missed. It may not be deleted, but just not fully detected, so a confusion within its feature-class may result (like substituting ing for ed). Proof-reader's errors are a case of not fully processing an orthographic feature; but as a matter of fact, spelling is generally noticed in reading, when the time is not limited. Something looks odd and we go back after a sentence or two and verify the mistake.

A classic piece of evidence for ordered, indepdent processing of word features is so-called "semantic satiation" or "loss of meaning," where presentation time is exaggeratedly prolonged, instead of cut short. If I put a printed word in front of you and tell you to stare at it for five minutes, its "meaning" is said to slip away. What happens is that first the semantic features go, then the phonological features go, then one is left finally with the graphic features only, and even these will eventually fragment. There is a very interesting implication here. Meaning, for an adult reader, is embedded in the word. He doesn't begin by decoding it letter by letter; the concept symbolized by the word "hits him." It is specified for him in stimulus information.

The Stroop test is further evidence of this (Stroop, 1935). In this experiment, two features are put in opposition, meaning and a graphic feature, the color of the ink the word is printed in. The subject is shown an array of words and asked to name the color each word is printed in, going from left to right as in reading--green, red, blue, and so on. But the words themselves are the names of colors. When the name and the color of the ink conflict, the subject is in trouble. The name comes first to mind and his performance is badly slowed up compared to just giving the name of a color patch. There seems to be less interference in this task with very young readers. The meaning isn't yet as firmly embedded in, or specified by, the word on the page.

The second hypothesis is that there is a developmental change with age and schooling in feature analysis and pick-up. At an early age, phonological features of a word seem to have more control, in the sense of yielding greater generalization, than semantic ones. Riess (1946), using a conditioned GSR technique, found greater generalization at eight years of age to homonyms than to synonyms, but by adolescence the situation was reversed and semantic similarity became more effective. Perhaps the younger Ss simply had less knowledge of similarity of meaning. Rice and DiVesta (1965) controlled for this in an experiment using a paired associates method, making sure that the homonyms, synonyms, and antonyms used were recognized as such by the subjects. In the younger children (third and fifth grades) generalization occurred as a result of phonetic but not semantic similarity. But semantic generalization became increasingly apparent in older age groups. Felzen and Anisfeld (1970) confirmed these findings using a continuous recognition method. Children in third and sixth grade listened to a list of words and judged for each word whether it had appeared before on the list. False recognition of phonetically related (rhyming) words was more frequent for third graders than false recognition of semantically related words, but semantic similarity was more effective in producing errors for sixth graders.

We can find further evidence for developmental shifts by examining reading errors at progressive levels of instruction. Errors in oral reading through the first grade were studied by Rose Marie Weber (1970) and Andrew Biemiller (1971). The earliest errors, at the beginning of first grade, can generally be attributed to meaningful context. When the child reaches a word he does not "know," he uses all the semantic information at his disposal (context of the words already decoded, pictures, and so on) and guesses. He produces a word that makes sense, both semantically and syntactically, but bears no resemblance otherwise to the one on the page. A little later the child stops when he reaches an unknown word and simply says nothing. This is a transition to a stage where errors become determined by graphic similarity. The child is engrossed by discriminating letters and by correspondences of letters and sounds. Semantic features of the word are temporarily lowered in priority as the child strives to "break the code." (This is the period when he may be chided by the teacher for "reading without expression," but the stage nevertheless marks progress.)

Semantic features return as the decoding process becomes easier and the orthographic features demand a lesser share of the child's attention. But the orthographic and syntactic rule systems probably do not operate fully as important structural constraints until later. The influence of orthographic structure begins to be quite apparent by third grade. Besides the evidence from tachistoscopic experiments (Gibson, Osser, & Pick, 1963), it has been shown by Rosinski and Wheeler (personal communication) that third graders can judge correctly 90 percent of the time which of two pseudo-words is "more like a word." First graders, however, make this judgement at a chance level.

In the Gibson and Guinet experiment (1971), morphological inflections such as verb endings were found to operate as unitary features of a written word. This function was more apparent in the fifth graders than the third graders, and most apparent in college students.

Finally, syntactic constraints like phrase structure and grammatical conventions--the word's role in the sentence--have been shown in studies with the eye-voice span to operate in reading, and again this rule system increases in usefulness as reading skill progresses, being noticeably more functional in fifth than in third graders. The influence of phrase boundaries, for example, only shows up after grade 2 (Levin & Turner, 1968). The young reader does appear, then, to show a developmental sequence in the pick-up of word features. This progression is no doubt not as fixed as I may have implied and probably begins before very long to be influenced by the reader's task, in accordance with my next hypothesis.

My third hypothesis holds that the order of pick-up of word-features changes with the task. To put it a little differently, priorities of pick-up are geared stratetically to task utility. I repeat now my earlier argument that perception is an active search for information, that the perceptual strategy that develops will be as economical as possible and that ordering of priorities is adaptive and self-regulating. Common sense suggests that this is so--when we are looking for a weather report in the newspaper, we assign a very low priority to graphic features of the words we read. But in addition there exists a large number of experiments which go to prove my point, some of my own and many by others.*

We can influence priorities by instructions, of course. In a tachistoscopic experiment, if a subject is told to guess at words shown him, features like meaning and word frequency are evident. If he is told to report only what is literally seen graphic features are advanced in priority (Haber, 1965). But quite aside from instructions, different tasks seem to have acquired their own priorities in the cognitive economy either in the course of development or by learning

*There is so much evidence for this third hypothesis that I can only give examples of it here, roughly one for each feature class. My apologies to the authors whose results are relevant but not included.

during the task. All a word's features have their importance for one task or another.

Phonological features, we have learned in recent years, have a high priority for short-term memory. If I am trying to hang onto a telephone number and someone speaks to me before I have dialed it, it is lost. Conrad (1964) and others have shown in a number of experiments that acoustic similarity produces confusions in short-term memory even when the material is presented visually (Baddeley, 1968). Graphic and semantic confusions in short term memory, on the other hand, are infrequent (Baddeley, 1966). Auditory presentation has an advantage over visual presentation in short-term memory (Murdock, 1967), even when memory is tested by recognition (Murdock, 1968). But as lists are made longer (e.g., 14 vs. 7 items) acoustic confusability effects disappear entirely (Anderson, 1969); Ss no longer give the acoustic feature priority, but adopt some other strategy.* Articulatory similarity may play a role in short-term memory tasks, as well (Crowder & Morton, 1969). I shall class this with phonological information, but it undoubtedly plays its own role, distinct from acoustic similarity, in some tasks.

A word's pronounceability, I suggested earlier, is one of its phonological features, and it strongly facilitates pick-up with tachistoscopically-presented displays. Does it do so equally for another task, like later recall or recognition? Gibson, Bishop, Schiff, & Smith (1964) tried to separate pronounceability and semantic reference and to compare their effects in these two kinds of tasks. Trigrams were prepared which either rated high in pronounceability (like MIB), or in referential meaning (like the initials IBM), or in neither (like MBI). In one experiment, they were presented tachistoscopically and recognition thresholds were obtained. Pronounceability very effectively facilitated accurate perception of the trigram. Meaning helped little.** In another experiment, the same items were presented to subjects for two seconds each and 24 hours later recall was tested. This time the effect of meaning and pronounceability was reversed. Meaning facilitated recall far more

---

*Strategies which reduce the information will be adopted in short-term memory, when possible. Baddeley (1971) has shown that redundant strings of letters will be organized in groups (e.g., B-E-D as the word bed, rather than as three independent letters). Effects of acoustic confusability then drop out.

**An experiment by Pynte & Noizet (1971) also found a strong effect of pronounceability on tachistoscopic recognition of trigrams while finding meaningful sets of initials also facilitating. But this type of meaning was effective principally when the trigram was unpronounceable and had little effect when it was.

than pronounceability and there was evidence of categorizing, for some subjects made up sets of initials, like FDR, that had not been in the list.

While phonological features of words dominate pick-up in short-term memory, there is considerable evidence to show that they have low priority for long-term memory. Acoustic similarity has little or no interfering effect in a paired-associate retroactive-inhibition paradigm (Bruce & Murdock, 1968; Dale & Baddeley, 1969), whereas semantic similarity does. Wickens, Ory, & Graf (1970) found that acoustic similarity had some negative effects in a transfer paradigm, but the effect was slight compared to semantic similarity in the same paradigm. Sharing taxonomic category membership was a powerful influence on the subject's ability to recall items of a list, for either good or ill depending on task relations.

Concrete words are in general superior to those of abstract meaning in almost any long-term memory task, such as PA learning, free recall, serial recall, or recognition (Tulving & Madigan, 1970, p. 452), but this is not true for tachistoscopic recognition, where pronounceability has such a strong advantage (Paivio & O'Neill, 1970).

The utility (and utilization) of semantic features of a word for later recall has often been demonstrated by evidence of clustering in long-term recall. This was brought out cleverly in an experiment by Hyde & Jenkins (1970). In this experiment, the subjects were sometimes asked to do two tasks at once. They were presented with a list of words for later recall. In two conditions, Ss had to extract some graphic information about a word as it was presented--either estimate its length (number of letters), or detect the presence or absence of the letter E. Another group had to rate the word as it was presented for pleasantness or unpleasantness.

Compared to a control group with no second task, recall was greatly reduced for the first two groups and so was the amount of organization in recall as measured by clustering of words in categories. But the task of rating words as pleasant or unpleasant did not reduce recall nor organization in recall as compared with a control group that had no incidental task. When the subject was performing a second task that gave priority to semantic features of the word, neither recall nor its organization suffered. But when the second task required attention to a word's graphic features, like detecting e's or estimating word-length, the semantic pick-up which is apparently vital to later recall of words was blocked.

The Hyde and Jenkins experiment suggests that features of the same class, like semantic features of all kinds, are picked up together, while different feature-classes are processed sequentially (though probably overlapping one another). The value of the word--pleasant or unpleasant--could apparently be assessed at the same time as pick-up of semantic categories of the kind that operate in clustering. I will consider this again in connection with a different task, visual search.

Over the past five years, I have been particularly interested in visual search tasks, partly because I am interested in how perceptual search develops in children and also because they offer a good opportunity for studying what the subject learns incidentally. A lot of perceptual learning goes on during this task. Visual search also provides a fine opportunity for comparing pick-up of different types of word-features. I have made such comparisons in a number of experiments. The task is similar to one used repeatedly by Ulric Neisser (1963) and involves scanning systematically through a matrix of letters for a target letter or word.

When the subject is asked to search for a target letter embedded in a context of other letters by scanning down a column of letters arranged five or six to a row, he very quickly sets his priorities for graphic features. Even seven-year-old children do this. Gibson & Yonas (1966) compared the effect of high and low graphic similarity of context letters to the target letter. Graphic similarity slows the scanning rate enormously, both for adults and children. Given an opportunity for practice, the subject will learn to scan for a single very economical distinctive features as Yonas (1969) and Schapiro (1970) have shown in appropriate transfer experiments.

What about phonological features of the letters in this task? They seem to be virtually, if not entirely ignored. Gibson & Yonas (1966) tried to produce interference by exposing the subject to a voice pronouncing letters that sounded like the target letter while he scanned the list visually. There was no effect at all on scanning rate, even in children of seven years who might have been expected to subvocalize while reading. Kaplan, Yonas, & Shurcliff (1966) compared the effect of high and low acoustic similarity of context letters to the target. That is, the target was embedded in a context of letters that rhymed with it (B, V, D, and so on) or in a context of letters that sounded unlike it. Acoustic similarity did not slow scanning rate at all, in contrast to a powerful effect of graphic similarity.

An experiment by Krueger (1970) found some effect of acoustic confusability in visual search, but it is possible that acoustic was confounded with visual confusability (for instance, C and G, and M and N have both types of confusability and were used within the same target-list set). In the paper by Kaplan et al. (1966) this factor was controlled. Changing or maintaining target items from trial to trial (Krueger changed them, while Kaplan et al. did not) is also an important task variable, since practice with a target or target set to be discriminated visually from another set (Yonas, 1969) is very effective in reducing search to the most economical visual distinctive feature.

Subjects typically remember context letters in the scanning task very poorly, even when they are tested by recognition with the letters presented to them visually (Schapiro, 1970). I wondered whether introduction of some structure of a higher order in the context would not bring it to the fore perceptually. If subjects can learn to take advantage of the most economical possible single graphic feature that distinguishes the target, will they not also discover and use superordinate orthographic structure if it is present?

Yvette Tenney and I performed an experiment in which we compared scanning rate for a letter target, either embedded in letter-strings which, though not meaningful, were orthographically possible words and were pronounceable; or with the target embedded in strings of the same letters scrambled so as to be unpronounceable. I had thought that orthographic structure might be picked up along with graphic structure, which is so salient in this task. If so, it might facilitate search, because the subject could filter the irrelevant context in larger units--strip it off in bigger chunks, so to speak.

On the other hand, if the subject subvocally articulated the pronounceable context items, scanning rate ought to be slowed down. (I have been trying to find out for a long time whether redundant orthographic structure and pronounceability are necessarily functionally tied--my bête noire, in fact.) We ran 76 children in the fifth grade in this experiment. There was no difference in mean rate of scan between the two conditions. What was happening? It would appear that the children were not articulating the pronounceable items. Like lower-order phonological features, pronounceability of a letter-string in its literal sense of pronunciation seems not to influence the kind of verbal processing that goes on in this task. But what about the orthographic structure as such? Does it go unnoticed or can it not be used without an accompanying act of articulation that would be uneconomical? I think the latter may be the case here. When the children were questioned after the experiment they appeared, when context strings were pronounceable, to have been aware of it. Sometimes they commented on it spontaneously. But still it did not affect the mean scanning rate. The child could of course have processed the whole string as a unit, and then searched it for the target letter as a second step. This would contradict the suggestion I made earlier, that a given class of word features gets processed simultaneously. It is also contradicted by recent experiments of Reicher (1969), Wheeler (1970), and Krueger (1970).

On the other hand, the children may have found the orthographic structure early in the task (since they so often commented on it), but learned to disregard it, perhaps because they could not use it without articulation--a handicap in a scanning task where speed is emphasized. This would be a kind of perceptual learning involving an inhibitory or filtering process, one of the three processes for perceptual selectivity that I hypothesized earlier. It may be that adult readers can use the structure without articulation and would speed up their scanning rate even though the children did not.

What would happen if semantic structure were introduced in the scanning task? Would it be picked up, when presented incidentally? Would it speed the search, and if so would it transfer as a strategy--that is, lead to a search for similar structure in a second task? These questions were investigated in another experiment with Yvette Tenney. We introduced semantic structure by building categorical meaning into the context to be searched through in looking for a target word.

Context words, in one condition, were all the same length and all belonged to the same semantic category--e.g., kinds of fruit. We used categorical material that had previously been shown to cluster well in recall experiments. The target word was the name of an animal, and it varied from trial to trial. The S was told to search for the name of an animal, rather than for a specified word, because we found in preliminary work that S searched for nothing but graphic features if he was given a specific target word. We wanted to facilitate pick-up of the categorical relation among context words, if it could be done. A control group had context words chosen at random as regards meaning, but equated with the other condition for frequency, length, and as many graphic attributes as possible.

The results of this experiment ruled out unequivocally the possibility that categorical meaning plays any role at all in a search task of this sort. Semantic structure of the kind we introduced is evidently not an economical feature for a search task when the words are presented in a list, as unconnected prose. It did not speed up scanning rate and was not used by the Ss. Many Ss did not even notice the categorical relation within the context although they were told to use category membership for locating the target. What they actually did, it turned out, was very economical indeed. They looked, after a few trials, not for a word of any particular meaning, but simply for any combination of letters that had not appeared before.

Subjects do learn in this task. They learn the strategy that is most economical for the task. The conclusion is inescapable that semantic features of words have little utility for a search task of this type and are ignored in favor of graphic features that do. It is not that structure is never utilized; graphic structure in the sense of a redundant graphic feature that helps differentiate the target from the background is picked up and used as Schapiro (1970) has shown. But perceptual strategy in this task sets the graphic features a high priority, and other features--semantic, acoustic, even orthographic, low.

What happens in a search task if words are presented in a passage of connected prose and the S asked to search for semantic, or graphic or phonemic targets? Cohen (1970) tried this with all three types of targets and with combinations of them. The semantic target was a word of a given category (e.g., an animal)--in fact, 10 different words belonging to the category were to be cancelled in a meaningful paragraph. In this situation, search for the semantic feature was faster than search for a graphic feature (a letter), and both were much faster than searching for a phonological feature. The semantic feature had the advantage over the other two of redundancy with syntactic and topical predictability, but for adults it is not surprising if meaning is detected early in reading a meaningful connected passage, especially when S is searching for a member of a concept group and not a specific word. The effect of even small changes in task variables in shifting feature priorities is impressive and is witness to the remarkable adaptiveness of human linguistic processing.

One may ask, finally, if there is any task where syntactic features of words have priority. In laboratory tasks as far as I know, this has not been the case, although they are certainly picked up (Gibson & Guinet, 1971). A feature like part-of-speech (noun vs. verb or adjective) is not very effective for producing clustering in recall (Cofer & Bruce, 1965), generalization in verbal conditioning or interference in short term memory (Wickens, Clark, Hill, & Wittlinger, 1968). That may be, one could object, because syntactical information is generally spread over a string of words and is imperfect in one. But although a word always tells us more in context, it appears that its different features are picked up independently and differentially. I find that there is one real-life task in which I seem, willy-nilly, to give first priority to syntactic information. That is, reading students' papers, or a thesis. A split infinitive, or a singular verb following the word "data" distracts me so that I lose the meaning!

Conclusion

Does perceptual learning occur in word perception? I think it does--both during development and within a task, without instruction or even apparent intention. Words contain many kinds of information, and we learn to perceive them as a complex of features. Words should not be thought of as made up of elements of a given length, or as bits of sense-data to be "coded" into something, but rather as entities possessing information classifiable as phonological, graphic, semantic, and syntactic features. All these kinds of information are in words, I think, in the sense that a word *specifies* its information. The perceiver *does* something, indeed, but he does not invent the information.

Word perception, like other kinds of perception, is active, searching for the relevant information in stimulation. Perceptual learning with words, like other examples of perceptual learning, develops toward the strategy that is most economical. This means that priorities for features shift adaptively, with practice in a task, toward those that have most utility for it.

References

Anderson, N. S. The influence of acoustic similarity on serial recall of letter sequences. Quarterly Journal of Experimental Psychology, 1969, 21, 248-255.

Anisfeld, M., & Knapp, M. Association, synonymity, and directionality in false recognition. Journal of Experimantal Psychology, 1968, 77, 171-179.

Baddeley, A. D. Short term memory for word sequences as a function of acoustic, semantic and formal similarity. Quarterly Journal of Experimental Psychology, 1966, 18, 362-365.

Baddeley, A. D. How does acoustic similarity influence short-term memory? Quarterly Journal of Experimental Psychology, 1968, 20, 249-264.

Baddeley, A. D. Language habits, acoustic confusability, and immediate memory for redundant letter sequences. Psychonomic Science, 1971, 22, 120-121.

Biemiller, A. The development of the use of graphic and contextual information as children learn to read. Reading Research Quarterly, 1970, 6, 75-96.

Bloom, L. Language development: Form and function in emerging grammars. Cambridge: MIT Press, 1970.

Bower, G. H. A multicomponent theory of the memory trace. In K. W. Spence & J. T. Spence (Eds.), The psychology of learning and motivation. Vol. 1. New York: Academic Press, 1967.

Brown, R. The first sentences of child and chimpanzees. In R. Brown, Psycholinguistics. New York: The Free Press, 1970.

Bruce, D., & Murdock, B. B. Acoustic similarity effects on memory for paired associates. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 627-631.

Cofer, C. N., & Bruce, D. R. Form-class as the basis for clustering in the recall of non-associated words. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 386-389.

Cohen, G. Search times for combinations of visual, phonemic, and semantic targets in reading prose. Perception and Psychophysics, 1970, 8, 370-372.

Conrad, R. Acoustic confusions in immediate memory. British Journal of Psychology, 1964, 55, 75-84.

Crowder, R. G., & Morton, J. Precategorical acoustic storage. Perception and Psychophysics, 1969, 5, 365-373.

Dale, H. C. A., & Baddeley, A. D. Acoustic similarity in long-term paired-associate learning. Psychonomic Science, 1969, 16, 209-211.

Felzen, E., & Anisfeld, M. Semantic and phonetic relations in the false recognition of words by third- and sixth-grade children. Developmental Psychology, 1970, 3, 163-168.

Fillenbaum, S. Words as feature complexes: False recognition of antonyms and synonyms. Journal of Experimental Psychology, 1969, 82, 400-402.

Gibson, E. J. Principles of perceptual learning and development. New York: Appleton-Century-Crofts, 1969.

Gibson, E. J. The ontogeny of reading. American Psychologist, 1970, 25, 136-143.

Gibson, E. J., Bishop, C., Schiff, W., & Smith, J. Comparison of meaningfulness and pronunciability as grouping principles in the perception and retention of verbal material. Journal of Experimental Psychology, 1964, 67, 173-182.

Gibson, E. J., & Guinet, L. The perception of inflections in brief visual presentation of words. Journal of Verbal Learning and Verbal Behavior, in press.

Gibson, E. J., Pick, A., Osser, H., & Hammond, M. The role of grapheme-phoneme correspondence in the perception of words. American Journal of Psychology, 1962, 75, 554-570.

Gibson, E. J., Osser, H., & Pick, A. A study in the development of grapheme-phoneme correspondences. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 142-146.

Gibson, E. J., Schapiro, F., & Yonas, A. Confusion matrices for graphic patterns obtained with a latency measures. In The analysis of reading skill: A program of basic and applied research. Final report, Project No. 5-1213, Cornell University and U.S.O.E., 1968. Pp. 76-96.

Gibson, E. J., Shurcliff, A., & Yonas, A. Utilization of spelling patterns by deaf and hearing subjects. In H. Levin & J. P. Williams (Eds.), Basic studies on reading. New York: Basic Books, 1970.

Gibson, E. J., & Yonas, A. A developmental study of the effects of visual and auditory interference on a visual scanning task. Psychonomic Science, 1966, 5, 163-164.

Haber, R. N. Effect of prior knowledge of the stimulus on word-recognition processes. Journal of Experimental Psychology, 1965, 69, 282-286.

Hyde, T. S., & Jenkins, J. J. Differential effects of incidental tasks on the organization of recall of a list of highly associated words. Journal of Experimental Psychology, 1969, 82, 472-481.

Kaplan, G., Yonas, A., & Shurcliff, A. Visual and acoustic confusability in a visual search task. Perception and Psychophysics, 1966, 1, 172-174.

Katz, J. J., & Fodor, J. A. The structure of semantic theory. Language, 1963, 39, 170-210.

Krueger, L. E. Search time in a redundant visual display. Journal of Experimental Psychology, 1970, 83, 391-399.

Krueger, L. E. The effect of acoustic confusability on visual search. American Journal of Psychology, 1970, 83, 389-400.

Levin, H., & Turner, E. A. Sentence structure and the eye voice span. In The analysis of reading skill: A program of basic and applied research. Final report, Project No. 5-1213, Cornell University & U.S.O.E., 1968.

Murdock, B. B., Jr. Auditory and visual stores in short-term memory. Acta Psychologica, 1967, 27, 316-324.

Murdock, B. B., Jr. Modality effects in short-term memory: Storage or retrieval? Journal of Experimental Psychology, 1968, 77, 79-86.

Neisser, U. Decision-time without reaction-time: Experiments in visual scanning. American Journal of Psychology, 1963, 76, 376-385.

Paivio, A., & O'Neill, B. J. Visual recognition thresholds and dimensions of word meaning. Perception and Psychophysics, 1970, 8, 273-275.

Pynte, J., & Noizet, G. Trigrammes, syllables et sigles: Étude comparative de la facilité de leur perception. Mimeo, Lab. de Psych. Expérimentale, Aix-en-Provence, 1971.

Reicher, G. M. Perceptual recognition as a function of meaningfulness of stimulus material. Journal of Experimental Psychology, 1969, 81, 275-280.

Rice, U. M., & DiVesta, F. J. A developmental study of semantic and phonetic generalization in paired-associate learning. Child Development, 1965, 36, 721-730.

Riess, B. F. Genetic changes in semantic conditioning. Journal of Experimental Psychology, 1946, 36, 143-152.

Schapiro, F. Information extraction and filtering during perceptual learning in visual search. Unpublished Ph.D. dissertation, Cornell University, Ithaca, New York, 1970.

Stroop, J. R. Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 1935, 18, 643-661.

Tulving, E., & Madigan, S. A. Memory and verbal learning. Annual Review of Psychology, 1970, 21, 437-484.

Venezky, R. L. The structure of English orthography. The Hague: Mouton, 1970.

Weber, R. M. First graders' use of grammatical context in reading. In H. Levin & J. P. Williams (Eds.), Basic studies on reading. New York: Basic Books, 1970.

Wheeler, D. D. Processes in word recognition. Cognitive Psychology, 1970, 1, 59-85.

Wickens, D. D. Encoding categories of words: An empirical approach to meaning. Psychological Review, 1970, 77, 1-15.

Wickens, D. D., Clark, S. E., Hill, F. A., & Wittlinger, R. P. Grammatical class as an encoding category in short-term memory. Journal of Experimental Psychology, 1968, 78, 599-604.

Wickens, D. D., Ory, N. E., & Graf, S. A. Encoding by taxonomic and acoustic categories in long-term memory. Journal of Experimental Psychology, 1970, 84, 462-469.

Yonas, A. The acquisition of information-processing strategies in a time-dependent task. Unpublished Ph.D. dissertation, Cornell University, Ithaca, New York, 1969.

Perception of Inflections in Brief Visual Presentations of Words

Eleanor J. Gibson and Lynne Guinet

Abstract

An experiment was performed to investigate whether one aspect of syntax, verb inflection, creates units for perception in brief visual presentations of words and pseudowords. Inflections did tend to function as units, but not to increase the length of a word that could be read tachistoscopically. Priority of processing inflected endings occurred when the base was not a real word. It was concluded that words are read as complexes of features, of which syntax is one type, and that these features are processed independently.

Perhaps the most interesting, as well as important, problem for a psychological theory of reading is how "higher order" units are formed. The good reader does not read one letter at a time; he picks up superordinate chunks of larger size, varying with a number of structural attributes of the text he is perusing. These attributes may be of many kinds, such as graphological and orthographic structure, conceptual structure or meaning, syntactic and morphological rule structure; but whatever the attribute, it serves to define units that can be processed economically for the reader's purpose. Knowledge of the structural attributes that actually function for skilled readers would be invaluable.

The question we posed for experiment concerns the influence of morphological inflections in reading. Do the well-known rules for forming plurals, possessives, tenses, and so on operate to form a unit for reading? Berko (1958) found that preschool children had already achieved mastery of these rules, as witnessed by their ability to produce vocally appropriate inflections with unfamiliar pseudowords. Does this ability function for the young reader? When he has learned to decode the simple words in his texts, either letter by letter or as wholes, does he quickly learn to apply a rule for the inflected words? And does the rule produce a larger unit for him in the sense that he can take in more "at a glance"? Gibson, Osser, and Pick (1963) showed that children at the end of first grade were able to read unfamiliar trigrams that conformed to English orthography and were pronounceable, better than unpronounceable anagrams of them (e.g., <u>tup</u> vs. <u>ptu</u>). The trigrams were presented tachistoscopically. It was concluded that these children were already learning to perceive as a unit a string of three letters if they conformed to the phonological and orthographic rule system, even though they did not make a familiar word.

Children at the end of third grade had extended this ability in some cases to longer letter strings (four or five). A string of four letters is about as long as a third grader can perceive correctly with a short exposure even when the string is a word or a pronounceable pseudoword. But suppose the string were extended by a well-known ending applied to the base word? Would not a higher order unit of this sort still be perceptible even when the number of letters presented was increased? This is the question we tested in the following experiment with third- and fifth-grade children. There was reason to suppose (Gibson et al., 1963) that use was being made of orthographic rules to form higher units by third grade children. Could syntactic rules be used at this stage to create units for reading; and if not, would they be used by fifth-grade readers? We included a group of college students as well, to provide a comparison with adults.

Method

Construction of Test Items

A list of items was constructed to include three types of words (real words and pronounceable and unpronounceable pseudowords), four lengths of base words (three, four, five, and six letters), and three verb endings (s, ed, and ing). Twelve base words, three of each length, were used. All were verbs. They were common words that a third-grade child would know. From each of the base words, a pronounceable and unpronounceable pseudoword was constructed; e.g., from the base word put we have: put (real word), tup (pronounceable pseudoword), and ptu (unpronounceable pseudoword). This procedure yielded a list of 36 noninflected items. One base word of each length was assigned one of the verb endings, e.g., for the three-letter words: s was the ending for put, ed for ask, and ing for try. The ending was added to all three forms (real, pronounceable, and unpronounceable) of the base word. In the final list, there were six items from each base word: the real and pseudowords alone and the real and pseudowords with endings. For example, from the base word put were the items: put, tup, ptu, puts, tups, and ptus. The types of words and transformations are illustrated in Table 1. The final list had 72 items.

Table 1
Types of Base Words (Real and Pseudo) and Inflected Endings Presented Tachistoscopically to Third and Fifth Grade Children

	Base word	Inflected word	Base word	Inflected word	Base word	Inflected word
Real word	rain	rains	start	started	listen	listening
Pronounceable pseudoword	nair	nairs	trast	trasted	tensil	tensiling
Unpronounceable pseudoword	nrai	nrais	rtsta	rtstaed	tsleni	tsleniing

For the adult Ss the 72 items were arranged in a list in random order. Half of the Ss saw this order of items; the other Ss saw the same list in reverse order.

For the Grade 3 and 5 Ss, the list was divided into six sets of 12 items. Only one form of each word was used in each set. Three words of each base length, one real, one pronounceable, and one unpronounceable item, were in each set. Six of the 12 items in each set had verb endings. The children of each grade were divided into six groups and each group was shown three sets, or 36 words in all. The order of presentation of sets was rotated over Ss to counterbalance order and practice effects.

Four practice items (see, boat, fite, tpos) preceded the list of experimental items given to any S.

Procedure

The S was seated at a table facing an opaque screen about 1.2 m away. The items were projected on this screen. A Kodak Carousel projector with a tachistoscopic shutter was used. The slides were made by printing the letters, all lower case, on opaque paper. The exposure time was set at .067 sec for the children and .033 sec for the adults. The S initiated the exposure when he was ready by pressing a switch. The responses were written down by S.

When S was seated at the table, the following instructions were read to him:

"When you press the switch, you will see a string of letters on the screen. Try to read them and write them down. Sometimes they will make a real word and sometimes they won't but try to read them anyway. They will be on the screen for a very short time so be sure you are looking when you press the switch. Each time you want the next one to come, just press the switch."

The S was then shown how to work the switch and where to write down the letters. After any questions were answered, S was told to press the switch for the first practice item. The E checked that S saw the letters and wrote them down. After the practice items, S was reminded to write down the letters he saw on the screen. After all the items had been presented, S was interrogated to determine whether he noticed the endings, the different kinds of words, and the fact that some were anagrams of others.

Subjects

The Grade 3 and 5 Ss, 60 from each grade (30 boys and 30 girls), were drawn from classes at the Northeast School, Ithaca. The testing was done in the middle of the school year. The children were taken one at a time to a mobile laboratory stationed at the school. None of the children objected to going to the laboratory and most were enthusiastic about participating in the experiment.

The adult Ss were 20 male students in the introductory psychology course at Cornell. Participation in the experiment was considered part of the work of the course.

Results

Results were assessed by number of errors. An error was always defined as failure to reproduce the entire item correctly, no matter how many letters were deleted or substituted.

The overall effects of length of item, kind of word (real words and pronounceable and unpronounceable pseudowords), and age were similar to those in earlier experiments. For all groups there were more errors with longer words. Real words were easier than pseudowords, and pronounceable pseudowords were easier than unpronounceable ones. Errors decreased with grade. Table 2 summarizes these results in percentages. Both inflected and noninflected items

Table 2
Mean Percentage of Errors for Real and Pseudowords for the Three Age Groups[a]

	Real words	Pronounceable pseudowords	Unpronounceable pseudowords	Total
Gradd 3	41.67	65.00	80.14	62.27
Grade 5	12.36	46.39	64.03	40.93
Adults	2.08	17.29	37.71	19.03

[a]The number of words in each category (including both inflected and noninflected items) was 24. There were 30 observations taken on each item for the children of each grade (making 720 possible errors for each cell) and 20 on adults (480 possible errors for each cell). Percentages, rather than number of errors, are given so that children and adults may be compared.

are included. An analysis of variance was run comparing real, pronounceable, and unpronounceable items for Grade 3 and 5 Ss. Number of errors per item (a total of 30 was possible for each, since there were 60 Ss in each grade, each presented with one-half the items), kind of item (real, pronounceable, and unpronounceable) and grade were factors. Data for inflected and noninflected items were combined within word types. There was a significant effect of grade, $F(1, 138) = 22.88$, $p < .01$, and of kind of item, $F(2, 2) = 36.22$, $p < .02$, but there was no significant interaction.

If the verb inflections are perceptual units in reading, the adding of such an inflection might increase the length of word a child could perceive correctly, or make an inflected word easier than one of equivalent length without a suffix. This was tested by comparing the number of errors on words with inflections with the number of errors on words without inflections but of the same length (e.g., trying, six letters including ing, compared to listen, a six-letter item with no suffix). Data were combined for all three types of words. This comparison is shown in Table 3. The overall difference in errors is negligible for both Grade 3 and 5 Ss. This is

Table 3
Mean Number of Errors on Inflected Items Compared to Noninflected Items of Equivalent Length

Item length	Grade 3: Mean number errors: Inflected items	Grade 3: Mean number errors: Noninflected items	Grade 5: Mean number errors: Inflected items	Grade 5: Mean number errors: Noninflected items
4 letters	10.67	11.67	3.33	5.56
5 letters	20.00	19.33	11.33	12.67
6 letters	22.56	24.33	14.44	17.22
Mean of all[a]	19.72	18.44	11.56	11.81

[a]The number of items of comparable length in the inflected category increases as length increases. Thus, the mean for these items is not identical with the average of the three means given here.

the case, also, when the three types of items are examined separately. Inflected items fare better than noninflected when both are unpronounceable pseudowords, but the difference is not significant, and it is in the other direction for the other two types of item.

The errors in the verb inflections themselves were compared to errors in the comparable last letters of noninflected items of the same length, for example, the number of errors in the ing ending on three-letter base words was compared to the last three letters of six-letter words. The three types of item (real words and pronounceable and unpronounceable nonwords) have been combined in Table 4. If the Ss read the inflections as units, we would expect fewer errors on the verb endings than on the last few letters of noninflected items. In all nine comparisons in Table 4, errors are fewer for inflected endings than for non-inflected endings. Only one of these, however (Grade 3, ing), is significant ($p < .01$) by t test. The largest difference between errors on endings and last letters for all groups was on the comparison of ing to the last three letters of other words. The smallest difference in errors for all Ss was on the comparison of s to the last letter of four-letter items.

There would appear to be evidence here that inflected endings tend to be read as units compared to noninflected word endings, even though extension by inflection does not increase, overall,

Table 4
Mean Number of Errors in Verb Endings Compared to the Last Letters of Noninflected Items of Equal Length

Comparison	Grade 3	Grade 5	Adult
"ing"	7.67	5.67	0
last 3 letters	22.89	16.00	4.44
"ed"	16.17	8.33	.67
last 2 letters	19.22	12.50	2.61
"s"	10.11	4.78	.33
last letter	12.85	7.44	1.04

the length of word that can be read. To understand this, it is necessary to look at the distribution of errors in inflected and noninflected endings in the three types of item. Table 5 presents this distribution in percentages of total errors for inflected and noninflected endings for real words, pronounceable pseudowords, and unpronounceable pseudowords. It seems clear from this distribution that inflected endings gain an advantage, relatively, only when the item is a pseudoword, especially an unpronounceable one.

Table 5
Distribution of Errors in Inflected Endings and Noninflected Endings in Words of Equal Length, According to Item Type

	Grade 3 (%)		Grade 5 (%)		Adults (%)	
	Infl.	Noninfl.	Infl.	Noninfl.	Infl.	Noninfl.
Real word	28	18	10	6	0	1
Pseudoword (P)	36	36	36	36	14	29
Pseudoword (Up.)	36	46	54	58	86	70
Sum of above	100	100	100	100	100	100

It is not feasible to compare "intactness" of the beginning as compared with the end of a word by combining all cases and looking at serial position, since base-words differed in length, counts of an arbitrary number of letters violate morpheme boundaries, and errors frequently occurred in both initial and final portions, thus blurring the comparison. But the relative advantage of an ending to a beginning can be illustrated by looking at percentages of errors in the "base" compared to the ending in inflected words, discarding those cases where there are errors in both. For third- and fifth-grade children, the percentage of errors occurring on only the ending as compared to only the base of some typical inflected words are reported in Table 6.

As the examples illustrate, the distribution of errors located only in the base or the ending generally shifted as the base

Table 6
Percentage of Errors Occurring on Endings Only of Some Inflected Words

	Real	Pronounceable pseudoword	Unpronounceable pseudoword
	RAINS	NAIRS	NRAIS
Grade 3	23	9	7
Grade 5	100	0	10
	TRYING	RYTING	RTYING
Grade 3	33	0	12
Grade 5	100	39	28
	TALKED	KALTED	LKATED
Grade 3	100	30	29
Grade 5	50	29	30

word changed character. When the base word was orthographically and semantically "good" the base and the ending appeared to receive at least equal attention, or the base more. But when the base was a nonword, the inflected ending was more likely to be perceived intact than was the base, or than was an equal number of letters without morphological significance. With a word like listen, for instance, errors in the last three letters increased rather than decreased when the spelling was changed to create a nonword (tensil, tsleni). If just the first three or the last three letters are examined, 75% of the errors occurred on the last three for listen, but 100% for tensil and tsleni in Grade 3. In Grade 5, there were 50% on the last three for listen, 100% for tensil, and 90% for tsleni. When there is no unit of morphemic significance (either meaningful, like the ten in tensil or morphological like an ing ending), the S viewing the tachistoscopic display tends simply to read letter by letter, and the ending suffers relative to the beginning. It would seem as if some preattentive process (Neisser, 1967) signals the S's looking behavior.

In our examination of the errors, we noticed that where the ending was wrong, another ending sometimes had been substituted for it (e.g., if the item ended in ed, the response ended with ing). The verb inflections were also sometimes added to items that did not have inflections (these were also termed "substitute" endings). Table 7 summarizes the correct and substitute endings within erroneous responses for items that were actually inflected and items that were not.

The proportion of correct endings to errors (on inflected items) increased with age (Grade 3, .378; Grade 5, .506; adults, .754). The differences between the groups were significant (Grade 3 to 5, $Z = 4.6$; Grade 5 to adult, $Z = 5.9$). For the Grade 3 and 5

Ss, items ending in ed had the smallest proportion of correct endings and the largest proportion of substitute endings (usually ing). Most interesting (though not statistically significant) is the number of substitute endings on inflected items compared to those on noninflected words: Grade 3--66 on inflected words, 60 on other words; Grade 5--47 on inflected words, 31 on other words.

Table 7
Correct and Substitute Endings on Erroneous Responses for Inflected and Noninflected Items

Subjects	Items	Total errors	Number of correct endings	Number of substitute endings
Grade 3	Inflected	798	302	66
	Noninflected	547	---	60
Grade 5	Inflected	549	278	47
	Noninflected	335	---	31
Adults	Inflected	187	141	3
	Noninflected	87	---	3

If the Ss knew that three anagram-like forms of each word and three verb inflections were used in the list, it might have helped them to perceive the items correctly. The responses to questioning after the experiment indicated that no Ss were aware of the construction of the list. The endings were noticed by none of the Grade 3 Ss, a few of the Grade 5 Ss, and most of the adult Ss but only a few of these reported all three verb inflections. Most Ss noticed that there were real and pseudowords but none realized the extent of the anagram-like construction.

Discussion

These results might seem to be disappointing, since they provide little evidence for a simple carry-over of a unit-forming principle from speech to writing. Adding a familiar tense ending did not increase the length of a potentially readable word under these conditions. A word, real or at least pronounceable, without an inflection was just as readable as one with. Yet the ending itself gave evidence of being treated as a unit. Correct endings, with inflected words, were more frequent than the equivalent number of correct last letters on other words, and the proportion of correct inflected endings to the total number of errors increased with age.

We propose the hypothesis that the base word and the morphological inflection are separate features of a word: that they are picked up independently; that each is a unit, but that the base, if it is meaningful, has some priority for the actively engaged reader. The hypothesis is further supported by the finding that words with inflections have a tendency to get another inflection substituted erroneously. It is as if the reader detected that the base word was inflected but did not perceive the right ending.

It has several times been suggested in recent years (Anisfeld & Knapp, 1968; Bower, 1967; Fillenbaum, 1969; Katz & Fodor, 1963; Morton, 1969) that words are recognized, or "stored," or recalled as a complex of features. Some of these features, certainly, are learned as the child learns speech, and those he has learned in hearing and speaking might be expected to have parallels in reading. A written word has, for one thing, graphological features. It is composed of letters having forms (that may be confusible, like b and d), it has a "word-form" and a characteristic length, and the letters are combined within the word by orthographic rules. There may be confusing orthographic similarities, such as reversal ("saw" and "was"). Any or all of these graphological features may be perceived, or on the other hand, go unnoticed in reading. Although a word is perceived visually in reading, it also has potential phonological features. Letters or letter-clusters or syllables can be "sounded out"; the word itself has a characteristic sound (which can be rhymed with other words, for instance, and thus be confusable, or it may have homophones). The potential confusions within these feature sets are mentioned because errors traceable to them tell us something about what features are getting processed (the reader perhaps searches within a feature category but gets the wrong feature).

A word has many obvious semantic properties. It has categorical relations (it belongs to a class, like a kind of fruit). It has similarity and contrast relationships (synonyms and antonyms). It indicates values (pleasantness-unpleasantness; Osgood's semantic-differential scales). That a word possesses such semantic features and that there is a search to find the proper descriptor can be manifested, again, by relevant errors. (The word may have been peach, but I report pear; it may have been lemon, but I report melon, a dual-feature confusion caused both by sharing a category membership and by an orthographic reversal.)

A word, finally, has syntactic features. These include the part of speech (verb, adjective, noun), the role of the word in the surface structure of the sentence (subject or object), and morphological inflections of the sort we investigated in our experiment. These inflections mark the word's stem and they are indeed essential features or descriptors that serve to specify the word. Our data provided evidence that the base word can be noted as having an inflection but that this is sometimes erroneously confused with another inflection. The subject may correctly perceive "listen" as inflected but produce "listening" instead of "listened." This type of error occurred relatively more often with fifth graders than third graders. This finding implies that inflections are perceived as belonging to one class of word-features and that there is progress in the pick-up of syntactic features as reading skill increases. We did not find a still greater proportion of such errors in the college students, because by that time they are simply not making many errors, especially in the well-structured endings.

We suggest that features of a word are processed, in reading, independently and sequentially. For a given task, one class of features, e.g., meaning, might be given priority over another, e.g., syntactic tagging. This hypothesis has been elaborated and the evidence for it collected in a separate paper (Gibson, 1970). Let us now apply it to this experiment.

The subjects were not able, with a short exposure duration, to read a longer word when it had been extended by an inflected ending, or syntactic marker. The base word alone (when it was a real word) got the priority. But the inflected ending could nevertheless operate as a unit and got priority for pick-up when the base word was badly structured. Furthermore, there was evidence that the inflection was perceived as belonging to a class of features, morphological markers, because there was a tendency to substitute other markers of the same class.

When a person attempts to read a word with an inflected ending, then, he must pick up at least two kinds of information: features of the base word (orthographic and semantic) plus syntactic information given by the ending. We were wrong in expecting an inflection, however familiar and unitary, to increase the length of a word that could be read under these conditions. With sequential processing of features, the reader's task would be increased no matter what the base word's length.

Words should not, therefore, be thought of as made up of pieces of a given length--either letters or other subunits--but rather as entities possessing information that can be divided into classes of features. It seems likely, furthermore, that these features are picked up independently and sequentially by feature class.

References

Anisfeld, M., & Knapp, M. Association, synonymity, and directionality in false recognition. Journal of Experimental Psychology, 1968, 77, 171-179.

Berko, J. The child's learning of English morphology. Word, 1958, 14, 150-177.

Bower, G H. A multicomponent theory of the memory trace. In K. W. Spence and J. T. Spence (Eds.), The psychology of learning and motivation. Vol. 1. New York: Academic Press, 1967.

Fillenbaum, S. Words as feature complexes: False recognition of antonyms and synonyms. Journal of Experimental Psychology, 1969, 82, 400-402.

Gibson, E. J. Perceptual learning and the theory of word perception. Invited address, Division 3, American Psychological Association, September, 1970.

Gibson, E. J., Osser, H., & Pick, A. A study in the development of grapheme-phoneme correspondences. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 142-146.

Katz, J. J., & Fodor, J. A. The structure of semantic theory. Language, 1963, 39, 170-210.

Morton, J. Interaction of information in word recognition. Psychological Review, 1969, 76, 165-178.

Neisser, U. Cognitive psychology. New York: Appleton-Century-Crofts, 1967.

Is Discovery of Structure Reinforcing?
The Effect of Categorizable Context on Scanning
for Verbal Targets

Eleanor J. Gibson, Yvette Tenney, and Martha Zaslow

Abstract

An experiment was performed to study the effect of meaningful structure (a category) in the context words of a scanning task, and to see whether the structure, if discovered and used, would be carried over by S and influence processing in a second task. Results showed that semantic structure in the context does not facilitate search time even when S searches for a target word defined only by class, nor does it decrease search time in a second task where S might look for structure in vain, suggesting that it has little utility for the scanning task. Presence of a common category increased recall of context words, but less than expected. It was concluded that more economical strategies are adopted by S in a search task and that semantic features of words are processed minimally if at all.

It has been argued by E. J. Gibson (1969) that external reinforcement (e.g., food, shock, or correction by an E) is not a necessary condition for perceptual learning. Reinforcement in perceptual learning is the reduction of uncertainty, achieved by finding structure (economical distinctive features, or superordinate structure) that reduces the information to be processed; the finding of structure is thus automatically and internally rewarding. It was hypothesized that an individual who has made use of the structure in one situation will search for it again under similar circumstances with more persistence than one who has not.

The discovery of an invariant relation is obviously economical, particularly for any intellectual task that involves reading, where there are a great many separate items to be processed perceptually. The question is how to test the hypothesis? The traditional learning task depends on external reinforcement. Typically, the experimenter informs the subject when he is right or wrong and rewards him for adopting a desired strategy. Experiments on concept learning, for example, usually require the subject to find out what features or dimensions the experimenter has in mind. This type of learning may have little significance outside of the laboratory. In real life, an individual frequently forms concepts on his own. The internal reinforcement which results from his own discovery of an invariant or a good strategy may be much more important than any prompting or correction from an outsider.

It was proposed, therefore, to study internal reinforcement in incidental learning, where the subject is given neither explicit instruction nor external reinforcement such as the experimenter saying "right" or giving a tangible reward. One can manipulate the context in incidental learning so as to discover what conditions facilitate learning without instruction. The question is, what is learned in incidental learning? Is structure in the context facilitating? If so, what kind of structure? The hypothesis is that whatever reduces information in relation to the task (has utility for reduction of uncertainty) is potentially reinforcing and will, if discovered, be retained.

Previous experiments (Schapiro, 1970; Yonas, 1969) have shown that economical graphic features are picked up and used in a search task or in a discrimination-decision task, and facilitate performance (decrease latency, for instance) quite without instruction. We proposed in the present experiment to study the effect of semantic structure in the context of a search task (Neisser, 1964). Specifically, we asked whether a category shared by the background words would be picked up without instruction, would speed the search, would be recalled later, and would be sought for again in a second search task. Whether such semantic structure would facilitate the rate of search for a target word was an empirical question, but it seemed reasonable that a highly related context would facilitate finding an unrelated target. If the structure proved to be useful, several predictions would follow. Context is generally poorly recalled in a search task (Neisser & Beller, 1965; Schapiro, 1970), but a categorical relation among context items, if discovered, should improve recall. There should be a decrement in search rates on a subsequent task in which the context is no longer easy to categorize, if the Ss continue to look for structure. Since searching for a category involves attending to semantic aspects of the context words, Ss previously exposed to categorizable context should recall more of the uncategorized background words in the second task than a control group not previously exposed to categorized lists.

Method

The experimental design involved four conditions, each calling for different groups of Ss. All Ss scanned 20 lists of 30 words each, searching for a target word. Two of the groups had a second stage consisting of ten additional lists.

Group CI (categorical, one-stage) searched for a target word embedded in a set of 30 categorically related words, chosen from a pool of 11 fruit words (apple, banana, berry, cherry, fruit, grape, lemon, melon, orange, peach, prune). The order of context words was randomized from list to list. The S was told that the target word was an animal, but he was not told the name of the animal. There were 11 different target words (tiger, horse, kitty, panda, zebra, mouse, sheep, skunk, puppy, burro, camel). The target word was selected at random, but no word appeared as target on two consecutive lists. The position of the target word was varied from trial to trial so that a rate measure of scanning time could be obtained. After S had scanned 20 lists, he was asked to recall the context words.

Group UI (uncategorized, one-stage) searched for the same target words as group CI, in the same list-positions and order of trial, but the targets were embedded in a context of taxonomically unrelated words, drawn from a pool of 11 words (angel, banker, brick, cheese, front, grain, laugh, medal, office, plate, porch). They were the same length as the context words of group CI, had the same mean word frequency, and had as many graphic similarities as possible (initial letters, common clusters, etc.). After S had scanned 20 lists he was asked to recall the context words.

Group CII (categorical, two-stage) was identical with group CI in its stage 1 task, except that the Ss were not asked to recall the context words. This group progressed immediately to ten more lists which contained new target words and new context words. The context words this time were not members of a category (anger, board, bridge, corner, flame, grade, light, march, ocean, place, phone). All the target words were parts of the body (ankle, brain, mouth, heart, lungs, tooth, elbow, chest, wrist, thigh, cheek). Again the S was not told the specific target words but only the class. They were again positioned through the list to allow a rate measure of scanning time to be obtained. At the end of ten scanning trials, the Ss were asked to recall the context words of these lists, and finally, to recall the context words of the stage 1 lists.

Group UII (uncategorized, two-stage) was identical with group UI in its stage 1 task, except that the Ss were not asked to recall the context words. Stage 2 for this group was identical with stage 2 of Group CII so that they were comparable except for their preceding varying experience with categorized or uncategorized context in stage 1.

Apparatus and Procedure

The Ss were run in only one condition, 21 Ss in each condition. The S was brought to the experimental room and shown the scanning box (described in detail in Gibson and Yonas, 1966a), and given a demonstration of its operation. The S pressed a switch to turn on the light which made the list visible, scanned the list rapidly from top to bottom, and pressed the switch again as soon as he found the target word. The initial press started a Hunter Clock-Counter, and the second press stopped the counter and turned off the light. E then read the counter and reset it. The lists were typed with a primary typewriter in uppercase letters, and mounted on heavy cardboard to slide easily into the scanning box. Two practice trials were given before the stage 1 trials began, and one before stage 2. Instructions are given in detail below.

Condition CI and UI

"This piece of apparatus is called a scanning box. When you press this button, a light goes on in the box and you will see a list of words. Start at the top of the list and scan down, as quickly as possible, until you find a target word that I'll tell you to look for. As soon as you find it, press the button again. That will turn off the light and also stop a clock which tells me how long it took you to find the word.

When you are ready to begin, look at the dot of light here at the top. Then push the button and scan down until you find the word you're looking for.

All right, the word we want you to look for is the name of an animal. There will be one and only one animal word in each list. Try not to miss it, but scan as quickly as you can. If you get to the bottom without finding it, press the button and turn the light off anyhow, and let me know that you couldn't find the word.

Are you ready? Go ahead."

(After trial 1 and 2 E asked what the target word was, and again after trial 11.)

After 22 scans, E said: "Now I want you to try to recall all the words you saw that were not animal words. You may write them here." If S did not remark on the category, but belonged to condition CI, he was asked after the recall: "Did you notice anything that the background words--the words that weren't animals--had in common with each other?"

Condition CII and UII

Stage 1. Same as CI and UI.

Stage 2. "Now the target word will be changed. Look for a word that is a part of the body. Again, there will be only one in each list. Scan as quickly as you can, but try not to miss the target word. Turn off the light when you reach the bottom of the list even if you haven't found the word."

(E gave one practice trial and asked what the target word was.)

After 10 scans with stage 2 lists, E asked for recall of context as follows:

"Now I want you to try to recall the words you saw in the second part of the experiment that were not parts of the body. Please write down all the background words you saw as you were searching for parts of the body."

After S had completed this recall, E said:

"Now try to recall the background words from the first part of the experiment when you were searching for animal words." (At the end, if there was a category and S did not mention it, he was asked the same question as Ss in group CI.)

If an S scanned through a list without finding the target word, the list was withdrawn and S went on to the next list. The list was presented again at the end of the series without informing S. E watched S's eyes on the first few trials to make sure he did not try to go back over the list a second time.

Subjects

The Ss were 84 students from the Introductory Psychology course who served as part of the work of the course.

Results

As a check on the validity of the rate measure, the time to locate a target was correlated with the target's position in its list for all subjects in all conditions. The mean of these correlations was over .80. This measure seemed highly satisfactory, especially since experimental and control conditions were matched for target position within each list.

Mean time per trial and mean time per line (rate of scan) are presented for the four conditions, both stages, in Table 1. Although the scanning time in stage 1 appears to be slightly faster

Table 1
Speed of Scan

	Stage 1		Stage 2	
	Time/Trial (in sec)	Time/Line (in sec)	Time/Trial (in sec)	Time/Line (in sec)
Condition CI	4.220	.185		
Condition UI	4.553	.212		
Condition CII	4.486	.210	5.329	.232
Condition UII	4.572	.206	5.429	.219

for the groups with categorized context words compared to those with uncategorized (groups CI and CII vs. groups UI and UII), the difference in time per trial is not significant ($t = 1.066$, $df = 82$, $p = n.s.$). Neither is the difference in time per line ($t = .8782$, $df = 82$, $p = n.s.$). It is evident that presence of context words which all belong to one category, fruit, does not facilitate time to find a target word, even though the target word is not specified except as a member of a class, and a class different from that of the background words.

Is the group which searched in stage 2 for a word embedded in a random set of context words slowed down by previous experience with a categorizable context? This result would be expected if they were looking for a category where one did not exist. But when scanning rates for group CII and group UII are compared in stage 2, there is again no difference. For mean time per trial, $t = .374$, $df = 40$, $p = n.s.$ For mean time per line, $t = .714$, $df = 40$, $p = n.s.$ If Ss in group CII were looking for meaningful relations as a result of the stage 1 experience, it did not slow them down.

Does the presence of a categorizable context make a difference in the number of words in the context that S can recall correctly? A comparison of groups CI and UI suggests that it does (see Table 2). No S in group UI reached the mean of recall for group CI (the highest correct recall of any S in group UI was four words). The same trend appears in a comparison of stage 1 recall

Table 2
Number of Context Words Recalled Correctly (out of 11)

	Stage 1	Stage 2
Condition CI	5.10	
Condition UI	2.33	
Condition CII	2.43	2.19
Condition UII	.57	2.43

for conditions CII and UII (recall was delayed for both these groups and is consequently less than for CI and UI). There is no difference in recall for groups CII and UII in stage 2, and the number of words correctly recalled is comparable to condition UI, as might be expected since there was no common category in any of these contexts. The fact that CII Ss recalled no more words than UII Ss in stage 2 implies that the former were not doing more semantic processing than the latter, as they would if they were trying to find a new category.

It appears, then, that the presence of a common category does have some effect on the number of context words that an S can recall. Also, all the Ss in group CI said that they noticed the category. In group CII, four Ss did not remember having noted a category at all in stage 1. None of these four remembered any of the context words of stage 1.

It is consistent with previous experiments with the scanning task that context words are poorly remembered. Neisser and Beller (1965) suggested that words implicitly rejected as targets are not stored or even "registered," but the context words in their lists were not semantically related or repeated. The same 11 context words appeared in our lists over and over; yet even in the group given a common category and tested immediately for recall, a mean of less than half the possible words was recalled. What were the Ss doing as they searched for the target word?

Discussion

In preliminary experimenting, before the procedure described here was adopted, the Ss were given a specific target word to search for with the same background words as those used in the present experiment. If the target was "horse," for instance, S was shown the word in the same type as the list to be scanned on a 3 x 5 card and told "Here is your target for this trial." We soon discovered that under these conditions, the Ss did no semantic processing whatever. They simply looked for a word beginning with H, and when they found one they looked at the orthography just enough to reject it or accept it as the target. Hence, in the present experiment we adopted the procedure of asking S to look for an example of a class, thinking thus that he would be forced to go through a stage of semantic processing, find the category structure, and use it. Search time for seven Ss run in the preliminary experiment in condition CI was somewhat faster than for CI and CII (stage 1) Ss run under the present procedure. The mean list time of the preliminary Ss was 4.156 sec compared to 4.353 sec, by no means as great a difference as we had anticipated.

Voluntary, detailed reports from two Ss throw considerable light on what may have happened in the present experiment. These Ss reported that they quickly saw that the background words were repeated, albeit randomly. They then looked, not for an animal word, but for any word with novel orthography, noting especially the first letter. In half the lists, the animal word actually did begin with a different initial letter than the context words. Our attempts at forcing semantic processing were thus foiled in favor of a more economical strategy, looking simply for a novel target, which would work equally well for the two sorts of lists. Neisser and Lazar (1964) found, in fact, that Ss could search for "any unfamiliar symbol" as rapidly as for any numeral, though not as fast as for a fixed familiar symbol. "Numeral" is a class identification, but the class is a limited one defined by a small set of graphic features, so that processing physical details, rather than semantic properties, would still have been possible, though more extensive processing would be required than for a single fixed target. Many Ss in our experiment remarked on graphic details like word length or initial letters when asked what the words had in common, suggesting that the semantic features were not the most salient features.

It now seems to us, with the benefit of hindsight, that the scanning task is one in which semantic features of words are processed, if at all, with low priority. Graphic details are highly economical features to focus on in locating the physical word target (Gibson & Yonas, 1966b). In looking for a telephone number, or a word in the dictionary, we scan for the initial letter and process an item only enough to reject it or accept it. Recognizing its meaning is a deeper process which seems to have little utility for this task.

Gibson (in press) has described a theory of how words are perceived which is relevant to this conclusion. Words are complexes of features--graphic, phonological, syntactic, and semantic. These features may be perceived independently, and the feature strategically chosen as the one having greatest utility for the task will be processed first.

The original aim of the experiment--to investigate the possibility that discovery of structure can be reinforcing--was thus not fulfilled because it turned out that the structure provided by the experimenter had little utility for the task and the Ss found a more economical strategy. We are pursuing the original question with other tasks and with structural relations of potentially greater utility for the task.

References

Gibson, E. J. Perceptual learning and the theory of word perception. Journal of Cognitive Psychology, in press.

Gibson, E. J., & Yonas, A. A developmental study of visual search behavior. Perception and Psychophysics, 1966, 1, 169-171. a.

Gibson, E. J., & Yonas, A. A developmental study of the effects of visual and auditory interference on a visual scanning task. Psychonomic Science, 1966, 5, 163-164. b.

Neisser, U. Visual search. Scientific American, 1964, 210, 94-102.

Neisser, U., & Beller, H. K. Searching through word lists. British Journal of Psychology, 1965, 56, 349-358.

Neisser, U., & Lazar, R. Searching for novel targets. Perceptual and Motor Skills, 1964, 19, 427-432.

Schapiro, F. Information extraction and filtering during perceptual learning in visual search. Unpublished doctoral dissertation, Cornell University, 1970.

Yonas, A. The acquisition of information-processing strategies in a time-dependent task. Unpublished doctoral dissertation, Cornell University, 1969.

The Effect of Orthographic Structure on Letter Search

Eleanor J. Gibson, Yvette J. Tenney,
Roderick W. Barron[1], and Martha Zaslow

Abstract

Ss were asked to search for a single letter target embedded in a context of (a) letter strings that were orthographically well structured though not English words; or (b) letter strings that were poorly structured and unpronounceable. Neither college students nor fifth graders differed in scanning rate as a function of the context structure. Reports from Ss revealed that orthographic structure was frequently noted, but was used inconsistently, if at all, and without benefit to the search. It was concluded that optimal strategies for processing verbal features change with the task and that pick-up of larger structural units is incompatible with fast letter search.

Earlier experiments have shown that orthographic structure, even in nonwords, is an important determiner of the ability to read a briefly displayed string of letters (Gibson & Guinet, 1971; Gibson, Osser, & Pick, 1963; Gibson, Pick, Osser, & Hammond, 1962; Gibson, Shurcliff, & Yonas, 1970). The explanation proposed was that as reading skill progresses the reader learns to process units larger than a single letter by taking advantage of letter dependencies, especially clusters constrained by rules of English orthography to a given position in the word. Does this kind of processing occur only when the S is set to read the string and report on it, or is the rule structure, once learned, picked up automatically, without special intention? Is proficient reading an automatic process in which the graphological and phonological rules operate in decoding without any particular attention or notice on the reader's part?

We proposed to investigate the question by examining the effect of orthographic structure in the background display of a visual search task. The scanning task used by Neisser (1964) was chosen, since it has been shown (Neisser & Beller, 1965; Schapiro, 1970) that non-target items are poorly remembered at the conclusion of the search and thus appear to be relatively unattended. However, Schapiro (1970) showed that contextual structural information that is useful for locating a target can be picked up without instructions. Subjects in his experiment searched for a single target letter embedded in rows of six-letter strings. These strings were made up of six different background letters, randomly permuted in every row, and re-randomized for every list. When S reached asymptote, a new context letter was introduced without his knowledge. This letter was positioned randomly either above, below, to the right, or the left of the target letter in each list so that it could serve

1. U.S.P.H.S. Postdoctoral Fellow (MH47179-01A1).

as a kind of "marker" or pointer for the target if noticed. Despite having reached an apparent asymptote, all the Ss proceeded to reduce their search time still farther after the introduction of the new letter that provided redundant information about the target location.

We wondered whether good orthographic structure, as opposed to the random arrangement of context letters, would be noticed and whether the structure would facilitate the search for a target letter by allowing the background to be processed in larger units. We also asked whether there would be an age difference in automatic, uninstructed pick-up of orthography. A target letter was therefore searched for in two types of background. In one, letter strings were constructed according to rules of English orthography and were thus "pronounceable" pseudowords. In the second, letter strings were constructed which were anagrams of these pseudowords but were un-English in orthography and unpronounceable. Children in the fifth grade and adults were compared on the task. Since the orthographically possible strings were also phonologically possible, it was of further interest to discover whether the Ss would pronounce them subvocally in scanning for the target letter.

Method

Subjects were assigned to one of two experimental conditions. In condition OS (orthographic structure), the horizontal strings of letters in the background were orthographically well structured and pronounceable. In condition NS (no structure), the horizontal strings of letters were orthographically unstructed and unpronounceable. It was originally planned to run Ss in both conditions, counterbalancing the order. Preliminary experimenting, however, suggested that the structure or non-structure of the condition to which S was first exposed led to expectations which influenced his behavior on the second condition and contaminated the results. A recently reported experiment (Aderman & Smith, 1971) confirms the potential effect of expectation on pick-up or non-pickup of structure. Therefore, we decided to run Ss in only one condition in order to give them a better chance to perceive the background structure when it was present.

Construction of Lists

A single target letter (N) was assigned for all lists. The lists were prepared in horizontal strings five letter wide and 30 rows from top to bottom. The five letters in each horizontal string (item) in condition OS were arranged to conform to English orthographic structure (e.g., GLURK). The corresponding item of the corresponding NS list was composed of these same letters permuted so as to be un-English (e.g., RKUGL). All the OS words were monosyllabic. The items containing the target letter were similarly constructed.

The target letter occurred an equal number of times (four) in each of the five horizontal positions in an item. Four different word structures were utilized to allow the N to appear in all five positions while maintaining the pronounceability of the item. The consonant-vowel arrangements were: 1) CVVCC (e.g., NEARL); 2) CCVCC (e.g., SNORK, SLAND); 3) CVCCC (e.g., BENTH); 4) CCVVC (e.g., THAIN). The corresponding NS item was of the same CV structure, with the target letter in the identical position.

Context items conformed to the same CV structures and were distributed over the four CV types in the same proportion as the target items. This rather elaborate control of CV structure was necessary because we had reason to think that Ss would notice very subtle changes in CV structure if they could be used to differentiate an item containing a target from other items.

A pool of context items was prepared containing 60 OS items paired with 60 corresponding NS items. Every letter of the alphabet was used, except N. A pool of target items was prepared containing 20 OS items, with the target letter N appearing in each one, paired with 20 corresponding NS items.

Twenty OS lists and 20 corresponding NS lists were constructed by drawing one item from the target pool (without replacement) and 29 items from the context pool (with replacement) for each list. No context item was permitted to appear on two successive lists or twice on the same list, so each context item appeared at most 10 times during the 20 trials. The positions of the context items were randomly determined for each list. The target item appeared in a different row, randomly determined, on each list so that a rate measure of scanning could be obtained (Neisser, 1964). The target letter's position was thus randomized both horizontally and vertically over the lists, but with equal distribution horizontally, and with representation in all segments vertically.

The lists were printed on white unlined paper with the standard business type of an IBM 1403 printer and were pasted on heavy cardboard for insertion in the scanning box.

Procedure

The S was taught how to operate the scanning box (described in detail in Gibson & Yonas, 1966a) and given three practice trials. A trial was initiated when the S pressed a button that turned on the light in the scanning box and also started a Hunter Clock-Counter. When S found the target, he pressed the switch again, turning off the light and the clock. The target letter on the practice trials was the same as on the experimental trials, but the context items and target items were different. An S's practice lists conformed to the construction rules for the condition to which he was assigned (OS or NS). The S then scanned through the 20 experimental lists, while E recorded his latencies. E asked him afterwards if he had noticed anything about the list construction and whether he had pronounced any of the horizontal strings to himself.

Instructions

The fifth grade Ss were instructed as follows:

The name of this game is Find the letter "N" as fast as you can. (E holds up a card displaying the letter N.) Let me show you how it works.

Hold this block in your lap and push the button with your thumb. Push it again and see what happens. You may practice pushing the button.

Turn on the light and look into the window. What do you see? (A practice list appears in the display window. E records whether S reports letters or tries to pronounce words.)

Look at the top line. Do you see the letter N anywhere in that line? Look at the second line. Is the N there? How about the third? Fourth? (If S did not see it there, E said, "Are you sure?")

Good. You have found the letter N. Push the button and turn off the light. (E inserts a new practice card.)

Do you see the light? (E points out lighted fixation point.) Watch it closely. When you are ready, push the button and start looking down the list. This time the N will be in a different place. As soon as you find the N push the button again. Ready? Go ahead. Where was it?

Always be sure to start at the top and look down the list. (E illustrates with finger moving straight down lighted window.)

If you get all the way to the bottom of the list (E illustrates with finger) and you still have not found the N, don't back up and look for it, but push the button and say "I didn't see the N." (E changes to practice card 3.)

OK? Let's try it again. Watch the dot. When you are ready, push the button and start looking for the N.

Instructions for the college Ss were essentially the same but the wording was adapted for adults.

If S missed a target, the list was presented again when all the other lists had been scanned through. E watched S's eyes for the first few trials to check whether he actually scanned from the top downward.

Subjects

There were 76 Ss from the fifth grade, 38 in each condition, and the same number from the introductory psychology course at Cornell University.

Results

The validity of the measure of scanning (mean rate of scan per line) was tested by correlating the mean time of a group of Ss to locate the target in each of the 20 vertically distributed list positions with the target position. The time should, of course, increase as the target's position is more remote from the initial position at the top of the scanning box. The correlations were .85 (condition OS, fifth grade), .86 (condition NS, fifth grade), .91 (condition OS, adult), and .90 (condition NS, adult). These correlations seem very satisfactory. They reveal, in addition, that fifth grade children are able to scan as instructed just about as systematically as adults.

The mean rate of scan per line to find the target letter is presented in Table 1 for both conditions and both age samples.

Table 1

Mean Rate of Scan per Line (in sec) to Locate a Target Letter

	Condition OS	Condition NS
Fifth grade Ss	.81	.88
Adult Ss	.38	.37

It appears that scanning rate was not affected in either age group by the orthographic structure of the context items. The only obvious difference is between the scanning rates of the children and the adults. An analysis of variance confirmed the trends mentioned, showing that age is the only significant variable ($F = 123.89$; $df = 1,148$; $p = < .001$). The fifth graders, although they were able to scan systematically down the list, did so much more slowly than college students.

Did the Ss in condition OS notice that the horizontal strings of letters conformed to English orthographic structure and were pronounceable as syllables? Did they, in addition, articulate them subvocally? An experiment by Baddeley (1971) on immediate memory for redundant and non-redundant sequences of letter names showed that acoustic confusions between letter names disappeared as redundancy within the strings increased. Baddeley suggested that Ss were coding the letters into larger units of speech sounds, such as syllables.

When our Ss were first shown a practice list, they were asked "What do you see?" In condition OS, 91% of the fifth grade Ss responded "words" whereas in condition NS, 88% responded "letters." During the experiment, some fifth graders in the pronounceable condition spontaneously reported the whole word in which the target letter was embedded, but the extent to which the whole word was subvocally pronounced without report is unclear. A few fifth graders who scanned exceptionally slowly in the pronounceable condition reported that they were saying the letters to themselves, one by one. A few pilot Ss who were given both kinds of list spontaneously remarked at the end that the pronounceable lists were "much easier because I can say the words to myself." Thus it seems plausible that some of the fifth graders treated the pronounceable strings as words and even pronounced them subvocally.

Of the adult Ss in condition OS, 74% said they saw "words," 10% said "nonsense syllables," and there were a few idiosyncratic responses such as "list" and "names." In condition NS, 89% said "letters" and again a few said "nonsense syllables." The distinction was not as clear cut as with the children, due to the greater variety of answers available to the university students. Reports about treatment of the pronounceable items revealed inconsistencies. About half the Ss thought they pronounced the items as words occasionally, but not all the time. It appeared from their remarks that some of these Ss vacillated from a "word" strategy to a "letter" strategy within a list, or changed their strategy part way through the experiment.

If some of the Ss were processing the pronounceable words as units, whereas others were proceeding literally letter by letter, the contrasting strategies might be reflected in the mean latencies of individual Ss. Although we were not able to classify each S on this basis with any confidence, it seemed possible that such a difference would show up in a frequency distribution of mean latencies for the individual Ss in each condition. If the distribution in condition OS were bimodal, as compared to a unimodal distribution for condition NS, it would suggest the use of two different processing strategies on condition OS. For each age group the range of latencies in each condition was divided into ten equal intervals and the number of scores falling into each interval was plotted. Distributions for both OS and NS conditions had only one mode in both age groups and the OS and NS distributions were similar in shape. Thus the comparison affords no convincing evidence for a division of Ss in condition OS into two types.

Discussion

It has been known for nearly a hundred years (Cattell, 1885) that the number of letters that can be read in a brief visual exposure is significantly greater when the letters form a word than when they do not. Recently, several experiments have shown that a target letter is more easily identified when it appears in a word than when it appears in a string of unrelated letters or is displayed alone (Lott & Smith, 1970; Reicher, 1969; Wheeler, 1970). James & Smith (1970), however, obtained apparently contradictory results. They measured latency of report on presence or absence of a probe letter in a 5-letter string and found no advantage for a word compared to an unpronounceable anagram of that word.

All these experiments differ from the present one in that the type of redundancy studied was the kind found in real, familiar words as opposed to non-words. It is possible that the meaning and familiarity of the real words, rather than their orthographic structure, could account for the positive results. Yet English orthography is rule-like and does provide strong intraword redundancy (Garner, 1962, p. 240 ff.). Since we were interested specifically in orthographic structure, two sets of pseudowords were contrasted that differed only in this type of redundancy and were not confounded with other types.

Why did the orthographically structured context in our experiment not facilitate discovery of the target letter, even for the adult Ss? To understand the results of the present experiment it is essential to consider the nature of the task. The task calls for discrimination, not recognition of meaning or even identification. The S was asked to scan a list of letter-strings searching for a specified single target letter. It is possible, in condition OS, to code the pronounceable letter strings as words. But is this coding, however economical for picking up the string, compatible with deciding whether a given letter is a part of it? The answer seems to be no. The analysis of an orthographic unit into its constituent parts evidently required additional processing time. The advantage of a structured background which can be dealt with in larger units is cancelled by the necessity of further processing for a specific letter. A search for the graphic features which distinguish the target letter N from the set of background letters is probably more economical in this task.

At first glance our results appear to be contradicted by experimental findings published by Krueger (1970) midway during our investigation. He reports that his Ss searched more rapidly for a target letter through real words and through pseudo-words with a high degree of approximation to English than through scrambled letter strings. There is an important difference concerning the target letter between his experiment and ours. He used a different target for each list while we kept the target the same throughout. His subjects, unlike ours, did not have the opportunity to learn to search for the distinctive graphic features of a single target letter. However, the James and Smith experiment also presented different letter targets for each display and found no advantage in the time required to locate the target in words rather than in their unpronounceable anagrams. Both our experiment and the James and Smith experiment used different Ss for the pronounceable items than for the unpronounceable items. Krueger's Ss were shown both types of items, randomly presented, which may have led to a different strategy.

Krueger concludes (p. 398):

> The fact that Ss exploited sequential letter and word dependencies to search faster indicates that they were not restricting their attention to the letter shape being employing a selective filter to exclude all other, unwanted information, but rather were engaging in broader encoding operations, similar perhaps to those in reading.

Our Ss were clearly not engaged in coding operations similar to reading in the usual sense.

While no direct test for subvocal articulation of the pronounceable strings during scanning was provided, it was expected that articulation, if it occurred, would increase latency. Since the two conditions did not differ in search rate, it seems unlikely that the typical S coded the pronounceable strings into subvocal speech. We know that articulation of letter names does not occur in the scanning task in such a way as to interfere with scanning rate (Gibson & Yonas, 1966b); Kaplan, Yonas, & Shurcliff, 1966). Phonological encoding of letter names and, when possible, of larger pronounceable linguistic units does occur in short term memory tasks (Baddeley, 1971), but typically not in a search task.

To return to the original question posed, does a proficient reader automatically process orthographic structures, whatever the specific goal his task has set him? The answer is not simple. Nearly all the Ss in this experiment did notice the higher order structure present in the OS condition and many of the adults thought that they used it some of the time. But the results indicate that the structure certainly did not reduce search time. In accordance with the theory (Gibson, in press) that the perceptual process is highly adaptive and tends to utilize only that information most economical for satisfying task demands, we conclude that the pick up of orthographic structure in this task was incompatible with the most economical search for a single target letter and hence was not used. The results of this experiment lend further support to the notion that the sequence in which the various features of words are processed depends on the task.

References

Aderman, D., & Smith, E. E. Expectancy as a determinant of functional units in perceptual recognition. Cognitive Psychology, 1971, 2, 117-129.

Baddeley, A. D. Language habits, acoustic confusability, and immediate memory for redundant letter sequences. Psychonomic Science, 1971, 22, 120-121.

Cattell, J. M. Ueber die Zeit der Erkennung und Bennenung von Schriftzeichen, Bildern und Farben. Philosophische Studien, 1885, 2, 635-650.

Garner, W. R. Uncertainty and structure as psychological concepts. New York: Wiley, 1962.

Gibson, E. J. Perceptual learning and the theory of word perception. Cognitive Psychology, in press.

Gibson, E. J., & Guinet, L. The perception of inflections in brief visual presentations of words. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 182-189.

Gibson, E. J., Osser, H., & Pick, A. A study in the development of grapheme-phoneme correspondences. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 142-146.

Gibson, E. J., Pick, A., Osser, H., & Hammond, M. The role of grapheme-phoneme correspondence in the perception of words. American Journal of Psychology, 1962, 75, 554-570.

Gibson, E. J., Shurcliff, A., & Yonas, A. Utilization of spelling patterns by deaf and hearing subjects. In H. Levin & J. P. Williams (Eds.), Basic studies on reading. New York: Basic Books, 1970.

Gibson, E. J., & Yonas, A. A developmental study of visual search behavior. Perception and Psychophysics, 1966, 1, 169-171. (a)

Gibson, E. J., & Yonas, A. A developmental study of the effects of visual and auditory interference on a visual scanning task. Psychonomic Science, 1966, 5, 163-164. (b)

James, C. T., & Smith, D. E. Sequential dependencies in letter search. Journal of Experimental Psychology, 1970, 85, 56-60.

Kaplan, G., Yonas, A., & Shurcliff, A. Visual and acoustic confusability in a visual search task. Perception and Psychophysics, 1966, 1, 172-174.

Krueger, L. E. Search time in a redundant visual display. Journal of Experimental Psychology, 1970, 83, 391-399.

Lott, D., & Smith, F. Knowledge of intraword redundancy by beginning readers. Psychonomic Science, 1970, 19, 343-344.

Neisser, U. Visual search. Scientific American, 1964, 210, 94-102.

Neisser, U., & Beller, H. K. Searching through word lists. British Journal of Psychology, 1965, 56, 349-358.

Reicher, G. M. Perceptual recognition as a function of meaningfulness of stimulus material. Journal of Experimental Psychology, 1969, 81, 275-280.

Schapiro, F. Information extraction and filtering during perceptual learning in visual search. Unpublished doctoral dissertation, Cornell University, 1970.

Wheeler, D. D. Processes in word recognition. Cognitive Psychology, 1970, 1, 59-85.

Is Discovery of Structure Reinforcing? The Role of Semantic and Syntactic Structure in Anagram Solution

E. J. Gibson, Yvette J. Tenney, and Ruth Sharabany

Abstract

Two experiments were performed to study the effect of syntactic and semantic structure on the solution of sets of anagrams. It was expected that solution would be facilitated when such structure was present; that it would be discovered by S and that S would search for such structure in a new but similar task. The Ss were third and fourth grade children. In Experiment I six sets of five anagrams arranged in sentence order were presented successively to two groups of Ss. The same anagrams were presented in scrambled order to a third group. One of the groups with sentence order was given a hint to facilitate finding the sentence structure. This group discovered the order and afterward, when given the opportunity, used sentence structure to facilitate new solutions. Few Ss in the no-hint group discovered the structure. In Experiment II, common categories were present for one group on six successive sets of anagrams (a new category for each set). This group profited very significantly compared to a control group without categorical presentation, giving evidence of selective utilization of taxonomic structure and persistence of the strategy on new problems. How discovery of structure could be utilized to aid solution is discussed.

The revival of interest in cognition over the last decade has generated renewed concern about the nature of reinforcement in learning. Externally applied reinforcement, even in the form of correction by an adult, appears to be the exception rather than the rule in perceptual learning, language learning (Brown, Cazden, & Bellugi, 1969) and the every-day learning of concepts. What internal process substitutes in these cases for external reinforcement, to select what is learned? It has been proposed (Gibson, 1969) that reinforcement in perceptual and perhaps all cognitive learning is primarily the reduction of uncertainty--the discovery of structure (e.g., an invariant relation or a recurring distinctive feature) that reduces information and increases the economy of cognitive activity.

In order to investigate this hypothesis, experiments were planned to test whether discovery of structure was reinforcing in the sense that a successful strategy (one that has utility and economy for the task at hand) would be adopted and carried over in an appropriate transfer situation.

Earlier attempts to explore this problem experimentally (Gibson, Tenney, Barron, & Zaslow, 1971; Gibson, Tenney, & Zaslow, 1971) failed to get to the heart of the problem, but revealed several important points. One, the structure that the experimenter builds in for S to find must have direct utility for the task; second, structural relations that are economical for one task may not be for another, even when the same material is involved; and third, structure is not necessarily picked up by an S (especially a child) just because an experimenter has put it there. For instance, recurrent spelling patterns which seem obvious to an adult and certainly reduce information in a most useful way in reading, are not easily discovered by the beginning reader (Gibson, Farber, & Shepela, 1967). How structure is discovered is a baffling problem in its own right. How little we know about it is exemplified by the current controversies over how syntax in language is learned by the young child.

While pondering what task would be more suitable for providing the opportunity to discover structure than the ones we had previously tried, anagram solution occurred to us as a good one. Early laboratory manuals (Foster & Tinker, 1929; Langfeld & Allport, 1916) made use of this task and described experiments which compared the effectiveness of solution with and without categorical relations among the anagrams presented. In Foster and Tinker's manual, the S in one condition was instructed to look for "eating words" or "house words." In the other condition S was given no set nor was there a common category. Solution when a set was given was said to be twice as fast as without it. Langfeld and Allport included a similar experiment, but the S was not given specific instructions. A set was expected to develop as solutions of related words were achieved.

Even more to the purpose are a series of experiments by Rees and Israel (1935). In their first experiment, S was given instructions to look for "eating" or "nature" words and the effect of the set was measured by the number of appropriate solutions made from ambiguous anagrams. But in other experiments, the set was allowed to develop without instruction via the use of anagrams whose only solutions were all meaningfully related. Sets were established equally as effectively by this method as by verbal instruction. Effectiveness was again measured by the percent of appropriate solutions in a test series of ambiguous anagrams. An interesting outcome of one of these experiments was that the S was not necessarily aware of the operation of an effective set. Some categories or relations were more effective than others. When the relation was part of speech (all nouns or all verbs) no set for solution was established at all. The Ss were unaware of the grouping by grammatical class and it appeared to have no useful intrinsic relation to the task. But when the relation was a common order of solution (permuting the letters of the anagram in a 54123 order), the set was 95% effective in a test series, although few Ss had any notion they were following a regular order. This type of intraword relation is obviously intrinsically useful to the task of anagram solution. For us, it had the drawback that it takes a fair amount of practice to build up. The drawback lay in the fact that we desired to test inter-problem transfer of a general strategy, and to work with children whose word-knowledge and whose span of interest were limited. There was also the objection that the letter order might be claimed to be effective because a chain of S-R associations was acquired through repetition of the same eye-movement pattern, and not because structural relations were picked up either consciously or unconsciously.

Our experiment thus made use of an anagram task in which semantic and syntactic relations could be built in, or omitted, with the expectation that solution would be facilitated or impeded as these relations were potentially present or absent or, after being present, removed. Two experiments were performed, with third and fourth grade children as Ss. In the first, a set of five anagrams was ordered so that sentence structure was potentially present (the anagrams were presented in sentence order) or absent (the same anagrams were presented in scrambled order). The expectation was that discovery of sentence structure in the set of anagrams should be reinforcing, and should generate a strategy of searching for sentence structure, because the redundancy of both word order and semantic relationship would reduce uncertainty and facilitate solution.

In the second experiment, semantic structure of a different kind was potentially present or absent. A set of five anagrams in one condition, when solved, all shared membership in a familiar, rather specific, category, such as animals, clothing, or fruit. In the other condition, the set of five had no common category membership. Again, the hypothesis was that finding a common category should facilitate solution and lead automatically to search for another category in a new set of anagrams.

Experiment I

Design

The experiment compared three groups of Ss, 16 in each group, all of whom attempted to solve 30 anagrams arranged in six sets of five anagrams each. In Group SS (sentence structure), each of the six sets of anagrams when solved in the order presented made a sentence, always beginning with a proper name and ending with a noun object (see Table 1). In Group NS (no structure), the same anagrams were presented in each set, but the order was scrambled so that their solution did not yield a sentence. In Group SS-NS, the first four sets of anagrams, when solved, produced a sentence. The last two, however, were scrambled as in Group NS. The expectations were that discovery of sentence structure in the first few sets in Group SS should generate a search for sentence structure, faster solution, and fewer failures to solve in the remaining sets; that Group NS should be slower than Group SS in the later sets and fail to find solutions more often; that Group SS-NS, if sentence structure was discovered in the course of the first four sets, should show interference in the last two sets compared to both Groups SS and NS, since they might be searching for structure that was not present.

Table 1

Words Presented as Anagrams in Sentence Order (Group SS) and in Scrambled Order (Group NS)

Group SS

Set 1	Set 2	Set 3	Set 4	Set 5	Set 6
Jill	Sally	Steve	Jan	Ted	Dick
helps	can	sells	baked	drank	makes
Mom	tell	ice	three	fresh	neat
clean	good	cold	big	cool	mud
house	jokes	pop	cakes	milk	pies

Group NS

Set 1	Set 2	Set 3	Set 4	Set 5	Set 6
clean	tell	cold	Jan	cool	neat
Mom	can	sells	cakes	Ted	mud
Jill	jokes	pop	three	milk	Dick
helps	Sally	ice	baked	fresh	pies
house	good	Steve	big	drank	makes

Material

All the words presented as anagrams were simple ones encountered before the end of third grade, and each had only one solution. They are presented in Table 1. Group SS-NS was given sets 1 through 4 exactly as in Group SS, and sets 5 and 6 as in Group NS.

Procedure

Each set of anagrams was arranged on a rectangular 12 x 12 in. metal surface, to be referred to as a "board." The letters were engraved on squares (1 x 1 cm.) of wood magnetized so that they adhered to the board when put in place (obtained from magnetized Scrabble games). The anagrams were arranged one above the other, the top one at the far left of the board, and each succeeding one moved slightly farther to the right. This diagonal arrangement from top to bottom was adopted to prevent, as far as possible, the appearance of a "list" which might (in fact probably did) hinder discovery of the sentence structure. Each anagram was covered with a strip of red ribbon attached to the side of the board.

The E presented a board to S with the anagrams properly arranged and covered. The S pulled away the top strip of ribbon and attempted to solve the anagram. He was allowed to rearrange the letters as he worked. E timed him and stopped him after 60 seconds if he had not stated the solution. If he had not, E arranged the letters to form the word and asked the S to name it. The words were left uncovered as S completed them, so that the whole ordered array was visible when the set was finished. The board was then removed and a new one presented.

Before S began the six experimental sets, he was given a practice board on which E had arranged S's own name as an anagram on the first row, and the letters AHT on the second. The instructions were as follows:

> These boards are made of a sticky material that rips open like this. (E picks up one to demonstrate.) The letters of this word are out of order. Do you know what the word is? (If S answered correctly, E said: "Good, you may put them in order." If S did not know, E said: "Go ahead and try rearranging the letters until you can tell what the word is.")
>
> OK. We'll try another word in a moment. Call out the word as soon as you think you know what it is. Don't be afraid to take a guess. I will stop this timer as soon as you call out the correct word. You can move the letters around on the board to help you, but I won't be watching. So be sure to call out the word as soon as you figure it out. Ok, try it (the second practice word, AHT).
>
> Don't forget to call out the word as soon as you can tell what it's going to be, even if the letters are still out of order on your workboard. I will stop the timer as soon as I hear you say the word. Then you can finish straightening out the letters afterward.
>
> Now we'll try some more words. They will be words that you know well and have probably come across many times in your reading.

In Group SS-NS a slightly different procedure was followed. Because we found in pilot experiments that some Ss (even adults) never noticed that the words as presented in Group SS made a sentence, we introduced a procedure that facilitated discovery of the sentence structure on the first four sets. (It would have been futile to look for interference on the last two sets if the structure had never been found.) Before uncovering each new anagram, the child was asked to read aloud all the words already displayed on the current board. This procedure was generally effective in revealing the sentence order, as determined by later questioning and by spontaneous remarks.

Not all of the children were able to read all of the words--so many, in fact, that E was forced to check whether the child could read those words which he failed to unscramble. Children who failed to read two or more of the words were not included in the results and were replaced. Nineteen Ss (nine girls and ten boys) had to be eliminated for this reason. They were distributed over all three groups so that their elimination could not have biased the results.

Subjects

The Ss had finished third grade and were enrolled in summer playground programs at Ithaca schools. They enjoyed the "game" and were well motivated.

Results

Table 2 compares the speed of solution for the three groups of Ss by boards (set 1 to set 6, in the order given), and by the summed time per group for the first four sets. (The sum of these four, rather than the sum of all six, is compared, because the last two sets of Group SS-NS differed in make-up from their first four.) A one-way analysis of variance was run on the measures for each row. The differences are significant on sets 3 and 4, and on the summed time for the first four sets. Group SS-NS (which was given a hint by being required to read the words all through) was the fastest in all three of these comparisons. Group SS, which had sentence structure in the word order but no hint, was next fastest. Group NS, which had scrambled word order, was slowest.

Table 2

Mean Time of Solution (in sec) for Sets of Anagrams 1 through 6 and Mean Summed Time for the First Four Sets

	Group SS	Group NS	Group SS-NS	F
Set 1	73.30	78.85	54.09	1.46
Set 2	98.84	102.80	74.87	1.18
Set 3	120.15	132.41	76.42	6.81*
Set 4	70.57	97.32	30.52	9.81*
Set 5	93.09	78.02	66.38	1.18
Set 6	127.36	131.90	138.48	0.22
Sum of 1-4	364.85	411.38	235.90	8.13*

*F significant at <.01.

The differences that were significant by F test were further tested in paired comparisons between Group SS and NS, SS and SS-NS, and NS and SS-NS by Tukey and Newman-Keuls tests. All the significant differences rose from comparison of Group SS-NS with the other groups (see Table 3). Group SS which had sentence order but no "hint" did not solve the anagrams significantly faster than the group given the anagrams in scrambled sentence order. Group SS-NS, which had to read the solved anagrams in order, solved faster than both the other groups in all the above comparisons. That they were not significantly

Table 3

Tukey and Newman-Keuls Comparisons of Differences between the Three Groups

	Groups SS and NS	Groups NS and SS-NS	Groups SS and SS-NS
Set 3	ns	*	*
Set 4	ns	*	*
Sum of 1-4	ns	*	*

*Significant by both Tukey and Newman-Keuls.

faster in set 2 suggests that the sentence structure is not perceived at once, even when S reads all the words in sequence. It appears that S needed at least one more set to discover that a sentence structure is present or will recur.

Errors (failures to solve in 60 sec) generally followed the same pattern as the time measure. Group SS-NS made fewest errors (a mean of 1.69 for the first four sets); Group SS next fewest (mean of 2.63 for the first four sets); and Group NS most (mean of 3.69 for the first four sets). The difference is significant by a one-way analysis of variance ($F = 4.49$, $df = 2,46$, $p = < .025$). Tukey and Newman-Keuls comparisons revealed that only the difference between Group NS and Group SS-NS was significant. When structure is perceived, with the help of a hint, errors are reduced significantly in comparison to a group given no potential sentence structure, but not in comparison to the group where structure was present but no hint given.

Another question of interest was whether Group SS-NS (structure plus hint for the first four sets, followed by two sets without sentence structure) slowed down on the last two sets, and perchance showed interference compared to Group NS which had the same two final scrambled sets. Table 4 shows the mean time spent on sets 5 and 6,

Table 4

Mean Time (in sec) on Sets 5 and 6, and 5 and 6 Combined

	Group SS	Group NS	Group SS-NS	F
Set 5	93.09	78.02	66.38	ns
Set 6	127.36	131.90	138.48	ns
5 and 6	220.45	209.92	204.87	ns

and on 5 and 6 combined. A one-way analysis of variance revealed no significant differences between the three groups. Group SS-NS was not significantly slower than the other two groups. Neither were errors (misses) significantly different.

What did the Ss report about noticing the sentence structure? In Group SS, where no hint was given, only 6 out of 16 Ss noticed the structure at some point (five girls and one boy). Since the others might have used the relationship without being able to verbalize it, E read them the first three words of the first four sets, and asked if they could complete each sentence. They remembered correctly a mean of 2 out of 8 possible words. Group NS was read the three initial words of the sentence and asked to guess how the sentence might have been completed. They had a mean of 1 out of 8 correct, less than Group SS, but the difference is hardly impressive. Group SS-NS, who had read the words aloud, all noticed the sentence structure at some point and remembered a mean of 6 of the possible 8 words.

Discussion

A striking finding in this experiment was the failure of most of the Ss to notice the sentence structure in Group SS. The few who did probably accounted for the slight (but not significant) superiority of Group SS over Group NS. But why did not all the Ss notice a simple, familiar, presumably obvious structure, clued by both syntax and meaning? The answer would seem to be, as in two earlier experiments where Ss did not use structure which we had thought patently there, that the structure was not intrinsically appropriate for the task. Syntax is not a kind of structure that normally has a strategic usefulness for the task of solving anagrams. The S concentrates on letters, clusters of high frequency and constraint by word position, and the process of recombining these. Nevertheless, he must consult semantic features at some point, to know that he has found a word. Processing word meaning would seem to be a necessary step in the task of solution, therefore, but syntactical relations between words are treated as irrelevant and not noticed.

The Ss who did find the syntax, by virtue of reading the words in sequence after solution, did indeed proceed to use it with profit. They were not instructed to, and thus bear out our original hypothesis that a superordinate structure, once found, will be noticed again and utilized without external instruction or reinforcement. The lack of evidence for interference in the last two trials of Group SS-NS may have been due to too brief an opportunity to allow the sentence structure to "sink in." Since S had only three chances to look for and find recurrence, he may have slipped back easily to an early strategy of simply recombining letters in the hope of finding a word.

It is worth noting that Ss quite often rearranged the letters so as to form the word, but still failed to recognize it as a word. This failure sometimes turned out to be due to S's poor reading ability, but not always. It is possible for even a good reader to arrange the letters correctly and not perceive a word. Like "semantic satiation," the processing in such a case does not seem to reach the semantic features of the word.

Because processing for semantic features of a single word does seem required for anagram solution, we decided to investigate the effect of a superordinate meaningful relationship (sharing a taxonomic category) to see if this type of structural relation, which should facilitate recognition of a correct arrangement as a meaningful word, would be discovered, used, and persist in a new problem.

Experiment II

Design

This experiment compared anagram solution in two conditions, one which contained categorical structure (Group CS) and one which did not (Group NS). Six categories were employed for Group CS, all with simple conceptual relations familiar to children in the fourth grade.* They were fruit, things to drink, animals, eating utensils, colors, and furniture. There were six sets, as in Experiment I, of five anagrams each (see Table 5). Group NS was presented with the

Table 5

Anagram Solutions by Category (Set) for Group CS and a Typical Scrambled Set for Group NS

Set (Category)	Group CS: Fruit	Drinks	Animals	Utensils	Colors	Furniture
	apple	water	bird	knife	blue	chair
	orange	milk	tiger	fork	red	bed
	grape	soda	bear	cup	green	sofa
	fig	coke	pig	dish	black	desk
	prune	juice	fox	glass	white	lamp
Set	**Group NS: A^1**	**A^2**	**A^3**	**A^4**	**A^5**	**A^6**
	dish	fork	fig	cup	sofa	fox
	grape	bed	bird	black	desk	green
	coke	prune	orange	bear	water	tiger
	lamp	soda	white	glass	knife	blue
	milk	apple	juice	chair	red	pig

*The categories and examples were tested first on a comparable population. The children were asked to give examples of a given category, and later asked the category membership of individual instances. We are indebted to Miss Alida Spaans for this work. The most frequently volunteered examples of a category were selected when possible, but if they had more than one anagram solution, the next most frequent item was chosen. The final lists of words were shown to the fourth grade teachers, who confirmed that they were familiar and readable by the children.

same anagrams, but the anagrams were distributed over the six sets so that no set had more than two words from the same category. All anagrams, as in Experiment I, had only one solution.*

Since categories (for children at least) differ in difficulty, especially when the instances included are constrained by the necessity of having only one anagram solution, the order of presenting the categories in Group CS was counterbalanced. Every category appeared in every order (e.g., color was presented, first, second, third, ..., sixth to different Ss). Five Ss received a given set in each of the six orders, for a total of 30 Ss.

Individual anagrams also differ greatly in ease of solution, and practice may have an effect. The arbitrary sets in Group NS were therefore constructed to control for these possible effects when comparing Groups CS and NS. The anagrams were counterbalanced over the sets for different Ss both by set and intra-set order so that any given anagram could be compared with itself in the same position over the sets in the two groups. Thus category difficulty, item difficulty (the specific anagram) and list position were controlled. The letters of a given word were always presented in the same scrambled order.

Procedure

The sets of anagrams were arranged on boards as in Experiment I. An empty practice board was presented first. E presented S with a pile of letters and asked him to choose those which formed his name and to put them in order on the board. Then E scrambled them, covered them with the ribbon, started the timer, and allowed him to uncover and rearrange them. Instructions otherwise were similar to Experiment I. The solved anagrams were left in view, as before, but S was never instructed to read them aloud. The time limit for a solution was 75 sec. If S had not solved the anagram in this interval, E stopped the timer and "helped" him solve it. S's spontaneous comments were recorded.

Subjects

The Ss were 60 children from the fourth grade (second term) in an Ithaca school, 30 in each group. Mean age was 9.6 years. The experiment was conducted in a mobile laboratory on the school grounds. Two Ss failed to read more than two of the words even after E had arranged them and were therefore dropped from the experiment and replaced.

---

*It was discovered after the experiment was in progress that one of the words (bear) had more than one solution. None of the Group CS Ss gave any other solution, however.

Results

Our expectation was that mean time to solution should be lower, overall, for the group presented with sets of anagrams belonging to common categories. It was also expected that the difference between the two groups should increase as the experiment progressed and Ss in Group CS had opportunity to discover category structure, search for it again, and find it recurrent. Figure 1 shows the mean time for solution per set for the two groups as a function of order of presentation of the set. Since each set order for Group CS combines all six categories averaged within it, and the same words, scrambled, in Group NS, the comparison is not confounded with either category or word difficulty. When an S failed to solve an anagram, latency was recorded as 75 sec. and the trial averaged in with the others. The mean solution time is less for Group CS throughout, and the difference between the two groups increases as the Ss progress through the sets. From set 3 on, the solution time for Group CS appears near asymptote, since 17 sec. is very short for anagram solution by a nine-year-old. Group NS appears to show some practice effect from set 1 to set 3, but solution time then rises again.

It was expected that time for solution within a set should drop for Ss in Group CS, as the S discovered the category. One word would have to be solved before S could even guess at a category, and two would be minimal for confirmation of it. Table 6 shows the mean time per word for the five words within a set, averaged over all sets,

Table 6

Mean Solution Time in sec Over all Sets for First, Second, Third, Fourth, and Fifth Words in the Set

Position	Group NS	Group CS
1	16.69	34.71
2	16.54	22.14
3	48.10	13.91
4	27.38	16.49
5	29.81	13.27

for Group CS and NS. It should be noted that the counterbalanced design provided that these should be the same words in each of the five positions, so difficulty of words as such is equated in the comparison. The words averaged in position 3, for instance, may have been harder than those in position 4, but this does not confound the comparison of the two groups. As expected, Group CS shows a trend toward an increasing gain over Group NS and its advantage does not occur until the third anagram within the set. The first two words, in fact, were solved more slowly than by Group NS, perhaps because S was searching for a new category, or perhaps because the last one persisted and impeded him.

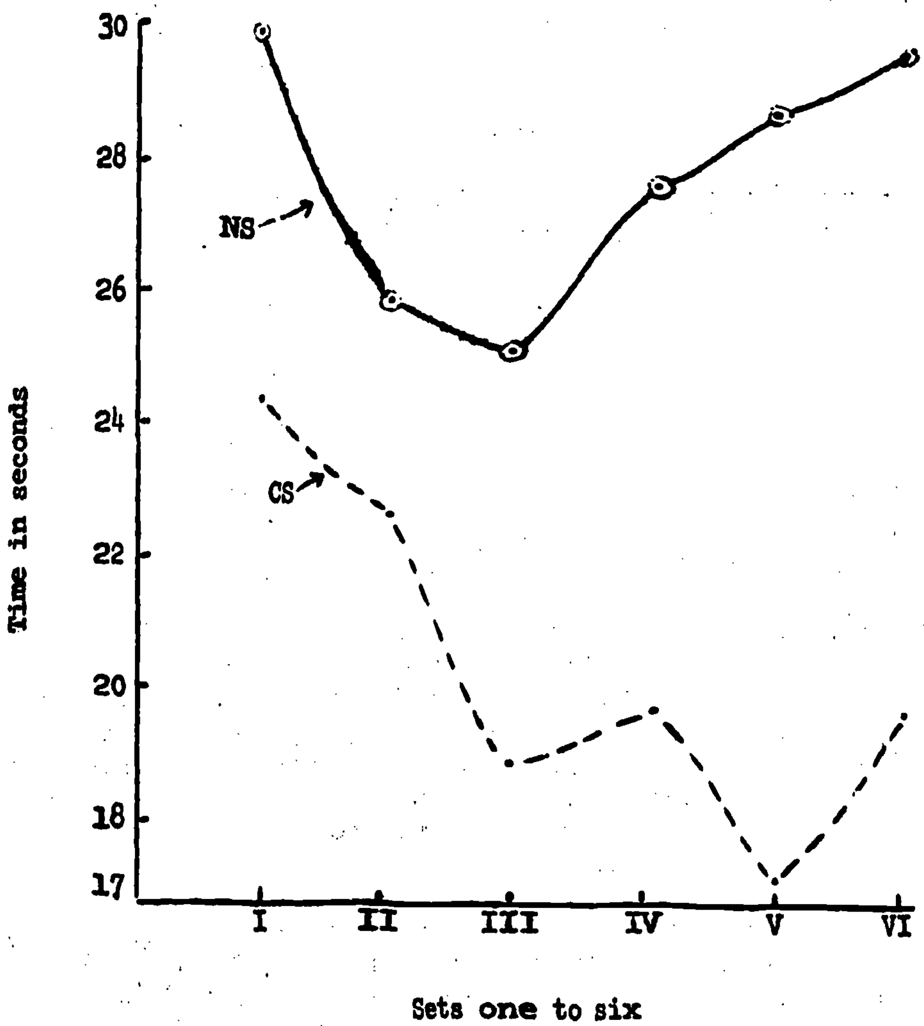

Figure 1. Mean solution time per set as a function of order of presentation of the set.

To test these trends, a three-way analysis of variance was run with groups (CS vs. NS), blocks (sets 1 to 6) and trials (anagrams 1 through 5 in a set) as factors, using the latency measure. The results are presented in Table 7. The analysis strongly confirms

Table 7

Analysis of Variance of Solution Latencies Comparing Groups, Blocks (in order) and Trials (in order)

	F	df	p
Groups	28.1688	1,58	<.001
Blocks	1.2546	5,290	n.s.
Trials	18.3525	4.232	<.001
Groups x Blocks	0.8000	5,290	n.s.
Groups x Trials	89.0403	4,232	<.001
Blocks x Trials	1.0979	20,1160	n.s.
Groups x Blocks x Trials	0.8721	20,1160	n.s.

the main trend, that Group CS solved the anagrams significantly faster than Group NS. The effect of blocks (order of sets) is not significant. In other words, there is no overall practice effect. As Figure 1 shows, the latencies for Group NS go up rather than down on the last three blocks. Although they drop for Group CS, there is not a significant interaction of blocks by group. Group CS's advantage shows up clearly on the first block, and evidently the apparent increase from then on is trivial.

The effect of trials (anagrams 1 to 5 in a set) is significant and the interaction of Groups x Trials especially so. Group CS's advantage as S goes down the board increases, indicating that S has found the category and is using it to advantage. The triple interaction (Groups x Blocks x Trials) is not significant, indicating that Ss in Group CS did not get better off on the last trial as they progressed from board to board. The advantage of the last trial is strongly present, evidently, on the first board.

One may check the above trends by looking at errors (failures to solve). An analysis of variance parallel to the above one was run on the error data. The trends mirror precisely those for the latency measure.

Another way of looking at the data is by category. A two-way analysis of variance was run with groups and categories as factors (see Table 8). The difference between category words is significant. Color words were easiest, followed by animals, dishes, furniture, drinks, and fruit, in that order. It is not that "fruit" is an ambiguous category for a child, but some of the words (probably "prune" and "fig") were hard or rather unfamiliar as words. Nevertheless, this effect of variable ease of words in the categories does not account

Table 8

Analysis of Variance of Solution Latencies
Comparing Groups and Categories

	F	df	p
Groups	28.1686	1,58	<.001
Categories	42.4334	5,290	<.001
Groups x Categories	3.3740	5,290	<.005

for the main effect of categories, because the interaction of group by category is significant. Group CS had a greater effect of word category, because the words were grouped for them in classes. A parallel analysis of variance was run on error data and showed precisely similar trends.

The differences by category between the two groups are largest for the easiest category and for the hardest one, but especially for the hardest one, by both measures. The latter finding suggests that knowledge of the category tipped the threshold to enable S to recognize something like "fig" as a word, whereas otherwise he would not have.

Did Ss in Group CS notice the categories within the sets? A number of them most assuredly did. Eleven of the 30 Ss volunteered comments indicating that they perceived the structure. Sample comments were:

> "Oh, these are colors!"
> "Once you got the first one, you got them all."
> "It keeps changing subjects." (After three sets were completed.)
> "I see a pattern in them. Do they all have a pattern?"
> "Oh! I see what's going on now." (While solving second word on second category.)
> "It must be a fruit."
> "I just saw blue and red and knew that all the rest must be colors."

Quite often, an S at the end of the set would spontaneously read sotto voce the whole set. He was never asked to and this never happened in Group NS, to E's knowledge. The first few Ss run were asked to recall all the words they could remember (this task had to be dropped because the extra time required was excessive). There seemed to be a tendency for Group CS to cluster, since appropriate errors occurred (pineapple, along with fruits; purple along with colors) and one S volunteered "I can't remember any more animals." Not all Ss, however, were so verbal about the categories and seemed to be less aware of them. Lack of spontaneous comment or even a lesser degree of awareness did not prevent their functioning, but may have been accompanied by a different strategy for solution. At least two "styles" of search were possible, as we shall discuss below.

Discussion

Is discovery of structure reinforcing? The results of these experiments confirm the view that finding relations in the material to be dealt with that create larger structural units (e.g., a sentence) or reduce the amount of information to be processed (e.g., a category) does lead to repetition of the successful strategy. The success is determined internally, by the S himself, for E did not instruct him to use such a strategy, tell him if he found one, or encourage him to search for another sentence or category if he had found one. This principle seems very obvious and simple as it has just been put, but it has nevertheless been overshadowed for three decades in favor of operant concepts demanding external reinforcers. It leaves open for further thought and research the problems, separate but related, of how structure is discovered and how it gets to be utilized. It is clear from earlier experiments that it will not be utilized if it does not increase the economy of cognitive processing; and that it is not very likely to be discovered if it has little intrinsic relevance for the task S is performing. The latter was the case with the sentence structure in Experiment I. Ss did not as a rule perceive it unless it was made very obvious, but once discovered it was facilitating and it was retained as a strategy. In a previous experiment, (Gibson, Tenney, & Zaslow, 1971) Ss were provided with categorical structure similar to that provided in the present Experiment II, but in a search task. The Ss generally noticed the category, but it did not facilitate search for a target word embedded in a display presented to S and it was ignored during the search.

In Experiment II, categorical relations were present within a set of anagrams in Group CS. They were found, and used with resulting economy in solution time and reduction of failures to solve. In anagram solution, S is searching for a word, but not a word defined for him and physically present in target form as in the detection-search task used by Gibson et al. Instructions in the anagram task were only to "find a word"--any word. What does S do, typically? He rearranges the scrambled letters, either mentally or by moving them about. Usually (if he is orthographically sophisticated), he tries potentially constrained clusters in highly probable places. For instance, if there is an S, a T, and an R, he tries them in initial position, where they often occur as a cluster in English orthography. If there is an N and a G, he tries them in final position for the same reason. If there is a vowel left over, he puts it in the middle and gets STRING or STRONG or STRENG or STRANG, etc. What if there are two vowels, an A and an E? He puts the vowel cluster EA, a frequent medial cluster, in the middle and gets STREANG. Is that a word? He consults his lexicon and decides that it is not. But the E can be moved to the end to yield another orthographically legal configuration, STRANGE. He consults his lexicon again and after a little time decides that it is. The configuration that he arranges must not only be orthographically and phonologically correct, it must also have an assigned meaning in the lexicon.

Making a pronounceable combination that is not a word can be disconcerting. A number of children in the present experiment, given the anagram for MILK, came up first with KLIM. It is perfectly pronounceable and could be a word, or a trade name (beginning like Kleenex). But a mental search through some kind of semantic network leads to a negative decision. How does this semantic search take place? Here is the step in the information processing where presence of a known category (or a sentence, given the first two words) yields economy. S does not have to search through the whole network. If he has found a category that says the word has the semantic features of inanimate, liquid, and potable, it is easy to reject KLIM and confirm MILK. When S rejects KLIM as a word, if he has no meaningful category in mind he may sit and stare at the letters, the pronounceable but wrong discovery inhibiting new possibilities.* But if he has a near-meaning to help, the mere placing of the M in initial position, with the vowel of course following, will "bring out the missing parts the way heat brings out anything written in lemon juice," to quote Brown and McNeil (The "Tip of the Tongue" Phenomenon, 1966). S knows most of the important semantic features of the word; an initial orthographic arrangement of letters, which in any case carries most information for orthographic and phonetic features of words, will bring the missing features to the surface at once, giving S closure and satisfaction at his knowledge of the exact match he has achieved.

If semantic features (Katz & Fodor, 1963) are assigned to words so that they can be classified in various ways, S is saved many steps by knowing what featural properties to look for, provided semantic features are processed in some sort of sequence. Sequential processing by some sort of hierarchical structure has been suggested and has been receiving investigation (Collins & Quillian, 1969; Quillian, 1967; Schaeffer & Wallace, 1969, 1970).

If searching for a word by semantic features were performed in parallel, that is, by checking all features classes at once, it is less easy to see why knowing a word's taxonomic class membership should help in the anagram task. If the search is sequential it should follow that the narrower the class (or, the farther down the conceptual hierarchy) the greater should be the facilitation. The "nature" class used by Rees and Israel, for instance, is very broad indeed and was in fact not very effective for all their Ss. It was actually slightly more effective when uninstructed, probably because

*Such inhibition has been found in anagram solution where the letters to be rearranged were presented to S in the form of an alternative word rather than scrambled letters (Beilin & Horn, 1962).

of the vagueness of the term "nature." The words in the series (e.g., a sequence like "plant, grass, ferns") defined the appropriate taxonomy better. The fact that less than five instances, in our experiment, were enough to establish a category and lead to faster discovery and utilization is probably due to the narrower categories employed.

It is worth noting again the apparent inutility of a grammatical form class in this task (Rees & Israel, 1935). Perhaps Ss in Experiment I did not profit by the syntactical relations as such. Knowing that a proper noun will come first, and a verb third does not, perhaps, help. The search that yields the final decision is a semantic one. Further experiments varying syntactic structure and semantic features such as class size and order in a potential semantic hierarchy should be instructive.

If the category assigned is a small and exhaustive one, like directions (North, South, East, West) or day of the week, S has another economical strategy available. He can use the category list to direct the search and skip the orthographic processing, by simply going through the list and selecting an instance. That some such process occasionally occurred is attested by the remark of one child: "I wish I could remember another silverware." She had found "knife" and "fork," but the category was broader than she thought. Colors were the nearest to an exhaustive category in this experiment. Color words in Group CS were the most speedily solved, and at least one S indicated that he was consulting a mental list of colors. So, this procedure can be highly economical, but it would not be with a category like "nature." However, all these strategies--the use of orthographic constraints, use of a known set of semantic features to eliminate part of the search through a semantic network, or simply checking a fixed list of instances--are cases of reduction of uncertainty by means of structural relations that result in economy of cognitive processing.

References

Beilin, H., and Horn, R. Transition probability effects in anagram problem solving. Journal of Experimental Psychology, 1962, 63, 514-518.

Brown, R., Cazden, C. B., and Bellugi, U. The child's grammar from I to III. In J. P. Hill (Ed.), 1967 Minnesota symposium on child psychology. Minneapolis: University of Minnesota Press, 1969. Pp. 28-73.

Brown, R., and McNeil, D. The "tip of the tongue" phenomenon. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 325-337.

Collins, A. M., and Quillian, M. R. Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 240-247.

Ervin, S. M. Changes with age in the verbal determinants of word-association. American Journal of Psychology, 1961, 74, 361-372.

Foster, W. S., and Tinker, M. A. Experiments in psychology. New York: Holt, 1929.

Gibson, E. J. Principles of perceptual learning and development. New York: Appleton-Century-Crofts, 1969.

Gibson, E. J., Farber, J., and Shepela, S. Test of a learning set procedure for the abstraction of spelling patterns. Project Literacy Reports, 1967, No. 8, 21-30.

Gibson, E. J., Tenney, Y. J., and Zaslow, M. Is discovery of structure reinforcing? The effect of categorizable context on scanning for verbal targets. Mss. prepared for Final Report, Research Grant OE29420446-1071-010 from U. S. Office of Education to Cornell University, 1971.

Gibson, E. J., Tenney, Y. J., Barron, R. W., and Zaslow, M. The effect of orthographic structure on letter search. Mss. prepared for Final Report, Research Grant OE29420446-1071-010 from U. S. Office of Education to Cornell University, 1971.

Katz, J. J., and Fodor, J. A. The structure of semantic theory. Language, 1963, 39, 170-210.

Langfeld, H. S., and Allport, F. H. An elementary laboratory course in psychology. Boston: Houghton Mifflin, 1916.

McNeil, D. A study of word association. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 548-557.

Quillian, M. R. Word concepts: A theory and simulation of some basic semantic capabilities. Behavioral Science, 1967, 12, 410-430.

Rees, H. J., and Israel, H. E. An investigation of the establishment and operation of mental sets. Pp. 1-26 in "Studies in Psychology from Smith College" (J. J. Gibson, Ed.), Psychol. Monogr., 1935, 46, (Whole No. 6).

Schaeffer, B., and Wallace, R. Semantic similarity and the comparison of word meanings. Journal of Experimental Psychology, 1969, 82, 343-346.

Schaeffer, B., and Wallace, R. The comparison of word meanings. Journal of Experimental Psychology, 1970, 86, 144-152.

Modality and Pattern Type: Some Problems in Temporal Pattern Perception in Children

Patricia Carlson

Abstract

Investigations of various tasks involving Morse Code-like stimuli are reviewed. The chief problems studied have been effects of visual versus auditory mode, and their interactions with spatial versus temporal presentation, crossmodal versus intramodal matching, and various effects of the patterns themselves. A study is reported in which it was found that in matching temporal stimuli second-graders perform better with auditory than with visual presentation, and that redundant presentation does not differ from auditory presentation alone. In terms of pattern effects, marked interactions with task requirements are found in the literature. The only pattern variable which was found to affect scores in the present study was relative preponderance of dots to dashes. Eliminated as sources of variance in this task were: absolute length of items, number of runs of identical elements, same or different number of runs within pairs, higher-order pattern structure such as reflection or complement, and preferred versus nonpreferred pattern, although all of the above have been found relevant to other tasks. It is suggested that this area of research, while interesting in itself, is not as directly relevant to reading research as has often been assumed, because of this sensitivity to task and because the spoken and written patterns involved in reading, as well as the relationships between them, are quite different from the patterns used in these studies.

Introduction

Recent investigations of auditory-visual matching, of auditory-visual redundancy, and of auditory-visual interference have been partly inspired by the presumed relevance of these abilities to learning to read. In reading, the patterned sequence of spoken language, in which the child is already fluent, is mapped to a patterned sequence of letters in a spatial, visual display. Several aspects of auditory-visual mapping have been studied with non-verbal displays: differences in visual versus auditory modality, differences in temporal versus spatial presentation, differences in pattern types, and interactions among these variables.

Visual Versus Auditory Modality and Temporal Versus Spatial Presentation

Birch and Belmont (1964) tapped out auditory patterns with a pencil, which were matched by 5- to 11-year-olds to linear spatial arrays of dot patterns. Blank and Bridger (1966) had children match temporally sequenced flashes of light to visually presented dot patterns. Both these studies found retarded readers poorer than normal ones.

A comparison of visual-spatial and auditory-temporal patterns was made by Muehl and Kremenak (1966). First-graders made same-different judgments, and were best at visual-visual matches, then at visual-auditory and auditory-visual matches, and finally worst at auditory-auditory matches. Sterritt, Martin, and Rudnick (1969) compared third-graders' performance on nine tests involving visual-spatial, visual-temporal, and auditory-temporal stimuli presented in various orders. Their data confirm the Muehl and Kremenak finding that visual-spatial matches are easiest. However, the results of the other tests are not clear from their published data.

Working with adults and with a different task (the subjects were asked to describe a repeating temporal pattern verbally), Garner and Gottwald (1968) found that at slow presentation rates (less than 2 per second) no modality effects were found; at higher rates, vision was poorer than auditory perception. The present study was designed in part to discover whether children also find visual temporal presentation more difficult than auditory-temporal, and whether simultaneous redundant presentation improves scores.

Since our study was run, Rubinstein and Gruenberg (1971), working with adults, reported testing intramodal and crossmodal matching using only temporal presentation. There were two different rates of speed, both well above the 2 per.sec. rate which marked the beginning of the relative decline in visual performance reported in the Garner and Gottwald study. Rubinstein and Gruenberg tried to equalize modality difficulty by adjusting the temporal relations for all visual stimuli to emphasize pattern differences; they also eliminated four subjects in each experiment because they did not meet certain visual criteria (four subjects perceived only three flashes instead of four at the rapid rate). With these adjustments, they found that at the higher rate, the visual-visual and auditory-auditory matches were not significantly different from each other, but at the slightly slower rate, the auditory-auditory matches improved and became significantly better than the visual ones. On the other hand, the crossmodal tasks, which were not different from the intramodal tasks at the slower rate, became significantly more difficult at the rapid rate.

To summarize, we know that for children of the age of most beginning readers, visual-spatial pattern matches are easiest, crossmodal (visual-spatial to auditory-temporal, auditory-temporal to visual-spatial) are intermediate, and auditory-temporal matches are most difficult. In addition, we know that for adults, visual temporal matches are more difficult than auditory temporal ones, and crossmodal tasks become more difficult than intramodal ones only at extremely rapid rates of presentation. Our study answers two remaining questions:

(1) Do children also find visual-temporal presentation more difficult than auditory-temporal?

(2) Does redundant presentation--i.e., simultaneous visual and auditory stimuli--improve performance or hinder it? If it is found that these matching tasks have relevance to learning to read, this might be useful information to educators, in addition to providing data about the integration of stimulus information from different modalities.

The remaining questions asked by our study involved the specific stimulus patterns themselves, to which we turn next.

The Effects of Particular Patterns

In the study already mentioned, Garner and Gottwald included an interesting pattern variable. Their stimuli consisted of repeating patterns which could be started at any point in the sequence. For example, basic pattern OXXOXXOXX could have initial sequence OXX, XXO, or XOX. The effect of this initial sequence turned out to be important only at the slower rates of presentation (under 2 per second), and disappeared at the higher rates. Garner and Gottwald suggest that the slower rates involve an intellectual "pattern learning" which should be distinguished from the direct "pattern perception" which occurs at the higher rates. They found, however, that the "preferred patterns" are similar across modalities and at different speeds. Royer and Garner (1970) elaborated on the idea of preferred patterns, finding that preferred organizations of long (9-element) repeating auditory temporal patterns included (1) balanced patterns with long runs at the beginnings and ends of the patterns, and (2) regularly increasing or decreasing length of run. Whenever a particular pattern was a preferred organization, so was its temporal reversal.

This pattern work is also interesting in view of a study by Shepard (1963). He analyzed Morse code learning at three stages. At the first stage--equivalent to same-different judgments about code letters by naive subjects--the responses were based primarily on the total number of elements in each stimulus, and the relative preponderance of dots versus dashes. At the second stage, when subjects were beginning to learn the 36 signals, the total number of elements remained important, but confusions increased among patterns which were similar except that one was a temporal reversal of the other ("reflection") or could be converted to the other by exchanging dots for dashes and vice versa ("complements"). These reflections and complements were not confused as often when presented in immediate succession for same-different judgments. The stimulus materials of the present study controlled for total number of elements by presenting paired items of equal length, and introduced the potential for errors of reflection and complementarity to see if they are characteristic primarily of the pattern-learning tasks of the type described by Shepard and by Garner and Gottwald, or have greater generality.

Incidentally, in the third stage of Morse code learning, in which presentation rate is increased, Shepard found that discrimination or length of runs of similar elements was the most frequent problem. This may reflect the same difficulty experienced by those of Rubinstein and Gruenberg's subjects who had to be eliminated from the fast rate experiment because they could not discriminate all four flashes. Extremely rapid rates of presentation seem to create new problems for subjects in both modalities.

Rubinstein and Gruenberg, in the study cited above, also included a pattern variable. Three patterns, paired sometimes with themselves and sometimes with each other, were presented to subjects in a matching task. Each pattern consisted of four clicks separated by long or short intervals. They describe the patterns in terms of the intervals; however, a description in terms of single, double, or triple clicks might be more appropriate phenomenologically.

Stimulus	Rubinstein and Gruenberg's description (A - long interval, B - short interval)	Alternative description (A - single, B - double, C - triple click)
1.	A B B	A C
2.	B A B	B B
3.	B B A	C A

Instead of being the same except for interval arrangements, Patterns 1 and 3 may have been more complex in the sense of having two types of element (single and triple) rather than the one type found in Pattern 2 (double). The importance of this description will be discussed below. It fits their results as easily as their description does: items beginning with Patterns 1 and 3 were equal in difficulty, and more difficult than items beginning with Pattern 2. When the same patterns were the final member of the stimulus pairs, they did not affect difficulty of matching.

In terms of pattern types, then, we have a number of hypotheses which might predict relative difficulty:

1. Longer items, in terms of total number of pattern elements, may be more difficult than shorter items.

2. If length is controlled, the higher-order features such as the pattern types found in Shepard's or Garner's learning tasks might emerge even at the immediate comparison level. The prediction would be that pairs which are reflections or complements of each other (i.e., are identical at a higher-order level) will be confused more often than pairs which differ in some other way.

3. Following Garner's suggestion, the number of runs of similar elements (as opposed to total number of elements) may be important. Perhaps pairs in which the number of runs is low (2 or

less) will be discriminated better than pairs in which the number of runs is larger (3 or more).

4. Pairs in which the number of runs is the same for each member might be more often confused than pairs in which the number of runs is different.

5. Garner's "preferred" patterns (i.e., symmetrical, balanced patterns with long runs at beginnings and ends of patterns, OR regularly increasing or decreasing length of runs) may be confused less often than nonpreferred patterns, at least when they are the initial member of the pair, as Rubinstein and Gruenfeld suggest.

6. According to Shepard's article, the most salient perceptual feature besides length in the immediate-comparison tasks he analyzed was the relative number of dots and dashes. This would predict that same-length pairs with equal ratios of the two types of unit should produce more errors, regardless of the pattern of arrangement of the units.

Our study was set up to analyze with statistical exactness only the first two pattern variable hypotheses listed above. However, the data can be looked at less formally to shed some light on which of the remaining four hypotheses are worth pursuing further.

Method

Material

Temporal patterns, consisting of arrangements of single and double tone units, were presented in pairs. The following types of pairs were utilized:

(1) The second pattern was identical with the first ("same").

(2) The second pattern was a "reflection" of the first; i.e., the order of the first pattern was reversed temporally in the second, last unit first, etc., as in the pair (.. . . . . ..); ("reflection").

(3) The second pattern was the complement of the first; i.e., the order of the pattern was maintained but every single unit of the first pattern was a double unit in the second, and vice versa, as in the pair (.. . . . .); ("complement").

(4) The second pattern differed from the first in some other way, but like the other types, the second pattern had the same number of tone units as the first (double units are counted as one unit); ("different").

Intervals after double units were slightly shorter than those after single units, in order to equalize absolute elapsed time as much as possible. Patterns with three units thus varied from approximately .81 second to .94 second in length, depending on whether the last unit of the pattern was single or double. Patterns with four units varied from about 1.19 seconds to 1.31 seconds, and those with five units varied from about 1.56 to 1.68 seconds.

Three different tests were composed. Each test consisted of 15 pairs divided as follows: 5 pairs in which the members had three units (two same, one reflection, one complement, one different); 5 pairs in which the members had four units (two same, one reflection, one complement, one different); and 5 pairs in which the members had five units (two same, one reflection, one complement, one different). The order was randomized within each test for both pattern type and number of units in the patterns. Each test could be presented in visual, auditory, or redundant modalities.

<u>Apparatus</u>

A Model 1-A Western Union, which activated a standard audio-signal generator set at 200 Hz, produced measured tones about .5 seconds long. These were recorded on a Wollensak tape recorder, then doubled in speed several times by successive retapings for a final tone length of about .06 seconds at 1600 Hz. Times for complete patterns are given above.

The subject sat in a slightly darkened room about 4.5 feet from a panel which included a small light bulb with dull reflector and a loudspeaker. The light bulb was activated by a transducer which converted the taped tones into the equivalent light flashes. Stimuli were presented with light alone, with sound alone, or together redundantly.

The experimenter controlled the tape recorder and recorded responses during the sessions. Each stimulus pair was played in its entirety, then the machine was stopped, <u>S</u> responded and then indicated to <u>E</u> when he was ready for the next stimulus pair.

<u>Procedure</u>

Each <u>S</u> was introduced to the stimulus materials with a group of simple examples. First he duplicated the pattern by clapping; then he was asked for same-different judgments on a few simple pairs. When he could perform satisfactorily on these warm-up items, the tests were run.

Each <u>S</u> performed all three tests, one in each modality. Order of test and order of modality were varied so that each <u>S</u> received a different combination of test and modality.

Subjects

The subjects were 15 children ranging in age from 7.4 to 8.11 who were attending a summer day camp.

Results

The results are presented in three tables. Table 1 shows the scores for modality, Table 2 the scores for number of pattern units (length), and Table 3 the scores for pattern type. The number of false "same" judgments made by each subject were ranked and compared by the Friedman non-parametric two-way analysis of variance.

Table 1

Error Scores by Modality of Presentation Across all other Variables

	Visual		Auditory		Redundant	
Subject	Errors	Rank	Errors	Rank	Errors	Rank
1	7	3	2	1	3	2
2	7	3	4	2	3	1
3	8	3	5	2	3	1
4	2	3	1	2	0	1
5	9	3	1	1	2	2
6	6	3	3	1	6	2
7	7	3	0	1	2	2
8	4	3	1	2	0	1
9	6	3	5	2	0	1
10	4	2.5	3	1	4	2.5
11	4	2.5	2.5	1	4	2.5
12	9	3	0	1.5	0	1.5
13	4	2.5	5	3	4	2.5
14	0	2	0	2	0	2
15	4	3	1	2	0	1
Σ Rj		42.5		24.5		25.0

$\chi_r^2 = 22.1$ for df = 2; $p < .001$

For modality, the χ_r^2 value was 22.1, $p < .001$, so the hypothesis that the modality of presentation has no effect is rejected. Examination of the data indicates a much higher number of errors in the visual mode; aural and redundant modes are almost equal.

For length of stimuli, the results show a χ_r^2 value of .93, $p < .70$, so we conclude that pairs of stimuli with 3-unit members are no easier to compare than 4- or 5-unit stimulus pairs in this task.

For pattern types, results show a χ_r^2 value of 2.7, $p < .30$, so we conclude that reflection and complement pattern types had no differential effect in this task.

Table 2

Error Scores by Length of Patterns Across all other Variables

	Length 3		Length 4		Length 5	
Subject	Errors	Rank	Errors	Rank	Errors	Rank
1	5	3	3	1.5	3	1.5
2	2	1	6	3	4	2
3	4	2	6	3	2	1
4	2	3	0	1.5	0	1.5
5	4	2.5	3	1	4	2.5
6	6	3	5	2	4	1
7	2	2	2	2	2	2
8	0	1	2	2.5	2	2.5
9	3	2	3	2	3	2
10	2	2	2	2	2	2
11	3	2	2	1	4	3
12	3	2	3	2	3	2
13	4	2.5	4	2.5	3	1
14	0	2	0	2	0	2
15	1	2	2	3	0	1
ΣR_j		32.0		31.0		27.0

$\chi_r^2 = .93$ for $df = 2$; $p < .70$

As noted above, we are not able to analyze our results confidently any further, since items were not equated for number of runs, preferred patterns, or relative preponderance of single to double units. However, mean scores for each subject were calculated for these other variables and the differences for each subject within each variable were subjected to a sign test to give us a rough idea of which hypotheses are most promising for further experimentation.

The number of runs hypothesis predicts that the 14 pairs in which the number of runs is low (2 or less) should be discriminated better than the 13 pairs in which the number of runs is larger (3 or more). Results show that only 4 out of 14 subjects performed in the predicted direction ($p = .18$). There is thus no support for this hypothesis in the data.

A related hypothesis predicts that the 21 pairs in which the number of runs is the same for both members of the pair might be confused more than the 6 pairs in which the first member has a different number of runs from the second. Results show that only 3 out of 13 subjects performed in the predicted direction ($p = .092$), which certainly provides no support for this hypothesis.

Table 3

Error Scores by Pattern Type across all Other Variables

Subject	Reflection Errors	Reflection Rank	Complement Errors	Complement Rank	Different Errors	Different Rank
1	3	1.5	3	1.5	4	3
2	5	3	3	1	4	2
3	4	2	3	1	6	3
4	1	2.5	0	1	1	2.5
5	4	2.5	4	2.5	3	1
6	6	2.5	3	1	6	2.5
7	2	1	4	3	3	2
8	2	3	1	1.5	1	1.5
9	4	3	2	1	3	2
10	2	1.5	3	3	2	1.5
11	4	3	3	2	2	1
12	3	2	3	2	3	2
13	3	2	3	2	3	2
14	0	2	0	2	0	2
15	2	3	0	1	1	2
ΣR_j		34.5		25.5		30.0

$\chi_r^2 = 2.7$ for df = 2; $p < .30$

Rubinstein and Gruenfeld suggest that their results on a similar task were in part a function of "goodness" of pattern as described by Garner. When it is the first member of the pair, they claim, a preferred pattern (symmetrical, or regularly increasing or decreasing runs) will be confused less often than a nonpreferred pattern. In means calculated from 20 preferred and 7 nonpreferred initial patterns in our data, 7 out of 14 subjects performed in the predicted direction. There is thus no support for this hypothesis in the data. The discrepancy of this finding with Rubenstein and Gruenfeld's results will be discussed below.

Finally, there is the relative preponderance hypothesis. Shepard suggested that the most important pattern feature in matching tasks (besides relative length, which we controlled) is the ratio of dots to dashes, regardless of arrangement. This hypothesis predicts that the 11 pairs with equal proportions of single to double units should be more confusable than the 16 pairs in which the first member has a different ratio of single to double units from the second member. Results show that 11 out of 14 subjects performed in the predicted direction, p = .06. We tentatively conclude that the data suggest that this hypothesis is the most promising one to account for pattern variables in matching tasks of the type investigated.

Discussion

Our findings, although largely negative, have implications both for the study of temporal pattern perception and for the relevance of similar studies to reading. In summary, we know the following facts:

1. In matching tasks involving Morse Code-like patterns, results with children indicate that:

a. Visual-spatial to visual-spatial matches are easiest;
b. Auditory-temporal to auditory-temporal matches are more difficult;
c. Visual-temporal to visual-temporal matches are most difficult;
d. Crossmodal matches are intermediate in difficulty, except at extremely rapid rates of presentation when they are more difficult than intramodal matches (for adults);
e. Redundant presentation of temporal patterns does not alter scores from the auditory-temporal level.

2. In terms of various pattern features that have been hypothesized to be relevant in this type of task, it was shown that neither increasing length (3 to 5 units) nor similar higher-order pattern features (reflection and complementarity) increased the number of errors. In addition, there was no hint in the data that certain other variables which have been found in other tasks using stimuli of this type contributed to the variance. The total number of runs, the similarity of number of runs across the two members of a pair, and the "preferred" patterns found by Garner seemed to have no effect. Since they are quite clearly important in the tasks Garner used, which involve mechanical reproduction of patterns in synchrony, verbal reports and verbal predictions, we should be cautious in discussing "temporal pattern perception" as though it were independent of task. Different features of the patterns seem to be relevant for different tasks. Matching, immediate verbal report at a slow rate of presentation, immediate verbal report at a more rapid rate of presentation, prediction of next element, recall, and recall at extremely rapid rates of presentation have been shown to have differential effects on the relevance of various pattern features. A lengthy discussion would be required to do justice to the implications of these differences. Meanwhile, we should note that the only hypothesis that even predicted results in the right direction was the one suggesting that relative preponderance of one type of unit to the other, regardless of arrangement, was the important variable for the matching task.

The Rubinstein and Gruenberg results seem at first glance to run counter to the above claim that Garner's pattern types are not influential in matching tasks. However, their results are probably better interpreted in terms of the "relative preponderance" hypothesis. The preferred pattern they mention is the pattern of intervals between clicks in their stimuli, which intuitively seems artificial. If the clicks are considered instead, we find that Pattern 2 (the easiest) consists of two similar units (two double clicks), whereas Patterns 1 and 3 consist of two different units (one single, one triple click). In terms of "relative preponderance" 1 and 3 are identical, 2 is different and therefore less confusable. It seems fair to say that the Rubinstein and Gruenfeld study supports the "relative preponderance" hypothesis for matching tasks, just as Shepard's analysis and the present study do.

Relevance to Learning to Read

As noted at the beginning of this paper, much of the interest in its topic has been stimulated by its presumed similarity to the task of learning to read, in which translations between a temporal, auditory display and a visual, spatial one must be made. However, I feel now with Pick (1970) that these tasks have very limited relevance to the topic of this paper and related ones for the following reasons:

1. We have seen that in terms of modality efficiency, the reading task, involving auditory-temporal and visual-spatial presentation, is already optimal. This is supported by results with sequential visual presentation of letters (Kolers & Katzman, 1966; Newman, 1966). The latter type of presentation is extremely difficult to read.

2. In terms of hypothesized crossmodal matching problems, we have found that there is difficulty beyond intramodal matching only at extremely rapid rates of presentation. This seems to have little relevance to the actual situation in which children are learning to read, for reading is self-paced. Further, even if some sort of crossmodal matching may be hypothesized as a process in reading, it must be at a much more abstract level than the one tapped by these tasks. Convergence could only occur at a semantic or syntactic level, if transfer is presumed to take place.

3. Most important is the extreme difference between the patterned stimulus strings of speech and of writing as compared with the stimulus patterns in these studies. We have seen that the relevant features of even these simple Morse Code-like patterns shift drastically with different task requirements. It seems foolish to try to generalize from these findings to the reading situation in which the task, the spoken and written patterned strings themselves, and the relationships that hold between the spoken and written patterns are so radically different. The surface information in spoken language and that in written language is related in a far more arbitrary, complex, and imperfect way than the information in the tones and light flashes of these studies. There are no amodally similar features in phonetic and graphic linguistic displays.

Work on code pattern perception and learning is, of course, interesting in its own right; but it should not be justified by its spurious similarity to the reading task, nor conclusions naively generalized to the reading process.

References

Birch, H. G., and Belmont, L. Auditory-visual integration in normal and retarded readers. Amer. J. Orthopsychiat., 1964, 34, 852-861.

Blank, M., and Bridger, W. Deficiencies in verbal labeling in retarded readers. Amer. J. Orthopsychiat., 1966, 36, 840-847.

Garner, W. R., and Gottwald, R. L. The perception and learning of temporal patterns. Quarterly Journal of Experimental Psychology, 1968, 20, 97-109.

Kolers, P. A., and Katzman, M. T. Naming sequentially presented letters and words. Language and Speech, 1966, 9, 84-95.

Muehl, S., and Kremenak, S. Ability to match information within and between auditory and visual sense modalities and subsequent reading achievement. Journal of Educational Psychology, 1966, 57, 230-238.

Newman, E. B. Speed of reading when the span of letters is restricted. American Journal of Psychology, 1966, 79, 272-278.

Pick, A. D. Some basic perceptual processes in reading. Young Children, 1970, 25, 162-181.

Royer, F. L., and Garner, W. R. Perceptual organization of nine-element temporal patterns. Perception and Psychophysics, 1970, 7(2), 115-120.

Rubinstein, L., and Gruenberg, E. M. Intramodal and crossmodal sensory transfer of visual and auditory temporal patterns. Perception and Psychophysics, 1971, 9, 385-389.

Shepard, R. N. Analysis of proximities as a technique for the study of information processing in man. Human Factors, 1963, 5, 33-48.

Sterritt, G., Martin, V., and Rudnick, M. Sequential pattern perception and reading. Reading disability and perception. Proceedings of the 13th annual convention of the International Reading Association, Newark, Delaware, 1969, 61-71.

The Effect of Redundant Spelling Patterns in a Verbal Discrimination Task

Eleanor J. Gibson and Nancy Rader

An important, if not the most important trend in perceptual development is progressive economy in the extraction of information from stimulation (Gibson, 1969). Any way in which information can be reduced appropriately for performance of the task at hand is adaptive and exemplifies the trend. Detection of distinctive features of things, extraction of invariants from transformations over time, and processing information in larger structural units are general cases of increasing economy. A particular case would be the discovery of recurrent invariant spelling patterns in English orthography (e.g., CVC followed by a marker E tells the reader that the vowel, whatever it may be, has the "hard" pronunciation), and the consequent simplification of the reader's task. Young children do not perceive these patterns easily or automatically (Gibson, Farber, & Shepela, 1967), but older children and adults have somehow incorporated them as part of their linguistic competence. Taxonomic categories of lexical items, recurrent spelling patterns, and relations that can occur with widely different elements (like being the "odd" one of a set) are sometimes referred to as "collative" features, because they reduce information. Developmentally, we suspect that they are perceived and used more and more effectively and that, once found, they are retained and transferred to similar situations. Discovery of the economical strategy automatically operates as a reinforcer in the absence of external intervention.

How this development of cognitive economy occurs, at what age, and under what conditions is little understood. To study it, one needs to set up a task in which redundancy, economical structure, or some rule-like principle can operate; to give the _S_ opportunity to discover it and to use it as an alternative to a possible but less economical strategy; and to present this opportunity "incidentally" in the sense that _E_ does not tell _S_ what the redundant feature or structure is, or when he has found it. The question is whether, how, and at what age the child will find and use the more economical strategy, and transfer it readily when a similar (but not identical) occasion occurs.

To investigate this question, the present experiment used a simple discrimination task which could be learned in either of two ways, made possible by presenting together alternative sets of cues. By selecting one set, the child had four unrelated choices to learn; by selecting the other set, which contained a verbal redundancy (common spelling pattern and rhyme), he had only a two-choice problem to solve. Following learning of the first task, a similar transfer task was presented which was immediately soluble if the economical (redundant cue) solution had been found and used in the first stage of learning.

Method

Task

The task presented S with four separate stimulus displays, projected from the rear one at a time on a small screen before him. He was faced with two buttons, one to his left and one to his right, one of which he must learn to press for each display. A light went on when S pressed the "correct" button, and he was also informed by E whether he had pressed the correct one or not.

The stimulus displays consisted of a printed word surrounded by a line contour. There were four contours, a rectangle, a circle, a triangle, and an octagon. The contour inscribed around a given word was always the same. All the words contained four letters and were ones presumably within the S's reading competence. The total vocabulary used in all conditions of the experiment was: look, book, cook, fall, tall, ball, call, tent, sing.

Design

There were two groups (conditions) in the experiment, one provided in its original learning with the economical, redundant cue (Group E-R) and a control group without it (Group C). Thus Group E-R had the possibility of selecting the more economical (collative) cue or not, while Group C could only learn the task as four separate discriminations or choices. Both groups were run in two stages, the original learning task and a transfer task. The transfer task was the same for both groups (see Table 1 for a summary of the design).

In Group E-R, the four words chosen as part of the stimulus display in Stage 1 were two rhyming pairs which were also spelled alike; look and book, fall and tall. As Table 1 shows, each word was circumscribed with a different contour which could, without regard to the word, have been used to "cue" the discrimination. The rhyming words were both paired with the same button, so that Ss who selected the collative feature as "cue" need learn only that one spelling pattern went with one button and the other pattern with the other.

In Group C, the four words chosen as part of the display in Stage 1 contained no rhyming pairs. They were look, fall, tent, and sing. The S could choose either the word or the contour as the "cue" for which button to press, but in either case he was forced to learn a four-choice discrimination; neither the word nor the contour was more economical.

In Stage 2 (transfer), the same contours were again displayed, but now each contour was paired with the opposite response button, so that S, if he had used contour as a cue, would have to learn to reverse his choice. New words were displayed, the same words within the same contours for both groups. There were two rhyming

Table 1

Summary of Experimental Design

Group E-R

Stage 1 Display	Stage 1 Response	Stage 2 Display	Stage 2 Response
LOOK	R	TOOK	R
BOOK	R	COOK	R
FALL	L	BALL	L
TALL	L	CALL	L

Group C

Stage 1 Display	Stage 1 Response	Stage 2 Display	Stage 2 Response
LOOK	R	TOOK	R
TENT	R	COOK	R
FALL	L	BALL	L
SING	L	CALL	L

pairs of words; took and cook, and ball and call. These fell into the same rhyme and spelling groups as the redundant pairs in Stage 1. The rhyming-response relation remained the same, so that an S in Group E-R could transfer directly on the basis of rhyme, but not on the basis of the contour. Group C had had no opportunity to observe rhyme and common spelling pattern in Stage 1, so an S in this group would have to find this redundancy, if he did, after he arrived at Stage 2 learning. Actually, Ss in Group C did have some opportunity to transfer on the basis of rhyme, however, for two of the individual words in their Stage 1 displays rhymed with one or the other of the rhyming pairs in Stage 2, and maintained the same response relationship. This was an error in design, for it lessened sharply the contrast between the two groups, making it more difficult to test the hypothesis that Group E-R (at some stage in development) should discover the rhyme-and-spelling redundancy and not only learn faster than Group C in Stage 1, but show more immediate transfer in Stage 2.

Procedure

The display apparatus and the response were demonstrated to S and the following instructions given:

> You are going to see something appear on this screen. What you will have to do is figure out which of these buttons goes with what comes on the screen. So each time you see something on the screen, press the button that you think goes with it. You'll be able to figure out which button is the right one because if you press the button that goes with what you see, this light will flash on telling you that that is the right button; and if you press the button that doesn't go with what you see, the light won't flash on and so you'll know that was the wrong button. At first you'll have to guess, but as you continue you'll be able to figure out which button is right.

The projected display remained on the screen for 2.5 sec. after the S made his choice of response. Following an inter-trial interval of 4 sec. the stimulus for the next trial appeared on the screen. The slides were prearranged in a random order, with the stipulation that no word appear more than two times consecutively and neither position be correct more than three times consecutively. Trials continued until S responded correctly on 10 consecutive trials, or until he had completed 60 trials. If he failed to reach criterion by 60 trials, learning was discontinued.

When Stage 1 was completed, the E told S that there would be a brief rest period. During this interval, E changed the slides and then explained to S that the experiment would continue as before but with some new slides. S was again run to a criterion of no errors for 10 consecutive trials or 60 trials if the criterion was not met by this time.

After Stage 2, the words were presented without the circumscribed contours to see whether any Ss had relied on contour alone as a cue. E also questioned S as to how he decided which button to press.

Subjects

The Ss were 40 children from the second grade and 40 from the third grade, half assigned to each condition. They were run individually in a mobile laboratory on the school grounds.

Results

The mean number of errors before reaching criterion are presented in Table 2 for both age groups, both experimental conditions, and both stages. The third graders made fewer errors than the second graders when the rhyming cue was available, but had approximately the

Table 2

Mean Number of Errors by Age Groups, Conditions, and Stages

	Condition E-R		Condition C	
	Stage 1	Stage 2	Stage 1	Stage 2
Grade 2	13.4	5.6	14.7	7.2
Grade 3	10.7	3.5	14.5	3.2
Grades Combined	12.05	4.51	14.6	5.2

same number of errors as the second graders in Stage 1 of Condition C, where there was no redundant cue. As for the two conditions, the trend is in the expected direction, but not very strong. The groups with the economical cue available in Stage 1 made fewer errors than the control groups during learning in Stage 1; a mean difference of 1.3 errors for second graders, and 3.8 errors for third graders. In Stage 2, the group which had had the economical cue available in Stage 1 surpassed the control group in the second grade, but this was not true among third graders. If a third grade S could take advantage of the redundancy in spelling patterns, he apparently did so in Stage 2 whether he had had previous relevant practice or not. (The previous practice was, as we pointed out earlier, to some extent relevant; if S noticed that two words in Stage 2 rhymed with a word in Stage 1, he could try the response that had been correct for the rhyming word in Stage 1 and find the problem solved for Stage 2.)

One can ask how many Ss solved the problem on Stage 2 immediately; that is, on or by Trial 2. There was no grade difference here, but there was a small difference between the two conditions. With Ss combined, 30% of Group E-R solved by Trial 2, whereas only 20% of Group C did.

Several analyses of variance were run to test these differences. A two-way analysis of errors in Stage 1 with condition and grade as factors yielded no significant differences. A similar analysis of errors in Stage 2 again yielded no significant differences. Despite the consistent predicted trends, therefore, we must conclude that the effect of the redundant verbal cue was not strong enough, by contrast with the control group, to override individual differences and to allow a significant developmental trend to appear.

Endeavors were made following Stage 2 to discover from the Ss what strategy they had used. It is noteworthy, and rather disappointing, that only 50% of the Ss in Group E-R reported using either rhyme or common spelling patterns and showed evidence of transfer by their use. Only 25% solved the problem immediately in Stage 1 and transferred perfectly in Stage 2. Spelling patterns that have very high saliency as common features for adults simply do not for children at this stage of reading development. Of the 50% of the Ss who did not report using the rhyme or common spelling, some said they used individual words, some used contour, and a few could give no report about the relation of cue to response.

The most promising trend in the results is Grade 3's tendency to surpass Grade 2 whenever the rhyme (redundant cue) was present, in both Stage 1 and Stage 2, but to perform no differently when a four-choice discrimination was the only available learning strategy. Separate two-factor mixed design analyses of variance were run with age as a between-S variable and stage as a within-S variable for both Groups E-R and C. Stage was a highly significant variable ($p < .001$) in both, so in general there was improvement from Stage 1 to Stage 2. But there was also a hint of an age by stage interaction ($p < .20$) in Group C which is explained by the failure of third graders to surpass second graders in Stage 1 of the control group. In the experimental group there was almost zero variance due to this potential interaction, consistent with the observed trend.

Discussion

The results thus point in the direction of an increased ability with age to take advantage of an economical strategy, but the evidence is weak. Why was the redundancy not significantly effective in reducing errors in Stage 1, at least? There are several possible reasons that might account for this result. It may have been relatively easy for children in Group C at the ages we tested to learn a four-choice discrimination. The fact that the two grades did not differ in Group C, Stage 1, is at least consistent with this notion. If this were true, it would have the effect of lowering the potential advantage of redundancy. However, a fair number of Ss did not meet the learning criterion within 60 trials (11 from Grade 2 and 8 from Grade 3), so we are more inclined to believe that the individual differences were so large as to make it hard to demonstrate variance due to selection of the redundant cue, or that a higher grade level was necessary for it to appear strongly.

Our first step in investigating these possibilities was to rerun the experiment with a within-S design, in order to cut down subject variance. Twelve Ss who had just completed the second or third grade were run, each S taking part in both Groups E-R and C with different sets of words. Stage 2 (transfer) was omitted, since time did not permit both two conditions and two stages. Half the Ss began with Condition E-R and half with Condition C. The Ss were equally distributed in this order by grade.

Due to exhaustion of the available subject population, only these 12 Ss could be run, but examination of the data made it clear that this design was fallible. The results were contaminated by the effect of the condition that was run first. Ss who began with the non-redundant four-choice discrimination actually made more errors, on the average, with the structured, redundant problem that followed than they had with the one which presented no economical solution. Those who began with the redundant, rhyming cues available tended to be slightly better on it than on the non-redundant four-choice discrimination that followed. Order of condition was thus a critical factor, and an unwanted one.

The next step, therefore, is to perform the experiment with new grade levels. Ss beginning Grade 2 and Grade 5 are being run, each S in only one condition. A two-stage transfer design was again selected. Several changes in material were made. The contours around the words have been omitted, and none of the words in Stage 1 of Condition C rhyme with words in Stage 2 of that condition, thus eliminating the possibility of transfer in Condition C that existed in the present experiment. New words have been chosen so that this experiment will provide two of the conditions for a second experiment described below. The words are presented in Table 3. Ss are also asked to read the words before the learning session begins. These changes should make it clear whether the economical strategy can be selected and utilized and whether stage of development (grade) is critical.

A further experiment has been planned to investigate the question of relative saliency of spelling patterns (orthographic structure) as cues. Was it the failure to attend to orthography that made only 50% of our Ss use the rhyming redundancy, or was it immaturity of the ability to adopt an economical strategy? If we compared printed words as cues with pictures, contrasting redundant with non-redundant conditions in both cases, some light should be thrown on this question. Pictures of objects that have rhyming names have been prepared on colored slides, with two sets available to make a transfer design possible, as before. The experiment will have four conditions: 1) a non-redundant four-choice printed-word condition; 2) a redundant four-choice printed word condition;* 3) a non-redundant four-choice

*These two conditions will have been provided by the foregoing experiment.

Table 3

Condition E-R

Stage 1		Stage 2	
Stimulus	Response	Stimulus	Response
KING	R	YARN	R
RING	R	BARN	R
CAKE	L	BOAT	L
RAKE	L	COAT	L

Condition C

Stage 1		Stage 2	
Stimulus	Response	Stimulus	Response
NOSE	R	YARN	R
CAKE	R	BARN	R
KING	L	BOAT	L
BELL	L	COAT	L

picture condition; and 4) a redundant four-choice picture condition. The design is illustrated above with the words. The pictures, when named, yield the same words.

It will be noted that the collative principle for transfer includes both spelling pattern and rhyme for the printed words, since that is the case in the reading situation which we wished to investigate. Ss in the picture conditions will be asked to name the pictures at the beginning of the experiment to make sure that they use the intended rhyming names and thus make a mediating phonetic cue possible. It can be expected that these Ss will spontaneously generate the names of the pictures and perceive the phonetic relationship, since Locke (1971a, b) has shown that four-year-old children recall more pictures whose names rhyme than nonrhyming control pictures. If second graders can use a collative principle, but do not yet perceive redundancy in spelling patterns, the picture condition should excell the word condition when redundancy is available. We expect that fifth graders will find the redundant spelling patterns at least equally facilitating.

Conclusion

Much work has gone into trying out and revising a technique for investigating the development of strategic use of economical, collative linguistic principles. The question is of interest for cognitive development in general and in particular for reading, which is the epitome of a cognitive task. It is the use of the economical principles and rules that accounts for the difference between the poor reader and the skilled one. If the design presently under trial is satisfactory, we shall be able to investigate other important collative principles for the reading task, such as related semantic features.

References

Gibson, E. J. Principles of perceptual learning and development. New York: Appleton-Century-Crofts, 1969.

Gibson, E. J., Farber, J., and Shepela, S. Test of a learning set procedure for the abstraction of spelling patterns. Project Literacy Reports, 1967, No. 8, 21-30.

Locke, J. L. Acoustic imagery in children's phonetically-mediated recall. Perceptual and Motor Skills, 1971, 32, 1000-1002. (a)

Locke, J. L. Phonetic mediation in four-year-old children. Psychonomic Science, 1971, 23, 409. (b)

The Involuntary Pick-up of Meaning from the Printed Word by Beginning Readers

Yvette J. Tenney

For the proficient reader the meaning of the printed word seems to be grasped so immediately and automatically that it is difficult to ignore. Meaning, however, does not seem to "jump out" for the beginning reader, who labors over each word. How and when meaning comes to reside in the printed word is a problem for a theory of reading acquisition (Gibson, 1970). In order to clarify the issue, an experiment was conducted on the development of immediacy of pick-up of meaning from the printed word.

A popular method for demonstrating the compellingness of the printed word is the Stroop test, in which the subject has to ignore the meaning of a color word and name the color of the ink in which it is written. When the meaning of the word is different from the color of the ink, interference results.

The Stroop test has been administered without modification to children of different ages (Comalli, Wapner, & Werner, 1962; Rand, Wapner, Werner, & McFarland, 1963; Schiller, 1966). In these studies children as young as seven years old did show strong interference effects. In fact, the younger the child, the more difficulty he had with the task. Presumably the older children developed more efficient strategies to overcome the difficulties. Rand et al. (1963) analyzed the kinds of errors contributing to the interference. Although the youngest group (around seven years old) made the greatest number of mistakes, they were lowest on one kind of error: actually reading the interfering word out loud. This finding seems to support the notion that the younger reader is less compelled to read what he sees.

Problems arise in trying to test even younger, more inexperienced readers. Only one Stroop test study involved a whole group of first graders (Schiller, 1966). These subjects, unlike their older schoolmates, did not show any interference. This result has interesting implications for a theory of reading, for it suggests that the proficient reader has learned something more than simply how to decode words to sounds: He has learned to pick-up the semantic information conveyed by the words much more directly and automatically (Gibson, 1970). Unfortunately, Schiller's report leaves open the possibility that his first-grade subjects could not read all the color words in the first place. He does not report reading errors, but he does show that the first graders, unlike any of the other children, took longer to read the list of color words (printed in black) than to name a comparable list of color patches. Usually reading words is much faster than naming colors. Ligon (1932) reports that the difference is constant across ages from first through ninth grade, so there is strong reason to suspect the words were too advanced for Schiller's first graders.

The present experiment was designed specifically to answer the question: Given that a beginning reader can read some words without hesitation, will he have trouble ignoring the meaning of those words, or will he still be able to look at them the same way he looks at words which he cannot read? A modified version of the Stroop test was devised, based upon some of the earliest words in the child's reading vocabulary. The subject's ability to read these particular words, among others, was assessed beforehand so that any failure to show interference could not be attributed to an inability to decode the words. The reading test also made possible an analysis of interference effects as a function of the reading level as well as the grade of the child.

Method

The Task

The subject is seated in a make-believe car with a gas pedal which hums when depressed by the foot and a horn which "beeps" when pressed with the finger. He learns to "make the car go" by stepping on the pedal when he sees green (once for a single green patch; twice for a sign with two green patches), to "beep on the horn" when he sees yellow (once or twice depending on the number of patches), and to make no response at all, i.e., "stop everything" when he sees one or two red patches. He is then instructed to ignore any writing which might appear on the signs and to pay attention only to the color of the writing. In the Interference condition, the signs actually contain the words: STOP, STOP STOP, GO, GO GO, BEEP, OR BEEP BEEP. Each sign is printed in red, green, or yellow ink at random, with the restriction that compatible word-color combinations occur only ten percent of the time. In the Control condition, the nonsense words VAPS, JY, and ZOBE are substituted for STOP, GO, and BEEP respectively.

Our task differs from the traditional Stroop test in several ways. The usual color words have been replaced by easier words. Secondly, motor responses to the colors substitute for the usual verbal color naming responses. Thus our subjects had to perform a motor act (make the car go, beep the horn, or stop all activity) in response to the symbolic meaning of the color of the ink. An increase in errors in the Interference condition compared with the Control condition can only be attributed to the involuntary pick up of meaning from the real words. This point is important because one could argue that the interference in the traditional Stroop test stems partly from an automatic reading response which interferes with the color naming response apart from consideration of meaning. For example, Klein (1964) demonstrated that a small but significant interference in color naming is found when non-color words are introduced. Since we were interested specifically in the pick up of meaning, rather than in the simpler act of decoding from a graphic to a phonetic representation, the motor task was chosen.

Materials

The color patches, words, and nonsense syllables were presented on 168 slides. The slides were Kodalith negatives made by photographing black stimuli on a white background. The colors were produced by placing a red, yellow or green filter over each slide. Since the slides were negatives, the printed letters were white and thus, seen through the filter, appeared to be printed in color on a black non-reflecting background.

Apparatus

The make-believe car consisted of a chair, a movable foot pedal which activated a doorbell when depressed, and a battery-powered bicycle horn. The slides were shown on a Carousel slide projector triggered by two Hunter timers. The projector could also be operated manually. Activation of the slide changer, foot pedal, and horn was registered on a Rustrack 4-track event recorder.

Procedure

Subjects were run individually in a laboratory trailer. A preliminary reading test was administered, consisting of twelve hand-printed words and nonsense words which the subject had to pronounce. A maximum of eight seconds was allowed for each word. The test consisted of STOP, GO, and BEEP (the critical words), VAPS, JY, and ZOBE (the corresponding nonsense words), STEP, SPOT, GOT, NO, and KEEP.

The subject was seated in the driver's seat of the make-believe car and was familiarized with the gas pedal and horn. After responding to verbal instruction "stop," "go," and "beep" given by E, he was shown the first traffic signs. The first set of 24 practice signs consisted of single and double patches of red, yellow, or green. He responded to each sign at his own pace and was corrected if he made a mistake. He then tried the same set, without correction, at a predetermined rate (two seconds per slide). The second set of 24 practice signs consisted of the critical words in incompatible colors mixed with the nonsense syllables. The subject was instructed in connection with this set: "These signs will have writing on them. But do not pay any attention to the writing. Just watch the colors the way you did before." The subject responded to these signs slowly with correction and then proceeded at the two-second rate without correction. After this practice, the subject was instructed to prepare himself for a long ride and told that his responses would be recorded. The test consisted of sixty signs presented at the two-second rate. The signs contained nonsense words for the Control subjects and critical words for the Interference subjects.

Design

Subjects were divided into different levels on the basis of the reading test. The following classification scheme was used:

1) reads none of the critical words,
2) reads one of the critical words,
3) reads two of the critical words
4) reads all three critical words, but only some of the other real words, and
5) reads all the real words.

Subjects within each reading level were assigned, in order of appearance, alternately to the Interference or the Control condition. Girls were assigned separately from boys, to insure an equal representation. The only difference between the two conditions occurred in the final test, where Interference subjects saw the critical words while Control subjects saw the corresponding nonsense words. The pattern of correct responses, determined by the color on the signs, was identical for the two conditions.

Subjects

Subjects who were completing first grade were tested in the late spring. Subjects who had completed second grade were tested during the summer and early fall. For convenience the two groups will be referred to simply as first and second graders.

Results

Reading Test

Seventy-two first graders were tested. Eight of these subjects could read less than two critical words; they were not counted in the experiment. Of the remaining sixty-four subjects, thirty-two could read two critical words (STOP and GO), and thirty-two could read all three critical words. (Six of these subjects in fact had perfect scores on the reading test.) These two groups will be referred to respectively as the below-average and the above-average first grade readers.

In the second grade, fifty-six subjects were tested. Only four were unable to read all the critical words; they were eliminated from the final analysis. Of the remaining fifty-two subjects, thirty-two had perfect scores on the reading test, while twenty subjects missed at least one of the non-critical words. These two groups constitute the above-and below-average second grade reading groups.

The average percentage of the real words read correctly by the subjects used in the experiment is given in Table 1. As a result of the procedure of assigning subjects to conditions, described under Design, half the subjects in each of these four derived classifications served in the Interference condition and half in the Control condition.

Table 1

Percent of Real Words Read Correctly
by First and Second Grade Groups

	First Grade	Second Grade
Below-average	40% (N = 32)	84% (N = 20)
Above-average	77% (N = 32)	100% (N = 32)

Analysis of Errors

The record of responses made by each subject in the 60-slide test was scored for number of errors. An error consisted of an incorrect response or a failure to respond. A single response to a double sign (e.g., BEEP BEEP) or a double response to a single sign was counted as an error. The data were examined in terms of grade (first versus second), reading level (above-average versus below-average) and condition (Interference versus Control). The average number of errors (out of a possible 60) made by Ss in each group is summarized in Table 2.

Table 2

Mean Number of Errors made by Ss in Grades 1 and 2
in Control and Interference Groups

	First Grade		Second Grade	
	Control	Interference	Control	Interference
Below average Readers	4.438	6.563	5.100	3.900
Above average Readers	3.688	6.500	4.000	5.875

Preliminary inspection of the data suggests an interference effect for all except the below-average second grade readers. In both grades, the below-average readers seem to show less of an interference effect than the above-average readers. Grade does not seem to matter so much as reading level. A three-way analysis of variance was conducted to determine the significance of these trends. Only the main effect of condition was significant ($F = 8.986$, df = 1,108, $p < .003$). A two-way analysis of variance, with grade and condition as factors, was performed separately for above-average readers and for below-average readers. The above-average readers showed a significant effect of condition ($F = 14.207$, df = 1,60, $p < .001$), with no other significant effects. The below-average readers showed no significant effects at all. Thus we can be confident that above-average readers in first and second grade are susceptible to interference. But we cannot conclude the below-average readers show significantly less of an effect than above-average readers, since the interaction between reading level and grade did not reach significance in the three-way analysis reported above ($F = 2.334$, $df = 1,108$, $p < .129$).

A two-way analysis of variance, with reading level and condition as factors, was performed separately for first graders and for second graders. The first graders showed a significant effect of condition ($F = 8.343$, $df = 160$, $p < .005$), but no significant interaction between condition and reading level. The interaction for the second graders, however, was significant ($F = 5.105$, $df = 1,48$, $p < .028$), indicating that the interference effect was significantly less for below-average readers than for above average readers in the second grade.

Discussion

It is clear from the data that our prediction that the interference effect would increase with the grade of the subject was not confirmed. Unlike Schiller's (1966) first graders, our youngest subjects did show a highly significant interference effect, confirming our suspicion that the words Schiller used exceeded the reading capacity of his subjects. In this respect, it is interesting that those of our subjects who could not read one of the critical words (BEEP) at the start of the experiment still experienced an interference effect which was statistically indistinguishable from their more advanced peers. Thus by the end of their first year of formal reading instruction these children seem to experience difficulty in ignoring the meaning of words which they can read.

A theory which claims that the meaning of the printed word only gradually becomes compelling could still be maintained if it could be shown that children at earlier stages of reading than those we tested were not disturbed by interfering words. For example kindergarteners or beginning first graders could be given a simpler test consisting only of the colors red, green and the words STOP, GO. Alternatively, a two-switch device could be manipulated in accordance with two contrasting colors and the words ON, OFF OR YES, NO. We considered using such a binary task, but abandoned the idea on the advice of Lee Brooks (personal communication), who claims that Stroop effects are difficult to achieve with less than three colors.

A puzzling aspect of our data is the appearance of a significant interaction between reading level and condition in the second grade sample, but not in the first. Our suspicion is that the below-average readers in the second grade were more atypical than those in the first grade. The first grade sample divided itself quite evenly into two reading groups. In the second grade, the majority of the subjects were able to read all the words, so there was no way of dividing them further. In order to locate a group of less advanced readers in a short period of time, the classroom teachers were consulted about possible candidates. The fact that the teachers were quite successful in sending us the appropriate children attests to the validity of our reading test. But the important point is that these children were probably considerably more below-average than the first grade subjects classified by us as such. Perhaps the immediacy of semantic pick-up is one aspect of the reading process which has not developed properly in children with true reading problems. It may be that these children labor so hard over the decoding process that the meaning is processed late or not at all.

Conclusion

We have been concerned with the child's sensitivity to the semantic content of a set of isolated words, presented in a setting which highlights their meaning. For a subject who is set to perform one of three possible acts--make the car go, beep the horn, or stop all activity--the message conveyed by the appearance of the word BEEP is all too obvious and irresistible. One does not have to be very good at extracting meaning from written words in order to be confused by this task. In the more typical event (if one can speak about typical events in reading), meaning comes not from isolated words, but from whole sentences and paragraphs. Although first graders seem to be as susceptible as older children to the idea contained in an isolated word, we cannot generalize to situations in which meaning has to be extracted from larger units of print.

References

Comalli, P., Wapner, S., & Werner, M. Interference effects of Stroop color-word test in childhood, adulthood, and aging. J. genet. Psychol., 1962, 100, 47-53.

Gibson, E. J. Perceptual learning and the theory of word perception. In press, Journal of Cognitive Psychology, 1971.

Klein, G. S. Semantic power measured through the interference of words with color naming. Amer. J. Psychol., 1964, 77, 576-588.

Ligon, E. M. A genetic study of color naming and word reading. Amer. J. Psychol., 1932, 44, 103-110.

Rand, G., Wapner, S., Werner, M., & McFarland, J. H. Age differences in performance on the Stroop color-word test. J. Pers., 1963, 31, 534-557.

Schiller, P. H. Developmental study of color-word interference. J. exp. Psychol., 1966, 72, 105-108.

A Developmental Study of the Ability to Perceive and Utilize Categorical Structure

Alida Spaans, Ph.D.
Cornell University, 1971

Abstract

The questions posed in this dissertation were whether elementary school children perceive and utilize structure of a kind that could facilitate performance of a new task; whether they would again utilize the structure in a different task; whether training facilitates the perception and subsequent utilization of structure; and whether type of material employed plays a role in pick-up and utilization of structure. To answer these questions, the perception and subsequent utilization of taxonomic structure in a search task followed by a free-recall task was studied. The categories and the instances employed were well known to the second, fourth, and sixth grade children who served as Ss.

Specifically, the study investigated, first, the initial quality of ordering which occurred upon introduction to the search task. Each child was presented with either (a) colored photographs of common objects, or (b) their corresponding printed verbal labels. He was instructed to place the items on a grid in front of him and to do so in a manner which would allow him to find any item as easily as possible. Secondly, the study investigated the improvement in quality of ordering of the items as a result of either of two relevant experiences. These consisted of (a) actual search for specific items (performance training), or (b) participation in a guided discussion about the items presented (perceptual-verbal study). For children in the performance training groups, the effect of the initial quality of ordering on time spent searching for items was determined by means of a correlation. Finally, each child was given an oral free-recall test in order to determine (a) the correlation between the quality of the last ordering on the grid and the degree of grouping (clustering) by taxonomic category in recall, (b) the correlation between the degree of grouping (clustering) by taxonomic category in recall and the number of items correctly recalled, and (c) the number of photos and words correctly recalled.

The results revealed that even second graders initially did order the items systematically, did benefit from the training, especially the perceptual-verbal study, and subsequently did cluster in recall. Therefore, the "production deficiency" hypothesis of Flavell, Beach, & Chinsky (1966) was not supported. Likewise, the positive correlations obtained indicated that younger children did utilize structure. Therefore, the "mediation deficiency" hypothesis of Flavell, Beach, & Chinsky (1966) was also not supported. Two of the correlations did not vary with age. The correlation between

clustering in recall and number of items correctly recalled did increase slightly with age. Thus, the effective utilization of structure may increase with age, at least for some tasks. However, the developmental improvement seemed to be a gradual process. There were no discontinuities in the data. Although with age, the quality of both categorical and alphabetic ordering improved, both types of training became more effective, items were located faster, and more items were recalled correctly, no support for a "stage" theory of development was found. While the distinctive features of objects appeared to be perceived faster with photographic than with verbal materials, as revealed by speed of ordering, the effect of type of material disappeared as soon as the meaning was perceived. It was concluded that children do increase with age in ability to perceive and utilize structure, but even second graders are capable of doing so.