Volume XII Number 1 Winter 1978

Visible Language

The research journal concerned with all that is involved in our being literate

ISSN 0022-2224

A special issue devoted to the interface of reading and listening.
Guest editor: Dr. Dominic W. Massaro.

Visible Language, Volume XII, Number 1, Winter 1978.
Published quarterly (Winter, Spring, Summer, and Autumn) by Visible Language, Box 1972 CMA, Cleveland, OH 44106.

Second-class postage paid at Cleveland, Ohio, and at additional mailing offices.

Merald E. Wrolstad, Ph.D., Editor and Publisher
P. O. Box 1972 CMA, Cleveland, OH 44106 USA

Dominic W. Massaro

A Stage Model of Reading and Listening

Visible Language, XII 1 (Winter 1978), pp. 3-26.
Author's address: Department of Psychology, University of Wisconsin, Madison, WI 53706.
0022-2224/78/0100-0003$02.00/0

Language processing is the abstraction of meaning from a physical signal such as a printed text or sequence of speech sounds. The goal of an information-processing model is to describe how language is processed, not simply what the reader or listener must know to understand language. Language processing is viewed as a sequence of internal processing stages or operations that occur between the language stimulus and meaning. The operations of a particular stage take time and transform the information in some way, making the transformed information available to the following stage of processing. In the present model the storage component describes the nature of the information at a particular stage of processing whereas the functional component describes the operations of a stage of processing. The information-processing model is used heuristically to incorporate data and theory from a variety of studies of language processing.

I. Introduction

Does it matter that I wrote this contribution rather than spoke it? Or does it make a difference that you are reading rather than listening? Or are you in fact, not only reading the article but simultaneously hearing it being spoken by the little homunculus in your head? Regardless of the modality of the input, this special journal issue does not offer convenient solutions to these and other important problems in communication. We will, however, present some recent research and theory on the psychological processes involved in listening to speech and reading printed text. Our goal is to stimulate your interest and involvement in our study.

One of the persistent questions about understanding language is whether the modality of input (what might be called the true surface structure) makes a difference. The first answer that comes to mind is why should it. The purpose of all language is to communicate and understand a message (or for some to camouflage a message). Language production and understanding processes must solve the same problems in both visible and audible language. Although reading and listening may have developed independently, the processes involved may still be analogous in the same way that analogous biological processes have developed in convergent evolution. It is commonly accepted that two organisms may develop similar solutions to the problem of survival even though they evoled independently of one another. As an example, the eye of the octopus (a cephalopod) and the eye of man (a mammal) function very similarly although they evolved completely independently of one another (Blakemore, 1977). Following this logic, the assumption that reading and listening can be viewed as similar processes does not necessitate an assumption of a common phylogenetic or ontogenetic evolution. Given that reading and listening solve the same problem, it is not unreasonable to assume that they are analogous rather than hierarchical solutions to the problem. Wrolstad (1976) presents a similar argument and supporting evidence.

Recent research on the processing of manual-visual languages such as American Sign Language (ASL) indicates that analogous solutions to language understanding extend beyond reading and listening. On every dimension that has been explored, remarkable parallels have been found between understanding signs and understanding speech. Lane, Boyes-Braem, and Bellugi (1976) found that perceptual confusions among signs can be described utilizing a distinctive feature system, analogous to systems developed for perceptual confusions in speech. There is also evidence that grammatical structure in sign language plays the same functional role that it does in spoken language (Tweney, Heiman, and Hoemann, 1977). These results support the claim that the processes of language understanding are relatively general and abstract—not tried uniquely to the input modality. The work on ASL encourages the belief that there are similar and analogous processes in all forms of language understanding.

Although it might seem reasonable to assume that understanding spoken and written language exploits similar or analogous comprehension processes and structures, the early stages of decoding the input should reveal some basic differences. This follows from the fact that modality-specific processes are necessary to transform the sound vibrations of speech and the light waves of print. Several other obvious differences come to mind. Spoken language comes in one ear and goes out the other, whereas the print remains available at the beck and call of a regressive eye movement or a flip of the page. It is true that some compulsive listener might record the message and capitalize on the rewind and play option for particularly difficult sections of a spoken message. In the information processing model presented here, however, we will draw similarities between even the earliest modality-specific stages of language processing. Returning to our argument of convergent evolution, it is not unreasonable that the same or analogous processes are exploited for decoding spoken and written language.

II. Information-Processing Model

Reading and listening can be defined as the abstraction of meaning from printed text and from speech, respectively. To derive or arrive at meaning from a spoken or written message requires a series of transformations of the energy signal arriving at the appropriate receptors. Language processing can be studied as a sequence of processing stages or operations that occur between the energy stimulus and meaning. In this framework, language processing can be understood only to the extent that each of these processing stages is described. In a previous effort an information-processing model was utilized for a theoretical analysis of speech perception, reading, and psycholinguistics (Massaro, 1975b). The model was used heuristically to incorporate data and theory from a variety of approaches to the study of language processing. The model should be conceptualized as an organizational structure for the state of the art in language processing. In this paper I will present a general overview of the information-processing model, and use the model to describe and incorporate some recent research.

Figure 1 presents a flow diagram of the temporal course of reading and listening. At each stage the system contains storage and functional components. The storage component represents the information available at a particular stage of processing. The functional component specifies the procedures and processes that operate on the information held in the corresponding storage component. The model distinguishes four functional components: feature detection, primary recognition, secondary recognition, and rehearsal-recoding. The corresponding storage component represents the information available to each of these stages of processing.

III. Feature Detection and Primary Recognition

The feature detection process transforms the energy pattern created by the language stimulus and transduced by the appropriate receptor system into a set of features held in preperceptual storage. Primary recognition evaluates and integrates these features into a percept which is held in synthesized memory. In speech, for example, the changes in sound pressure set the eardrums in motion and these mechanical vibrations are transduced into a set of neural impulses. It is assumed that the signal in the form of continuous changes in vibration pattern is transformed into a set of relatively discrete features. Features do not have to be relatively primitive such as the amount of energy in a particular frequency band, but they may include information about the direction and rate of frequency change. It would be possible, for example, to have a feature detector that responds to the rising first formant transition that is characteristic of the class of voiced stop consonants.

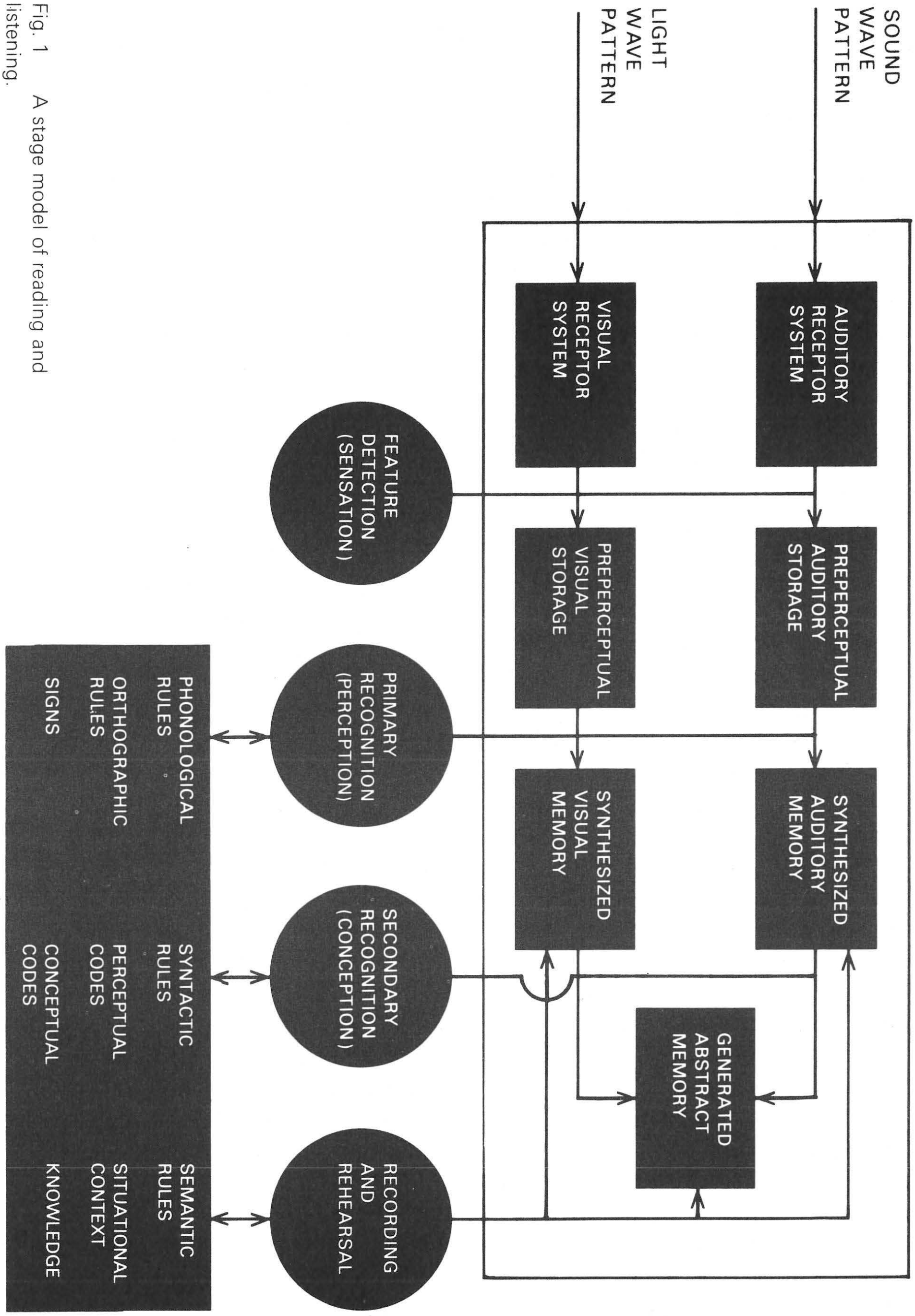


Fig. 1 A stage model of reading and listening.

A. Audible features

One traditional concern in speech research has been to determine the acoustic features that are utilized in perception. In terms of our model the feature detection process places features in a brief temporary storage called preperceptual auditory storage (PAS), which holds information from the feature detection process for about 250 msec. The primary recognition process integrates these features into a synthesized percept which is placed in synthesized auditory memory. One critical question is what features are utilized and a second important question is how are all of the features integrated together. Does the listener only process the least ambiguous feature and ignore all others, or are the features given equal weight, and so on? Despite the overwhelming amount of research on acoustic features, very little is known about how the listener puts together the multitude of acoustic features in the signal in order to arrive at a synthesized percept.

The integration of acoustic features has not been extensively studied for two apparent reasons. The first is that research in this area was highly influenced by linguistic descriptions of speech sounds in terms of binary all-or-none distinctive features (Jakobson, Fant, & Halle, 1961). One of the goals of distinctive feature theory was to describe all of the functional differences among speech sounds by a minimal number of distinctive features of the language. Therefore, distinctive features were designed to be general: if a distinctive-feature difference distinguished two phonemes in the language, that same distinction would also distinguish several other phoneme pairs. Given the distinctive feature of voicing, for example, the distinction of voiced versus voiceless can account for the differences between /z/ and /s/, /v/ and /f/, /b/ and /p/, and so on. The integration of information from two or more binary dimensions is a trivial problem. Integrating binary features from voicing and place of articulation, for example, could be carried out by simple logical conjunction. If the consonant /b/ were represented as voiced and labial and /p/ were represented as voiceless and labial, the identification of voiced labial sound would be /b/ whereas the identification of a voiceless labial sound would be /p/.

A second reason for the neglect of the integration problem is methodological. The primary method of study involved experiments in which the speech sound was varied along a single relevant dimension. In a typical study of voicing all voicing cues were made neutral except one, such as voice onset time and then this dimension was varied through the relevant values. Similarly, place of articulation was studied by neutralizing all cues but one, and then varying the remaining dimension through the appropriate values. Very few experiments independently varied both voicing cues and place cues within a particular experiment so that little information was available about how these cues were integrated into a synthesized percept.

More recently, we have initiated a series of experiments that are aimed more directly at the study of the integration of acoustic features in speech perception (Massaro & Cohen, 1976; Oden & Massaro, 1977). In contrast to the traditional linguistic description, we assume that the acoustic features held in preperceptual auditory storage (PAS) are continuous, so that a feature indicates the degree to which the quality is present in the speech sound. Rather than assuming that a feature is present or absent in PAS, it is necessary to describe a feature as a function of its degree of presence in PAS. This assumption is similar to Chomsky and Halle's (1968) distinction between the classificatory and phonetic function of distinctive features. The features are assumed to be binary in their classificatory function, but not in their phonetic or descriptive function. In the latter, features are multivalued representations that describe aspects of the speech sounds in the perceptual representation. Similarly, Ladefoged (1975) has also distinguished between the phonetic and phonemic level of feature description. A feature describing the phonetic quality of a sound has a value along a continuous scale whereas a feature classifying the phonemic composition is given a discrete value. In our framework the continuous features in PAS are transformed into discrete percepts in synthesized auditory memory (SAM) by the primary recognition process.

Given this theoretical description, acoustic features in PAS must be expressed as continuous values. That is to say, the listener will be able to hear the degree of presence or absence of a particular feature, even though his judgment in a forced choice task will be discrete. Oden and Massaro (1977) have used this description to describe acoustic features as fuzzy; that is to say, varying continuously from one speech sound to another. In this representation features are represented as fuzzy predicates which may be more or less true rather than only absolutely true or false (Zadeh, 1971). In terms of the model, fuzzy predicates represent the feature detection and evaluation process; each predicate is applied to the speech sound and specifies the degree to which it is true that the sound has a relevant acoustic feature. For example, rather than assuming that a sound is voiced or voiceless, the voicing feature of a sound is expressed as a fuzzy predicate.

$$P(\text{voiced}(S_{ij})) = .65 \qquad (1)$$

The predicate given by Equation 1 represents the fact that it is .65 true that speech sound S_{ij} is perceived to be voiced. In terms of our model, then, the feature detection process makes available a set of fuzzy predicates at the level of PAS. In addition to being concerned with the acoustic features in preperceptual storage this analysis of the feature evaluation process makes apparent that an important question in speech perception research is how the various continuous features are integrated into a synthesized percept.

As an example of the study of acoustic features, consider the dimension of voicing of speech sounds. In English the stops, fricatives, and affricates can be grouped into cognate pairs that have the same place and manner of articulation but contrast in voicing. The question of interest is what acoustic features are responsible for this distinction and how the various features are integrated together in order to provide the perceptual distinction. The integration question has not been extensively studied, however, since the common procedure in these experiments is to study just a single acoustic feature at a time. Consider two possible cues to the voicing distinction in stop consonant syllables: voice onset time (VOT), the time between the onset of the syllable and the onset of vocal cord vibration, and the fundamental frequency (F_O) of vocal cord vibration at its onset. Each of these cues has been shown to be functional in psychophysical experiments when all other cues have been held constant at neutral values. However, it is difficult to generalize these results to the perception of real speech, since no information is provided about the weight that these features will carry when other features are also present in the signal. To overcome this problem it is necessary to independently vary two or more acoustic features in the signal. The results of this type of experiment not only provide information about the cue value of one feature when other features are present in the signal, but also allow the investigator to evaluate how the various acoustic features are combined into an integrated percept. (For a further discussion see Massaro & Cohen, 1976, 1977; Oden & Massaro, 1977).

B. Audible features in fluent speech

The success of finding acoustic features in perception of isolated speech sounds might lead one to expect that perception of fluent speech is a straightforward process. Sound segments could be recognized on the basis of their features and the successive segments could be combined into higher-order units of words, phrases, and sentences. However, the acoustic structure of words in fluent speech differ significantly from the same words spoken in isolation. Two sources contribute to the large variation of words in fluent speech: coarticulation and psychological parsimony (Cole, & Jakimik, 1977; Ross, 1975).

In fluent speech the speech articulators must assume an ordered series of postures corresponding to the intended sounds, and the articulators cannot always reach their intended targets because of the influence of adjacent movements. Coarticulation refers to altering the articulation of one sound because of neighboring sounds. The words *did* and *you* spoken as /dId/ and /ju/ in isolation will be articulated as /dIdʒ u/ in combination because of palatalization. The alveolar stop followed by a front glide when combined produce the front-palatal affricate /dʒ /, even though a word boundary intervenes. Psychological show that his grouping of letters in terms of their legibility could predict performance on multi-letter items that are not words. If the uniqueness of words is responsible for Cosky's negative finding, then positive results should occur in nonword strings. Until this is demonstrated, Cosky's results can only be taken as a failure to find proof for the letter-mediation model; it cannot be taken as disproof.

In a well-known experiment carried out by Reicher (1969), subjects presented with either a single letter, a four-letter word, or a four-letter nonword flashed in a tachistoscope had to report what they saw. Reicher's contribution to this century-old task was to constrain the subject's choice by presenting two letter alternatives after each trial. Both alternatives would complete the display spelling words in the word condition so that performance on the word trials would not benefit from a simple guessing strategy. Even with these constraints, Reicher found a 10% advantage for recognition of a letter in a word over recognition of a letter in a nonword or a letter presented alone. These results have been described both in terms of whole-word and single-letter perceptual units. In terms of a perceptual unit the size of a word, it has supraletter features such as overall word shape which facilitate direct contact with the appropriate memory representation. Words are recognized better than letters or nonwords because the unique visual features of a word allow for easier recognition than the features of a single letter or a nonword.

The advantage of words over single letters and nonwords is not incompatible with the idea that the letter is a basic perceptual unit, however (Massaro, 1975b). In the present model the primary recognition process operates on a number of letters in parallel. The visual features read out at each letter position define a candidate set of possible letters for that position. The recognition process is not limited to featural information, but can also utilize knowledge about the orthographic structure of English spelling. The letter that is synthesized at each position, therefore, will not only correspond to the visual information that is available from feature detection and evaluation, but will also correspond to the orthographic constraints in the language. For example, consider the case in which the subject is given the lowercase string *coig* and has resolved just the circular envelope of the first letter and all of the last three letters. Given that *c, e,* and *o* are the only letters that are consistent with the circular envelope, these are the only possible letters at this position. If the reader further assumes that the string must conform to English orthography, only *c* is possible since the strings *ooig* and *eiog* are illegal English spellings. In this case the reader can synthesize *coig* since it is the only valid alternative. When the single letter *c* is presented, on the other hand, the perception of the envelope does not allow an unambiguous choice among *c, e,* and *o.* Accordingly, the reader is less likely to synthesize the correct alternative and will be correct only one out of three times. Although a word is recognized via its component letters, familiarity with the orthographic structure of words facilitates primary recognition of its letters relative to a single-letter or nonword.

IV. Secondary Recognition

Secondary recognition transforms synthesized percepts into meaningful forms in generated abstract memory. In speech perception it is assumed that the input is analyzed syllable by syllable for meaning. In reading letter sequences are closed off in word units. In both cases the secondary recognition process makes the transformation from percept to meaning by finding the best match between the perceptual information and the lexicon in long-term memory. Each word in the lexicon contains perceptual and conceptual codes. The concept recognized is a function of at least two independent sources of information: the perceptual information in synthesized memory and the semantic/syntactic context in the message.

A. Perceptual and contextual contributions to listening

Our conceptualization of speech processing is one that is perceptually, and, therefore, acoustically driven. We assume that the secondary recognition process operates syllable by syllable on the output of primary recognition. However, contextual constraints also exert a strong influence at this stage of processing, so that both contributions must be accounted for in describing how meaning is imposed on the spoken message. A series of recent studies has shown that abstracting meaning is a joint function of the perceptual and contextual information. In one experiment Cole (1973) asked subjects to push a button every time they heard a mispronunciation in a spoken rendering of Lewis Carroll's *Through the Looking Glass.* A mispronunciation involved changing a phoneme by 1, 2, or 4 distinctive features (for example, *confusion* mispronounced as *gunfusion, bunfusion,* and *sunfusion,* respectively). The probability of recognizing a one-feature mispronunciation was .3 whereas a four-feature change was recognized with probability .75. This result makes apparent the contribution of the perceptual information passed on by the primary recognition process. In our view some of the mispronunciations went unnoticed because the contribution of contextual information worked against the recognition of a mispronunciation. The syntactic/semantic context of the story would support a correct rendering of the mispronounced word, outweighing the perceptual information. In support of this idea all mispronunciations were correctly recognized when the syllables were isolated and removed from the passage.

Cole and Jakimik (1977) reasoned that the listener should be faster at detecting a mispronunciation to the extent that a word is predicted by its preceding context. This follows from the idea that the quickest way to detect a mispro-

Fig. 2 Predicted and observed percentage of fluent restorations as a function of the amount of feature change, the syllable, and the contextual constraint of the mispronunciation (observed data from Marslen-Wilson and Welsh, in press).

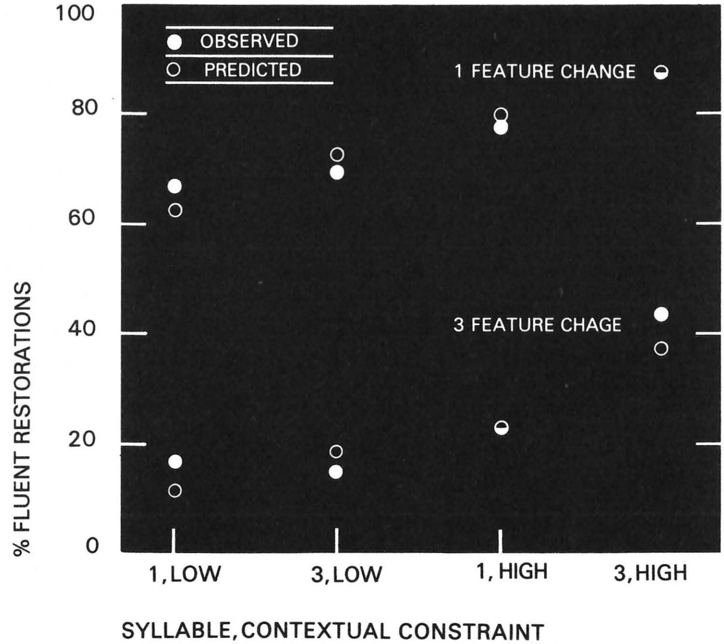
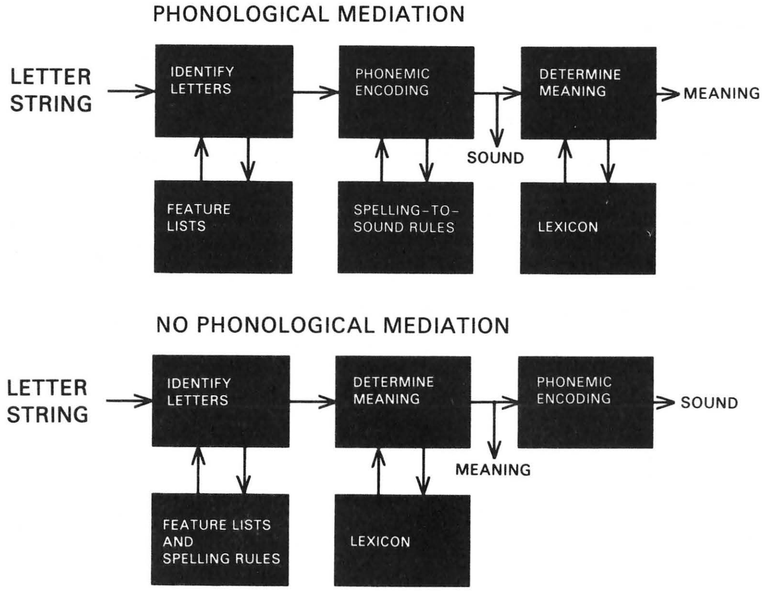


Fig. 3 Two stage models of the role of phonological mediation in reading.



the outcomes of these analyses were combined with the higher-order constraints. The fact that higher-order constraints in the passage influence shadowing does not mean that some analyses do not begin before others. More importantly, their view might be interpreted to mean that higher-order analyses modify the output of lower-level analyses. However, a quantitative model that assumes that both levels of analyses are functionally independent can accurately describe the results of their experiments. Figure 2 also presents the predictions of a quantitative formulation of the independence model (see Massaro, 1977, for the exact form of the model). The model assumes that the information passed on by the feature detection and evaluation process is equivalent regardless of the higher order constraints in the message. Therefore, it is not necessary to assume that higher-order constraints allow the subject to selectively attend to or selectively process certain acoustic properties of the speech input. In this model higher-order constraints do not modify the nature of low-level perceptual analyses performed on the input data.

B. Phonological mediation in reading

A persistent question in reading-related research is the extent to which the reader translates print into some kind of speech code before meaning is accessed. A similar but not identical question is the extent to which the speech code is *necessary* for the derivation of meaning. Figure 3 presents two extreme answers to the phonological mediation question. In the first model letters are identified and mapped into a speech code using spelling-to-sound rules, and meaning is determined on the basis of the derived speech code. In the second model meaning is determined from the letter resolution, and a speech code is not made available until after meaning has been accessed.

Gough and Cosky (1976) believe that they have accumulated some new data in support of the phonological-mediation view of Gough (1972). Subjects were asked to read aloud as rapidly as possible words that violated or obeyed spelling-to-sound rules. If phonological mediation occurs, regular words which conform to spelling-to-sound rules should be converted to a speech code faster than exception words which violate the rules. Accordingly, the time to comprehend the word and name it aloud should take longer for the words that violate spelling-to-sound rules. In support of their hypothesis, the pronunciation times for exception words averaged 27 msec longer than the pronunciation times for regular words. However, there is no assurance that differences in pronunciation time result from differences in word recognition time. The differences in reaction time could also have resulted from differences in the time for response selection and programming after the word had already been identified (see Massaro, 1975b, p. 262).

but not pseudowords. Frederiksen may have found much larger differences than Theios because blocking the words in a session would encourage the subjects to pronounce the words via lexical access. Randomizing words and pseudowords in the Theios and Muise study might have encouraged pronouncing some of the words by way of spelling-to-sound rules rather than by way of lexical access. In agreement with this interpretation, Frederiksen and Kroll (1976) found a larger effect of word frequency on naming RTs when only words were presented in a block of trials relative to a random mixture of word and pseudoword trials.

Green and Shallice (1976) asked subjects to judge whether two words rhymed or whether they belonged to the same broad semantic category. Mispelling the words as homophones produced a much larger decrement in the semantic than the rhyming task. If lexical access occurs via phonological coding, there is no reason that the semantic task should have been slowed more by misspelling than the rhyming task was. The fact that the rhyming task was performed about twice as fast as the semantic task shows that lexical access was not necessary in the former task although it was in the latter. Spelling-to-sound rules would have been sufficient to perform the rhyming task, and misspelling should have very little effect on this process. In support of this, misspelling the words increased reaction times by only 11 percent. Lexical access should be drastically influenced by misspelling however, if it occurs via a visual code. Reaction times were slowed by 58 percent in the semantic task, arguing against the idea of phonological or speech recoding in lexical access and derivation of meaning. The results support other negative findings on the necessity of phonemic encoding in processing written language for meaning (see Massaro, 1975a).

V. Rehearsal and Recoding

In the present model, the same abstract structure stores the meaning of both listening and reading. Generated abstract memory (GAM) in our memory corresponds to the working memory of contemporary information processing theory. Rehearsal and recoding processes operate at this stage to maintain and build semantic/syntactic structures. There is good evidence that this memory has a limited capacity, holding about 5 ± 2 chunks of information. For a more detailed discussion of processing at this stage, see Massaro (1975a, Chapter 27).

Although GAM is assumed to be abstract relative to SAM and SVM, the nature of the information appears to be tied to the surface structure of the language rather than in terms of underlying meaning that is language independent. Some relevant research comes from work experiments carried out with bilingual subjects (Dornic, 1975, provides an excellent review). Recall from immediate memory (supposedly tapping GAM) does not differ for unilingual and bilingual lists, whereas recall of items assumed to be no longer in GAM is poorer in bilingual than unilingual lists (Tulving & Colotla, 1970). Similarly, Kintsch and Kintsch (1969) showed that the semantic relationship between the words in different languages did not influence immediate memory, but did affect recall of items no longer active in GAM. Saegert, Hamayan, and Ahmar (1975) showed that multilingual subjects remembered the specific language of words in a mixed language list of unrelated words, but this information was forgotten when the words were presented in sentence contexts. Dornic (1975) points out that surface structure and item information are integrally related in immediate memory; subjects seldom report translations for the words. If the items are remembered, so are the appropriate surface structure forms.

In our model, GAM has a "limited capacity" and the learning and memory for information is a direct function of rehearsal and recoding processes. Memory of an item will increase with the time spent operating on that item, and will decrease with the time spent operating on other "unrelated" items. This "limited capacity" rule has provided a reasonable description of the acquisition and forgetting of information in GAM (cf. Massaro, 1975a, Chapter 27). A critical question for the recoding operation centers around the size of the units that are recoded. It seems unlikely that recoding occurs word by word given that many words are ambiguous until later context disambiguates their meaning.

VI. Conclusion

It seems valuable to attack reading and listening with similar methodological and theoretical forces in the framework of an information-processing model. Our concern is with *how* the reader and listener perform, and with the dynamics of this performance. Although the surface structure of written text and speech present questions unique to each skill, the apparent similarities in deep structure offer the hope of a single framework for understanding both reading and listening.

VII. Preview of Contributions on Reading and Listening

One reason that speech has been considered primary and reading and writing secondary is the supposedly uniqueness of certain speech perception phenomena. At the top of the list has been the categorical perception of speech sounds. Categorical perception refers to a basic perceptual limitation in the perception of speech sounds. Certain sounds cannot be discriminated from one another unless they are, in fact, categorized differently. For example, the two sounds /ba/ and /pa/ can be synthesized electronically so that they differ only along a single dimension called voice onset time (VOT, the time between the onset of the stop release and the onset of vocal cord vibration in real speech). If two of the sounds differ by a VOT of 10 msec, they will not be discriminated from each other if

References

Blakemore, C.
The baffled brain. In R. L. Gregory and E. H. Gombrich (Eds.), *Illusion in nature and art,*
London: Duckworth, 1973.

Blakemore, C.
Mechanics of the mind.
New York: Cambridge University Press, 1977.

Blesser, B.; Shillman, R.; Kuklinski, T.; Cox, C.; Eden, M., & Ventura, J.
A theoretical approach for character recognition based on phenomenalogical attributes.
International Journal of Man-Machine Studies,
1974, 6, 701-714.

Cattell, J. M.
The time it takes to think.
Popular Science Monthly, 1888, 32, 488-491. Reprinted in James McKeen Cattell, *Man of science* Vol. 2, addresses and formal papers, Lancaster, PA: The science press, 1947.

Chomsky, N., & Halle, M.
The sound pattern of English.
New York: Harper & Row, 1968.

Cole, R. A.
Listening for mispronunciations: a measure of what we hear during speech.
Perception & Psychophysics,
1973, 13, 153-156.

Cole, R. A. & Jakimik, J.
Understanding speech: how words are heard.
Technical report,
Department of Psychology, Carnegie-Mellon University, 1977.

Cosky, M. J.
The role of letter recognition in word recognition.
Memory & Cognition,
1976, 4, 207-214.

Dornic, S.
Human information processing and bilingualism.
Report from the Institute of Applied Psychology, the University of Stockholm, No. 67, 1975.

Eriksen, C. W.; Pollack, M.D., and Montague, W. E.
Implicit speech: mechanism in perceptual encoding.
Journal of Experiment Psychology,
1970, 84, 502-507.

Frederiksen, J. R.
Decoding skills and lexical retrieval. Paper presented at Psychonomic Society, St. Louis, November, 1976.

Frederiksen, J. R., and Kroll, J. F.
Spelling and sound: approaches to the internal lexicon.
Journal of Experimental Psychology: Human Perception and Performance,
1976, 2, 361-379.

Gough, P. B.
One second of reading. In J. F. Kavanagh & I. G. Mattingly (Eds.)
Language by ear and by eye,
Cambridge, Mass., MIT Press, 1972.

Gough, P. B., & Cosky, M. J.
One second of reading again. In N. J. Castillan, Jr.; D. B. Pisoni, and G. R. Potts, (Eds.).
Cognitive Theory, Vol. 2, Hillsdale, N.J., Erlbaum, 1976.

Green, D. W., & Shallice, T.
Direct visual access in reading for meaning.
Memory and Cognition,
1976, 4, 753-758.

Hubel, D. N., & Wiesel, T. N.
Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex.
Journal of Physiology,
1962, 160, 106-154.

Huey, E. B.
The psychology and pedagogy of reading.
New York: MacMillan, 1908 (Reprinted Cambridge: MIT press, 1968).

Jakobson, R.; Fant, C. G. M., & Halle, M.
Preliminaries to speech analysis: the distinctive features and their correlates.
Cambridge, Mass.: MIT Press, 1961.

Keyser, S. J., & Halle, M.
What we do when we speak. In P. A. Kolers & M. Eden (Eds.)
Recognizing Patterns.
Cambridge, Mass.: MIT Press, 1968.

Kintsch, W., & Kintsch, E.
Interlingual interference and memory processes.
Journal of Verbal Learning and Verbal Behavior,
1969, 8, 16-19.

Klapp, S. T.
Syllable-dependent pronunciation latencies in number naming: a replication.
Journal of Experimental Psychology.
1974, 102, 1138-1140.

Klapp, S. T.; Anderson, W. G., & Berrian, R. W.
Implicit speech in reading, reconsidered.
Journal of Experimental Psychology,
1973, 100, 368-374.

Ladefoged, P.
A course in phonetics.
New York: Harcourt, Brace & Jovanovich, 1975.

Lane, H.; Boyes-Braem, P., & Bellugi, U.
Preliminaries to a distinctive feature analyses of handshapes in American Sign Language.
Cognitive Psychology,
1976, 8, 263-289.

Lieberman, P.
Some effects of semantic and grammatical context on the production and perception of speech,
Language and Speech,
1968, 6, 172-187.

Lieberman, P.
Intonation, perception, and language.
Cambridge, Mass.: MIT Press, 1967.

Lindsay, P. H., & Norman, D. A.
Human information processing.
New York: Academic Press, 1977.

Marslen-Wilson, W. D.
Linguistic structure and speech shadowing at very short latencies.
Nature,
1973, 244, 522-523.

Marslen-Wilson, W. D.
Sentence perception as an interactive parallel process.
Science,
1975, 189, 226-228.

Marslen-Wilson, W. D., & Welsh, A.
Processing interactions and lexical access during word recognition in continuous speech.
Cognitive Psychology,
1978, 10, 29-63.

Massaro, D. W.
Experimental psychology and information processing.
Chicago: Rand-McNally, 1975(a).

Massaro, D. W.
Understanding language: an information-processing model of speech perception, reading, and psycholinguistics.
New York: Academic Press, 1975(b).

Massaro, D. W.
Reading and listening.
Technical Report No. 423.,
Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, Madison, Wisconsin, 1977.

Massaro, D. W., & Cohen, M. M.
The contribution of fundamental frequency and voice onset time to the /zi/-/si/ distinction.
Journal of the Acoustical Society of America,
1976, 60, 704-717.

Massaro, D. W., & Cohen, M. M.
Voice onset time and fundamental frequency as cues to the /zi/-/si/ distinction.
Perception and Psychophysics,
1977, 22, 373-382.

Mayzner, M. S., & Tresselt, M. E.
Tables of single-letter and diagram frequency counts for various word-length and letter-position combinations.
Psychonomic Monograph Supplements,
1965, *1*, 13-32.

Naus, M. S., & Shillman, R. J.
Why a Y is not a V: a new look at the distinctive features of letters.
Journal of Experimental Psychology: Human Perception and Performance,
1976, 2, 394-400.

Oden, G. C., & Massaro, D. W.
Integration of place and voicing information in identifying synthetic stop-consonant syllables.
WHIPP Report #1,
Wisconsin Human Information Processing Program, July, 1977.
[Also *Psychological Review,* 1978, 85, 172-191.]

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

Rosenberg, S., & Lambert, W. E.
Contextual constraints and perception of speech.
Journal of Experimental Psychology,
1974, 102, 178-180.

Ross, J. R.
Parallels in phonological and semantic organization. In J. F. Kavanagh & J. E. Cutting (Eds.)
The role of speech in language,
Cambridge, Mass.: MIT Press, 1975.

Saegert, J.; Hamayan, E., & Ahmar, H.
Memory for language of input in polyglots.
Journal of Experimental Psychology: Human Learning and Memory,
1975, 1, 607-613.

Shillman, R. J.; Cox, C.; Kuklinski, T.; Ventura, J.; Blesser, B., & Eden, M.
A bibliography in character recognition. Techniques for describing characters.
Visible Language,
1974, 8, 151-166.

Spencer, H.
The visible word.
New York, Hastings, 1969.

Sternberg, S.
The discovery of processing stages: extensions of Donders' method.
Acta Psychologica,
1969, 30, 276-315.

Theios, J., & Muise, J. G.
The word identification process in reading. In N. J. Castellan, Jr.; D. B. Pisoni, & G. R. Potts (Eds.)
Cognitive Theory,
Vol. 2, Hillsdale, N.J., Erlbaum, 1976.

Treisman, A. M.
Verbal responses and contextual constraints in language.
Journal of Verbal Learning and Verbal Behavior,
1965, 4, 118-128.

Tulving, E., & Colotla, V. A.
Free recall of bilingual lists.
Cognitive Psychology,
1970, 1, 86-98.

Tweney, R. D.; Heiman, G. W., & Hoemann, H. W.
Psychological processing of sign language: effects of visual description on sign intelligibility.
Journal of Experimental Psychology: General,
1977, 106, 255-268.

Umeda, N.
Consonant duration in American English.
Journal of Acoustical Society of America,
1977, 61, 846-858.

Wrolstad, M. E.
A manifesto for visible language.
Visible Language, 1976, 10, 5-40.

Zadeh, L. A.
Quantitative fuzzy semantics.
Information Sciences,
1971, 3, 159-176.

Phonemes and Alphanumeric Characters: Possible Components of Parallel Human Communication Systems

Visible Language, XII 1 (Winter 1978), pp. 27-42.
Author's address: Department of Psychology, SUNY at Binghamton, Binghamton, NY 13901.
0022-2224/78/0100-0027$02.00/0

Alphanumeric characters and phonemes can be viewed as information codes used by human communication systems. If such communication systems were designed to be effective, then we should expect to find certain characteristics which should be manifested in the nature of the physical representations of the information codes and in the perception of these codes. These characteristics are discussed in terms of their importance to communication systems in general and their manifestations in human audible and visible language. When viewed from the perspective of such communication systems, we should expect to find many parallels in the perception of alphanumeric characters and phonemes. This paper examines some of these parallels, drawing upon our knowledge of human perception.

The purpose of this paper is to examine the perception of components of written and spoken codes of human communication from a broad perspective in an attempt to identify possible similarities. We will focus on the perception of phonemes and alphanumeric stimuli as equivalent levels of components or codes of an effective communication system, and intentionally will side-step the important, often discussed but poorly-answered question whether these components represent the basic units of language.

A. An Overview

Human language normally involves two distinct sets of physical representations of its information code. One of these physical codes involves patterns of sounds varying in time, frequency, intensity, and complexity. The sound patterns are produced by the articulatory system and are perceived by the auditory system. The other code involves spatial patterns of lines, is written, and is perceived in a temporal and spatial manner through the visual system. One level of units for the spoken language is the phoneme, whereas it is the letter for the written language (at least for Western cultures).

There are many well-documented differences between the auditory and visual systems:

differences between the physical nature of auditory and visual stimuli, differences in the nature of the sensory transducers for the stimuli (Geldard, 1972), and differences in the nature of auditory and visual perception. There also are similarities, especially similarities between elements of auditory and visual perception (Julesz & Hirsh, 1972). Thus we must expect to find differences, as well as some similarities between the components of the language codes. However, for the perspective of this paper it is more important that there also exist many parallels between the components of the written and spoken languages as codes for an effective communication system. Independent of stimulus modality, such a communication system is faced with problems associated with pattern recognition. This paper examines some of these parallels in the organization and perception of spoken and written language in the context of such communication systems. We will begin by examining the desirable characteristics of effective communication systems. After summarizing relevant aspects of phoneme perception, we will examine similarities between phoneme perception and the perception of alphanumeric stimuli, as well as the perception of other meaningful stimuli.

B. An Effective Communication System

Let us examine phonemes and alphanumeric stimuli as codes used by a minimum error (therefore effective) human communication system. All communication systems must include a set of transmission processes, a set of reception processes, and a set of physically defined stimuli which represent the code which exists between transmission and reception. The information code must be represented by the overlapping subsets of the potential outputs of the transmitter and all of the stimuli which can be perceived by the receiver.

A minimum error system is most viable when there exists several attributes of the communication code. First, there should be major differences in the manner in which the various components of the language code are produced. In addition, the characteristics of the stimuli which result from these major differences in production also must be perceptually distinct. In this way there will tend to be a high degree of correspondence between the act of producing a given language component and the language component actually perceived. Secondly, although language components may be distinguished from one another by their manner of production, the way in which any specific component is produced must be flexible. By encompassing a broad (rather than a highly restrictive) range of production modes, structural variability of different transmitters may be tolerated. In a human system, where learning and practice are essential for the proper execution of the production sequences by the transmitter (especially by the young and inexperienced), such flexibility is crucial. Thus, we should expect to see some variability within, but basic differences between, language components (e.g., between different phonemes or different letters). It also is highly desirable for differences in the code to be designed around stimulus attributes which are highly discriminable by the receiver. Then members of different code classes will be more discriminable than members of the same code class. The phenomenon called categorical perception defines precisely this relationship (see below). Given these general attributes of a minimum error system, we would expect a high correlation between production and reception of the transmission code.

The conception of an effective communication system is not inconsistent with the notion that a portion of the communication system is composed of (or is under the control of) specialized processes unique to language. Our conception simply is that the communication code involves only a subset of all production and reception processes and the stimuli relevant for each. It implies that the transmission processes could be capable of producing stimuli other than those associated with the communication code and that the reception system could be capable of perceiving stimuli other than those produced by the transmission processes.

Another desirable attribute of the perceptual aspects of an effective communication system is the use of what G. A. Miller calls "chunking" (Miller, 1956). The stimuli which comprise the information code should not be processed and remembered in terms of the individual multitude of stimulus components which, together, are the physical representation of the code. The amount of information which would have to be processed per unit time under such a communication structure would be enormous. This point has been made by many early Gestalt theorists for visual perception (e.g., Wertheimer, 1923) and, more recently, by phonetic theorists (e.g., Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). Rather, the stimulus information should be processed in terms of organized figurally meaningful groupings (e.g., chunked in terms of phonetic or alphanumeric units). This implies that stimuli should be analyzed in terms of molar units, and that the molecular structure of the stimuli should tend to be ignored (unless attention is specifically focused on such aspects).

This description of an effective communication system appears to be consistent with the results of research investigating the production and perception of phonemes. It also is consistent with the production and perception of alphanumeric stimuli, with the phonetic and alphanumeric language systems paralleling each other (and potentially interacting) in terms of the basic attributes defined above for such communication systems.

C. Phoneme Perception

The characteristics of English phonemes important to the definition of the language code are well documented (e.g., Fant, 1968; Stevens, 1971; Denes and Pinson,

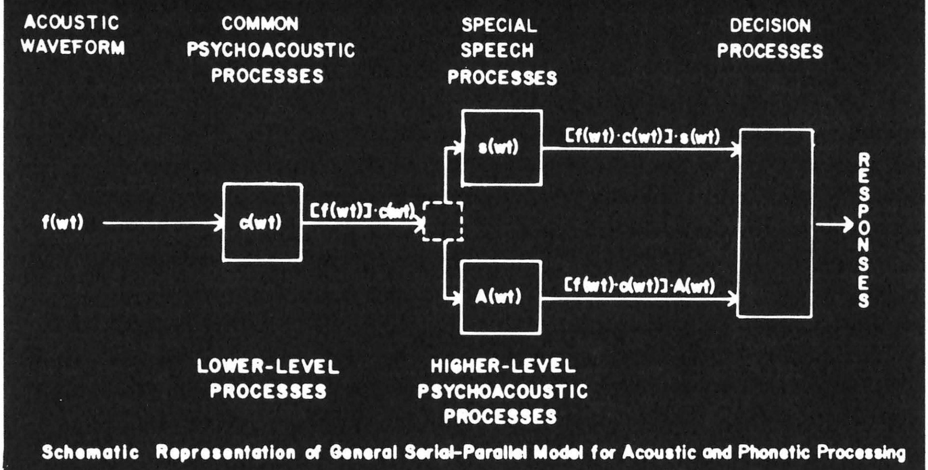
1973) and bear on this discussion only to the extent that they are relevant to the effective communication of information. More important to our comparison of audible and visible languages is the manner in which the perceptual processing of phonetic information occurs.

A simplified version of one general model commonly assumed to describe the relationship between the acoustic and the uniquely phonetic processing of the auditory code is diagrammed in Figure 1. The model, as diagrammed, is essentially the hybrid serial-to-parallel model described by Wood (1975). The initial acoustic waveform is an energy distribution which varies as a function (f) of both frequency (w) and time (t), and thus is described as f(wt). According to the model, f(wt) first undergoes a preliminary psychoacoustic level of processing.

In Figure 1 these preliminary processes have a transfer function described by c(wt), which represents every aspect of auditory stimulus processing common to all acoustic stimuli (whether or not phonetic). The resulting waveform is mathematically described in terms of the convolution of f(wt) and c(wt). This preliminary level of acoustic processing is assumed to involve primarily the simple transduction of the physical waveform of the stimulus into its neural code, and thus often is tacitly assumed to be relatively unimportant in determining the nature of auditory perception. Obviously, this is a simplification of the nature of preliminary acoustic processes which, at a minimum, are known to contain nonlinear components and to smear information in time and frequency. The neural representation which results from the action of these common psychoacoustic processes on the physical waveform is assumed to be fed into a simple decision mechanism (represented in terms of broken lines in Figure 1) which directs the waveform to either (but not both) of the two "higher-level" sets of processes (e.g., Studdert-Kennedy, Liberman, Harris, & Cooper, 1970). If the waveform is recognized as being phonetic in nature, it is fed into a set of special speech processes.

The higher level speech processes are assumed to interpret and classify the stimuli into phonetic categories, ignoring (and somehow permanently losing) information about the acoustic detail of the basic components of such stimuli. If the waveforms are not identified as being phonetic, they are passed through the psychoacoustic processes, then fed to decision processes which determine the appropriate responses for the given situation. There is a large body of literature which seems to generally support this hybrid serial-to-parallel model for acoustic and phonetic processing (e.g., Liberman & Studdert-Kennedy, 1977; Liberman & Pisoni, 1977). For instance, differences in the psychophysical functions describing the perception of phonetic and *selected samples* of nonphonetic, acoustic stimuli provide a major justification for some separation in higher-level acoustic and phonetic processes. An even stronger justification exists in the large body of literature which demonstrates a hemispheric special-

Fig. 1 Simple hybrid serial-parallel midel commonly assumed to describe the processing of phonetic and nonphonetic acoustic signals. Following an elementary analysis, the signals are subjected either to processes specialized for phoneme perception or to general, higher-level psychoacoustic processes (but not both).



ization in the cortical locus of control for the speech production and perception (although understanding the significance of these differences is dependent upon the demonstration of actual differences in perception of phonetic and nonphonetic stimuli). These differences would be defined in terms of the differences in the transfer function s(wt), for all uniquely phonetic processes, and a(wt), for all remaining processes *not* involved in phonetic perception.

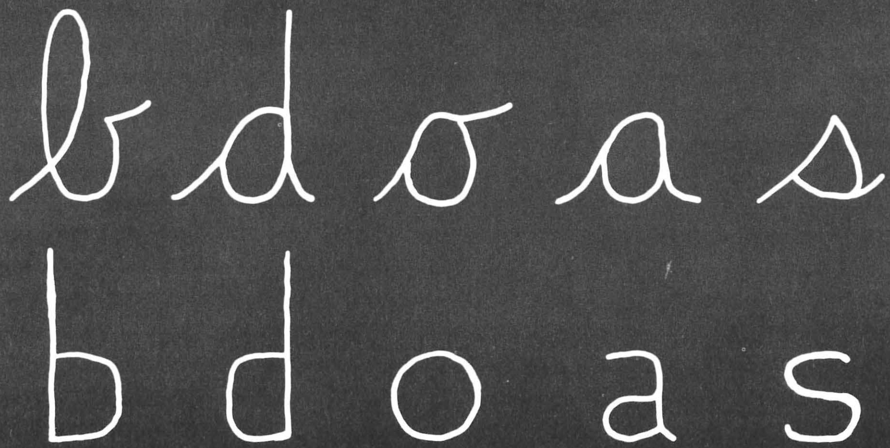
D. Comparisons of spoken and written language

1. *Hybrid Serial-Parallel Model for Written Language*

It can be argued that the general type of hybrid serial-parallel model developed for phoneme perception (see above) is equally applicable for the perception of visually presented letters. Research on stabilized retinal images provides one possible source of justification (e.g., Heckenmueller, 1965). The stabilization of the retinal image removes temporal variation in the proximal stimulus, thus maintaining a specific visual pattern on a given retinal location. Under such conditions, portions of the visual image tend to disappear differentially and selectively over time. For example, Pritchard (1961) found that when the stimulus was a random array of dots or lines, the stimulus components tend to be maintained in unpredictable but confined sections of the retinal field, and tend to lack any meaning or coherence (which was not present to begin with). However, when the stimulus was structured, such as the stimulus "HB " subjects reportedly saw residual stimuli, based on past experience, which were meaningful (e.g., H, B, 3, and 4). Such selective maintenance of stimulus organization must be a consequence of a higher level of organizational analysis probably based upon previous learning. Otherwise we would expect the fading of the visual field to be random. While these findings may imply that visual perception is maintained, or organized, on the basis of the meaningfulness of the stimuli, a more restrictive interpretation (which parallels the phonetic model) remains tenable. Such a hypothesis asserts that after identifying the visual stimuli as being alphanumeric, these stimuli are analyzed and recoded by a set of specialized higher level language processes (e.g., see Figure 1). The original "molecular" attributes (and missing portions) of the stimuli then will be lost (or ignored), with the stimuli being reorganized and completed as meaningful alphanumeric characters. The restructured stimuli then are perceived as complete members of the language code. Whatever the nature of the actual perceptual processes, this research does indicate that the perception of such stimuli (like the perception of phonemes) appear to be at least partially governed by prior knowledge, rather than simply being built up from the physical stimulus properties.

A second and probably stronger justification for an analogous special alphanumeric processor can be found in the principles of perceptual organization found in many modern (and not so modern) alphanumeric displays. With limited sets of unconnected dots or lines, one can display patterns of stimuli which are readily perceived not as individual lines or dots, but as complete letters or numbers. Could such perception of these "incomplete" patterns as "closed" alphanumeric figures be due to the activity of higher level language processes, such as described in Figure 1? In the Gestalt literature we can find further justification for this notion. If one views an apparent random pattern of stimuli, one sees a mosaic of components. However, once a pattern approximating a letter (or number, or face, etc.) is perceived, that pattern will tend to stand out as a "figure," with an identity separate from the background supplied by the remainder of the stimulus array (e.g., see Beardslee & Wertheimer, 1958). That figure now will tolerate a good deal of distortion, yet will maintain its apparent identity and will tend toward being perceived as complete (i.e., closure). Thus, we have the potential justification for hypothesizing the existence of a set of higher level alphanumeric processes analogous to the higher level phonetic processes outlined in Figure 1. Furthermore, the apparently similar perceptual organization of "other meaningful" visual stimuli can be assumed to be processed in a different but related manner which is analogous to the similarities and differences in the perception of phonetic and musical stimuli. Instead of following the logic, often used by phonetic theorists that there exist many different specialized systems, one could assume that we are describing attributes of a general perceptual system.

Fig. 2 Examples of cursive (upper row) and printed letters (lower row) demonstrating that the shapes of the patterns are more similar within types of production (i.e., rows) than within types of letters (columns).



2. Perceptual Invariance

Claims of the uniqueness of the spoken language code has tended to limit generalization of important aspects of the perception of the phonetic code to alpha-numeric code and to perception in general. For instance, the physical representation of a single consonant (e.g., /d/) may be a rising change in frequency in the context of one vowel and a falling change in frequency in the context of a different vowel (Liberman, et al., 1967). Thus, we appear to have a perceptual invariance in the context of gross physical differences between stimuli. At a recent symposium on language by ear and eye, it was claimed that, "There is nothing in reading corresponding to the lack of invariance between the speech signal and the underlying phonetic segments . . . " (Kavanagh & Mattingly, 1972, p. 53). However, a close examination of letters used in the English language reveals clear examples of a lack of apparent structural invariance in the definition of certain letters. Furthermore, this lack of invariance in the definition of certain letters is probably related to the production mechanism in an analogous manner that the lack of invariance in phonemes is related to their articulation.

Some examples are summarized in Figure 2. Printed lower case examples of the letters "s" and "a" are similar in structure and are difficult to produce when writing script. In developing an efficiently produced cursive alphabet, the cursive code for the letter "s" becomes more similar to the cursive code for the letter "a" than to the printed code for the letter "s." Likewise, the cursive code for the letter "a" is more similar to the cursive codes for the letters "d" and "s" than it is to the printed code for the letter "a" (see Figure 2). Examining the remainder of the cursive and printed alphabets, one can find a number of other examples in which there is an apparent lack of invariance within the code for specific letters. However, through experience we have learned to recognize and readily perceive different physical codes as being equivalent representations of the same letter. This finding is not surprising, since it is well known that given proximal stimulus patterns often can be produced by an infinite number of different distal stimuli, and a single distal stimulus can produce a large number of proximal stimuli (Hochberg, 1964). One of the basic premises of modern perceptual theory is the existence of transformations of the stimulus input to accomplish perceptual constancy in the context of considerable physical dissimilarity in the proximal stimuli (Gibson, 1966).

These observations lead us to question the uniqueness of an apparent perceptual invariance for phonetic code and to conjecture that the perceptual invariance we have described for physically dissimilar alphanumeric stimuli might involve an equivalent type of perceptual reorganization of physical stimulus information. This possible similarity leads us to look for other possible parallels in the structure of written and spoken codes for language. We believe that there is such a parallel in the phenomenon known as categorical perception.

Categorical Perception

Categorical perception has been claimed to be unique to phoneme perception (e.g., Liberman, Harris, Hoffman, and Griffith, 1957; Liberman, Harris, Kinney, and Lane, 1961) and is believed to occur when stimuli along a continuum are responded to solely on the basis of the absolute labeling or identification grouping of stimuli (Studdert-Kennedy, et al., 1970). Once recognized as being phonetic in nature, such stimuli are believed to be automatically encoded in terms of the specific phonemes (e.g., Figure 1). The encoding process is usually assumed either to involve a reflection back upon the articulatory mechanism (e.g., motor theory of speech perception [Liberman, et al., 1967]) or by the activation of special feature detectors (e.g., Eimas, 1975). This specialized encoding of phonetic stimuli involves a loss of the acoustic characteristics of the stimuli. As a result, phonetic stimuli are discriminable only to the extent to which they are categorized differentially by the phonetic encoding processes.

The following characteristics provide an operational definition of categorical perception: (a) a sharp labeling boundary along a given continuum, (b) approximately chance performance in discriminating stimuli within categories, and (c) a peak in discrimination performance at the category boundary (Studdert-Kennedy, et al., 1970). In practice, this operational definition has been met quite well by phonetic stimuli varying along a limited number of dimensions, although the discrimination peak sometimes is shifted relative to the labeling boundary (e.g., Liberman, et al., 1957), with discrimination performance typically exceeding that predicted from the labeling data (Cutting and Rosner, 1976).

Over the past few years, the notion that categorical perception is the unique property of phoneme perception has lost considerable credibility. Miller, Pastore, Wier, Kelly, and Dooling, (1974) demonstrated this phenomenon with non-phonetic noise-buzz sequences; both Cutting and Rosner (1974) and Pisoni (1977) found it with stimuli varying in relative onset times; and Pastore, Ahroon, Baffuto, Friedman, Puleo, and Fink (1977) found categorical perception when adding a sinusoidal reference to a sequence of clearly detectable tonal stimuli varying only in intensity. Pastore et al. (1977) also found categorical perception with an intermittent visual stimulus. Related effects have been demonstrated for phonetic stimuli with very young infants (Kuhl, 1976), with chinchillas (Kuhl and Miller, 1975), and with monkeys (Morse and Snowden, 1975), and have been demonstrated with adult humans for musical stimuli (e.g., Burns & Ward, 1974; Siegel & Siegel, 1975). These findings are not consistent with the notion of an absolute causality by a set of innate human processes unique to speech perception.

Recently, we have proposed that categorical perception can be associated with the imposition of some reference along a continuum of stimuli (Pastore, et al., 1977). This reference supplies information which is more precise (for the given purpose) than that normally available for use in discriminating stimuli along a given continuum. Thus, in the region of its imposition along the continuum, the reference creates a natural dichotomy for identification or labeling. We have argued that any efficient communication code would tend to develop naturally around such perceptual references. Specifically, it behooves any efficient communication system to have much clearer distinction between identified codes than within identified codes. We therefore ran a set of experiments in an attempt to demonstrate categorical perception with specific alphabetic stimuli.

If one measures the discriminability of a small constant increment in the length of a line as a function of the initial line length, one would expect to find that discrimination performance obeys Weber's Law (I/I=k). That is, an increment which is just detectable for a short initial line length will not be detectable when the initial line length is longer. If one now adds a constant reference to the longer line lengths, an increment which was not detectable for a given line length in isolation could now become detectable when viewed in close proximity to that fixed reference. Furthermore, the addition of the fixed reference can serve as a natural boundary between two categories (e.g., longer than, or shorter than, the reference). Examples of this situation are diagrammed in Figure 3, which are representations of the stimuli used in our experiments. When the line is vertical and the reference is in the form of a backwards letter "C," the category "shorter-than-reference" is perceived in terms of the common letter category "D," while the "longer-than-reference category" tends to be perceived as the letter category "P." Likewise, when the line is oblique and the reference is a short, straight line at an intersecting angle, we have the two letter categories "V" and "Y."

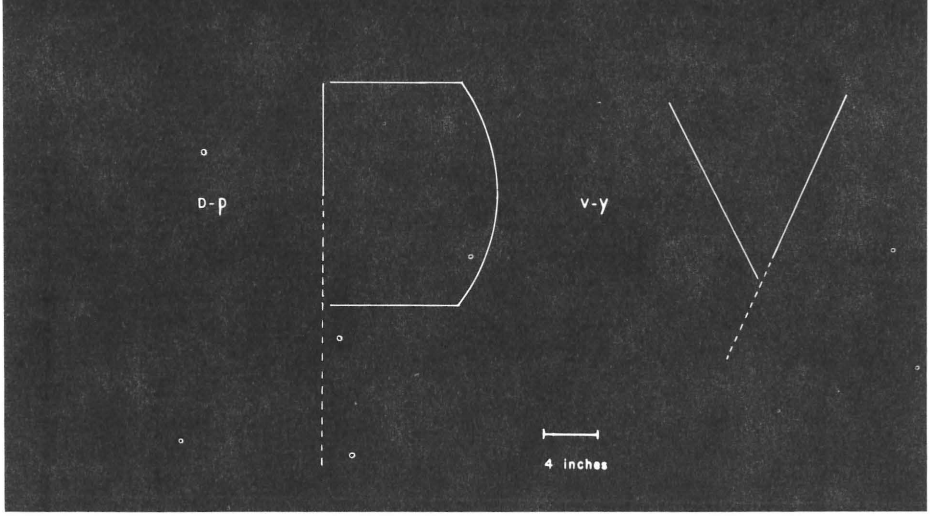
Method

Subjects. The four observers in this experiment were all psychophysically trained and had a minimum of one year of experience as observers in signal detection tasks in the psychoacoustics laboratory. The author served as one of the observers. The two subjects who wore corrective lenses for distance vision, wore such lenses during the experiment.

Procedure. We used standard ABX discrimination and single-stimulus labeling tasks in these experiments. The stimuli were projected with three (one for the labeling task) random-access carousel projectors on a large, flat, painted white wall, and were viewed from a distance of approximately twelve feet. In the ABX task, the three stimuli were presented simultaneously for one second with the A (left) and B (right) stimulus positions at the same vertical distance (approximately three feet) from the floor, while the X stimulus was presented in the horizontal center of these two and was displaced slightly upward (approximately four inches) from the A and B stimuli. The three stimuli (A, X, and B) covered a total horizontal distance of approximately nine feet (approximately 41° of visual angle). (The dimensions of the components of each stimulus may be obtained from Figure 3, which is drawn to scale, with the reference (4 inches) being based upon the projected image of the stimuli and corresponding to a visual angle of approximately 1.6°.) The size and spacing of the stimuli were chosen to prevent the observers from making *simple,* direct comparisons across stimuli and to force the observers to scan across stimuli. In most studies on categorical perception, the A, B, and X stimuli are presented sequentially, thus requiring that the observers either use memorial images of the stimuli, or make within stimulus judgments and compare these judgments. Thus, while acknowledging that alphanumeric stimuli are spatially displaced while phonetic stimuli are temporarily displaced, our procedure is analogous to that typically employed in categorical perception studies. In the ABX task, each pair of stimuli was presented eight times in an order which counterbalanced for the assignment of the stimuli to the A and B positions (for line length: A>B or A<B), and the identity of the X stimulus (X=A or X=B). In both the labeling and the ABX discrimination tasks the order of presentation of the stimuli (or pairs of stimuli) was randomized.

All stimuli were generated from carefully prepared ink line-drawings. In all cases two original sets of stimuli were generated. One

Fig. 3 Diagram, drawn to scale, of the stimuli employed in the D-P and V-Y experiments. The broken lines represent those portions of the stimuli which were varied.



set was photographed as the X stimuli, while the other set was photographed twice and used as the A and B sets of stimuli. Since any small errors or imperfection in the stimuli, would not be common to both the X and the comparison stimuli, this procedure minimized the possibility that the observers could use such imperfections in performing the discrimination. The stimuli are drawn to scale in Figure 3. Both sets of stimuli were varied in terms of the length of the major line component (broken line in Figure 3).

With the D-P set stimuli, the vertical line was varied, while with the V-Y stimuli, the positive diagonal was varied. Finally, the same tasks were run with stimuli compared only to the vertical line which varied in length between 18-3/4 and 23-1/4 inches (range of visual angles of approximately 7.5° to 9.3°). This provided a baseline condition against which the effects of imposing a fixed reference could be evaluated. The projected stimuli were adjusted so that in a progression of increasing line lengths, adjacent stimuli (and thus compared stimuli) would differ by 1/2 inch (or approximately 72′ of visual angle).

Results and Discussion

Figure 4 plots the results of the D-P discrimination task, with the data pooled across the four subjects.[1] For this D-P comparison, we ran the discrimination task for the three sets of stimuli which differed only in terms of the length of the vertical line (8-13 inches, 14-19 inches, 24-28 inches). All four observers independently reported that all of the stimuli with the shortest set of line lengths (8 to 13 inches) always were perceived as "D," while all the stimuli with the longest set of line lengths (24 to 28 inches) always were perceived as "P." Thus, the complete labeling tasks for the two extreme sets of stimuli were not run. In general, discrimination between adjacent stimuli drawn from within a single labeling category is clearly at or near chance (50% correct). The means (and standard deviations) were 65% correct (sd=9) and 53% correct (sd=10) for short and long sets of line lengths, respectively. The stimuli with vertical line length in the range of fourteen to nineteen inches straddle the boundary between the "D" and "P" categories. Here we find a sharp labeling boundary with an associated peak (92% correct) in discrimination performance. Therefore, these data meet the criteria for categorical perception: a sharp boundary between labeling categories, a peak in the discrimination function approximately at the labeling boundary, and close to chance discrimination performance within categories.

The ability of the same subjects to discriminate differences in the length of vertical lines presented alone (with approximately the same range of line lengths as used in the V-Y experiment described next) are diagrammed in the left frame of Figure 5. We could see little value in running these subjects with the longer and shorter sets of line lengths, since our subjects all indicated that for all comparisons they could do little more than guess as to the correct answer. As can be seen, discrimination performance is relatively uniform across the continuum and is close to chance (mean of 64% [sd=8.8]). The boundary between the categories "short" and "long" is not sharp, and there definitely is not a major peak in the discrimination function at or near that boundary. Therefore, it is evident that the imposition of the "backward-c" served as a reference or standard which created a sharp category boundary (between D and P) and a peak in the discrimination function at the category boundary. After the completion of the experiment, all of the observers reported having recognized and used these cues. It can be argued that such a reference makes the information code more efficient by heightening the discriminability of stimuli between categories, thus reducing the uncertainty concerning the specific symbol code in any visual presentation of a component of the language code.

The labeling data for the categories "long" and "short" are also shown in the left panel of the figure. Following Weber's Law [I/I=k], we would expect that the discrimination of the same half-inch increments in line length might be easier for shorter lines and more difficult for longer lines. This did not occur for the single lines, but there was a slight tendency toward this relationship in the overall D-P and V-Y data, although not to the degree reported by Samuel (1977) for phonetic stimuli.

The data from the V-Y continuum are shown in the right panel of Figure 5. As with the data from the D-P continuum, we find a very sharp boundary between the "V" and the "Y" categories, and a peak (100% correct) in the dis-

Fig. 4 Results of the labeling (circles) and discrimination (squares) tasks for the D-P experiment.

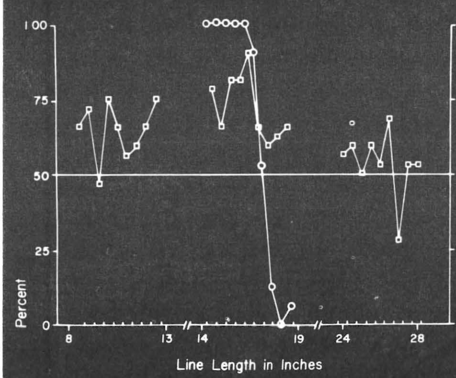
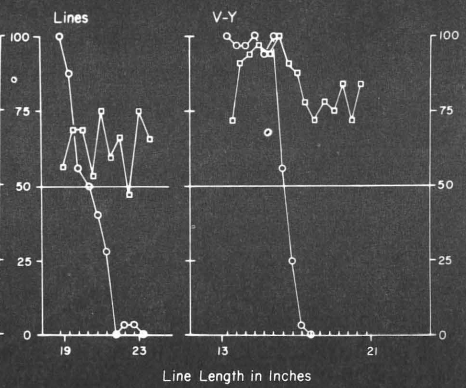


Fig. 5 The results for the labeling tasks (circles) and discrimination tasks (squares) for the line length (left panel) and V-Y experiments (right panel). The labels short and long (for Lines) and V or Y (for V-Y) were employed in the labeling tasks.



crimination function at this boundary. However, the discrimination peak is broader than that for the D-P stimuli and discrimination performance within categories is not reduced to chance, but only reduced from a peak of 100% at the category boundary, to approximately 75% within categories. This absence of the demonstration of chance performance within categories could argue against these results as unequivocally demonstrating categorical perception (Abramson, 1977). Our observers all reported being able to compare the distance between the lower end of each of the two diagonal lines and the theoretical intersection of these diagonal lines. Since the two diagonals were symmetrical around the vertical axis, it would appear that the stationary minor diagonal could serve as the visual reference both at the point of theoretical intersection with the major diagonal (as expected) and for a small range of shorter major diagonal line lengths. This extended reference did not alter the judgment of the categorical boundary between V and Y. It is also possible to argue on logical grounds that, while exhibiting a broad peak, these data represent as much a demonstration of categorical perception as the data from the D-P experiment above and data from most experiments proportedly demonstrating categorical perception. Based upon the data from the D-P experiment, it would seem reasonable to expect that chance performance within categories would have been found had we extended the range of line lengths to be more compatible with the range employed in the D-P experiment (i.e., 8 to 28 inches instead of 13 to 21 inches). Moreover, we again have a clear instance in which a portion of a letter, in this case the fixed diagonal line, served as a standard creating a sharp boundary and a heightened degree of discriminability at that category boundary. Thus, we have the action of the same type of cue for perceptual organization which logically should (and in this case probably is) an important factor in defining these components of the visual language code.

The results of the D-P and V-Y experiments demonstrate that the important phonetic phenomenon of categorical perception also exists for the components of the visible language code. We suspect that we also would find categorical perception for the line length continua between other pairs of visual stimuli, such as Y-X, D-b, and h-n. These findings should not be surprising, since both phonemes and letters are codes for effective communication systems. Such systems should have evolved around perceptual characteristics which would tend to enhance the distinction between different component codes. If we are correct in this conception of categorical perception, and in our conception of the perceptual basis for the equivalence of physically different forms of the language codes (described earlier), then we see little need for hypothesizing a hybrid serial-parallel model (Figure 1) for visible language. Rather, we conceive of many of the salient apsects of the perception of the audible and visible codes for language as expected manifestations of effective communication systems.

1 The data for the individual subjects were highly similar. These data may be obtained by writing the author.

References

Abramson, A. S.
Noncategorical perception of tone categories in Thai.
Journal of the Acoustical Society of America,
1977, *61,* 566(A).

Beardslee, D. C., & Wertheimer, M. (Eds.),
Readings in Perception.
Princeton, N. J., Van Nostrand, 1958.

Burns, E. W., & Ward, W. D.
Categorical perception of musical intervals.
Journal of the Acoustical Society of America,
1974, *55,* 456(A).

Cutting, J. E., & Rosner, B. S.
Categories and boundaries in speech and music.
Perception & Psychophysics,
1974, *16,* 564-570.

Cutting, J. E., & Rosner, B. S.
Discrimination functions predicted from categories in speech and music.
Perception & Psychophysics,
1976, *20,* 87-88.

Denes, P. B., & Pisoni, E. N.
The Speech Chain:
The Physics and Biology of Spoken Language.
New York: Anchor Press, 1972.

Eimas, P. D.
Auditory and phonetic coding of the cues for speech: Discrimination of the R-L distinction by young infants.
Perception & Psychophysics,
1975, *18,* 341-347.

Fant, G.
Analysis and synthesis of speech processes.
In *Manual of Phonetics,* Malmberg (Ed.), North-Holland Publishing Co., P. 173.

Geldard, F. A.
The Human Sense.
New York: Wiley, 1972.

Gibson, J. J.
The Senses Considered as Perceptual Systems.
Boston: Houghton Mifflin, 1966.

Hechenmueller, E. G.
Stabilization of the retinal image: A review of method, effect, and theory.
Psychological Bulletin,
1965, *63,* 157-169.

Hochberg, J. E.
Perception.
Englewood Cliffs, N.J.: Prentice-Hall, 1964.

Julesz, B., & Hirsh, I. J.
Visual and auditory perception—an essay of comparison. In E. E. David & P. N. Denes (Eds.),
Human Communication: A Unified View.
New York: McGraw-Hill, 1972.

Kavanagh, J. F., & Mattingly, I. G.
Language by Ear and Eye: The Relationships Between Speech and Reading.
Cambridge, Mass.: MIT Press, 1972.

Kuhl, P. K.
Speech perception in early infancy: The acquisition of speech-categories. In S. K. Hirsh, D. H. Eldredge, I. J. Hirsh, & S. R. Silverman (Eds.),
Hearing and Davis: Essays Honoring Hallowell Davis.
Washington University Press, St. Louis, Missouri, 1976, 265-280.

Kuhl, P. K., & Miller, J. D.
Speech perception by the chinchillas: Voiced-voiceless distinction in alveolar plosive consonants.
Science, 1975, *190,* 60-72.

Liberman, A. M.; Cooper, F. S.; Shankweiler, D. P., & Studdert-Kennedy, M.
Perception of the speech code.
Psychological Review,
1967, *74,* 431-461.

Liberman, A. M.; Harris, K. S.; Hoffman, H. S., & Griffith, B. C.
The discrimination of speech sounds within and across phoneme boundaries.
Journal of Experimental Psychology,
1957, *54,* 358-368.

Liberman, A. M.; Harris, K. S.; Kinney, J. A., & Land, H.
The discrimination of relative onset-time of the components of certain speech and nonspeech patterns.
Journal of Experimental Psychology,
1961, *61,* 379-388.

Liberman, A. M., & Pisoni, D. B.
Evidence for a special speech-perceiving subsystem in the human.
In: T. H. Bullock (Ed.),
Recognition of Complex Acoustic Signals.
Berlin: Dahlem Konferenzen, 1977.

Liberman, A. M., & Studdert-Kennedy, M.
Phonetic Perception. In:
Handbook of Sensory Psychology, Vol. VIII,
"Perception," R. Held, H. Leibowitz, and H. L. Teuber (Eds.), Heidelberg: Springer-Verlag, Inc., (1977).

MacMillian, N. A.; Kaplan, H. L., & Creelman, C. D.
The psychophysics of categorical perception.
Psychological Review,
1977, *84,* 452-471.

Miller, G. A.
The magical number seven, plus or minus two, or some limits on our capacity for processing information.
Psychological Review,
1956, *63,* 81-96.

Miller, J. D.; Pastore, R. E.; Wier, C. C; Kelley, E. J., & Dooling, R. J.
Discrimination and labeling of noise-buzz sequences with varying noise-lead times: An example of categorical perception.
Journal of the Acoustical Society of America,
1976, *60,* 410-417.

Morse, P. A., & Snowdon, C. T.
An investigation of categorical speech perception by rhesus monkeys.
Perception & Psychophysics,
1975, *17,* 9-16.

Pastore, R. E.
Categorical perception: A critical re-evaluation. In S. K. Hirsh, D. H. Eldredge, I. J. Hirsh, & S. R. Silverman (Eds.), *Hearing and Davis: Essays Honoring Hallowell Davis.* Washington University Press, St. Louis, Missouri, 1976.

Pastore, R. E.; Ahroon, W. A.; Buffuto, K. J.; Friedman, C. J.; Puleo, J. S., & Fink, E. A.
Common factor model of categorical perception.
Journal of Experimental Psychology: Human Perception & Performance, 1977, *4,* 686-696.

Pisoni, D. B.
Identification and discrimination of the relative onset time of two-component tones: Implications for voicing perception in stops.
Journal of the Acoustical Society of America, 1977, *61,* 1352-61.

Pritchard, R. M.
Stabilized images on the retina.
Scientific American, 1961, *204,* 72-78

Samuel, A. G.
The effect of discrimination training on speech perception: Non-categorical perception.
Perception & Psychophysics, 1977, *22,* 321-330.

Siegel, W., & Siegel, J. A.
Categorization of tonal intervals by musicians: Perceptual effect or response bias?
Journal of the Acoustical Society of America, 1975, *58,* S83 (A).

Stevens, K. N.
Perception of phonetic segments: Evidence from phonology, acoustics, and psychoacoustics. In D. C. Horton and J. J. Jenkins (Eds.), *The Perception of Language,* New York: Charles Merrill, 1971.

Studdert-Kennedy, M.
Speech Perception.
In Contemporary Issues in Experimental Phonetics, Lass (Ed.), 1976, p. 243.

Studdert-Kennedy, M.; Liberman, A. M.; Harris, K. S., & Cooper, F. S.
Motor theory of speech perception: A Reply to Lane's critical review.
Psychological Review, 1970, *77,* 234-249.

Wertheimer, M.
Untersuchungen zur Lehre von der Gestalt: II.
Psychologische Forschung, 1923, *4,* 301-340. (Abridged translation by M. Wertheimer, In D. C. Beardslee & M. Wertheimer (Eds.), *Readings in Perception.* Princeton, N.J.: Van Nostrad, 1958.

Wood, C. C.
Auditory and phonetic levels of processing in speech perception: Neurophysiological and information-processing analyses.
Journal of Experimental Psychology: Human Perception and Performance, 1975, *104,* 3-20.

From Print to Meaning and from Print to Sound, or How to Read Without Knowing How to Spell

Visible Language, XII 1 (Winter 1978), pp. 43-54.
Author's address: MRC Developmental Psychology Unit, London WCIH OAN, England.
0022-2224/78/0100-0043$02.00/0

Two groups of 12-year-olds, both of normal intelligence and reading age, were compared. One group consisted of good spellers, the other of poor spellers. The two groups were equally good at reading single words and sentences. However, they differed on other reading tasks, notably with nonsense words and other tasks involving conversion of print into sound. The differences indicated that the poor spellers were proficient at going from print directly to meaning, but were impaired at converting print to sound. In contrast, the good spellers showed mastery of both aspects of reading, converting print to meaning and converting print to sound.

Poor spelling ability can coexist with good and even excellent reading ability. This has not always been recognized because the combination of deficit and skill in two highly related functions is paradoxical enough to rouse suspicions. Most people suffering from spelling problems without reading problems seem to react to this handicap as if they too consider it paradoxical: they often disguise their spelling errors by poor handwriting; they have excuses for not writing at all; they make use of dictionaries at inordinate cost of effort and time; or they hold spelling in contempt as a pedantic and trivial matter. All this makes for successful camouflage of the problem, and it is therefore difficult to estimate its incidence. Nevertheless, in some circumstances this problem can be a serious handicap.

The question that deserves some explanation is how input processes (reading) and output processes (writing) can be divorced to such an extent that one functions well, the other poorly. To explore this question a series of experiments was carried out. The subjects were normal school children, aged 12, of average intelligence and with an average reading age as estimated by a standard reading test, the Schonell graded word list (Schonell, 1942). They were also given the Schonell Spelling test (Schonell,

Table 1

Performance on Schonell Graded Word Lists (Reading and Spelling).

		Good Spellers		Poor Spellers	
		$\overline{X}$	Sd	$\overline{X}$	Sd
n = 10	Reading Qu.	105.6	(5.6)	100.7	(9.0)
n = 10	Spelling Qu.	109.0	(6.1)	89.7	(3.9)

1942). Ten good spellers and ten poor spellers were selected from the population of 12-year-olds of three secondary schools, with the restriction that they were of equal reading achievement. The achievement in reading and spelling was expressed as the quotient of reading age or spelling age and chronological age, multiplied by 100. This is shown in Table I. By definition, there was no difference on reading ability between the two groups but a significant difference on spelling ability.

The present poor spellers differ from other poor spellers who, besides spelling difficulties, also suffer from reading problems. Poor spellers of this type differ qualitatively in both their spelling and reading from the children labelled "poor spellers" here (Frith, note 1). Nelson and Warrington (1973) have compared the spelling errors of similar groups of children, diagnosed as dyslexic. They found that those children who suffered from both severe reading and spelling problems showed an underlying language deficit, while those children who suffered mainly from spelling problems and had largely overcome their reading difficulties did not show a language deficit.

Experiment I

The obvious question to ask was how well the poor spellers would be able to read their own misspellings. It seemed possible that the orthographic structure of words in general was perhaps less important to poor spellers, not only for output but also for input processes. Therefore words were taken from the Schonell Spelling test that each child had previously written after dictation. For each child individually, his last 12 misspelled words and his last 12 correctly spelled words were chosen. These words were typed, presented in random order, and the children were asked to read the words as quickly as possible. Poor spellers found their own misspellings harder to read (78%) than the correctly spelled words (94%), $p<.02$. Good spellers found both equally easy (96% to 91%). The interaction just missed significance at the .05 level.

Thus, it appears that poor spellers benefit more from intact orthographical structure than do good spellers. Spelling conventions are not an irrelevant aspect of reading for the poor spellers. This sensitivity to orthographic structure is not reflected in their awareness of whether a misspelling has occurred. The poor spellers thought that 74% of their own misspellings, as presented in the list, were actually correct, while the good spellers thought this in only 50% of the cases. This was significantly fewer.

The difference between good and poor spellers in Experiment I might have been due simply to the fact that poor spellers made more serious mistakes. Hence good spellers, too, should have found them difficult to read. To eliminate this possibility in the second experiment, both groups had to read the same misspellings. These misspelled words were presented in the context of a story because single word reading is rather atypical. Another reason for embedding the critical words in text was that many misspellings would be highly ambiguous without the aid of context cues, for example, the misspelling "ar" could stand for air, arc, an, or, etc.

Experiment II

Reading prose aloud was the basic task in the second experiment. Embedded in the text were misspelled words which the children were asked to read aloud as if correct. To ensure that all children would read for meaning a variant of the "cloze" procedure, the "maze" technique (Guthrie, 1973) was employed. At frequent points in the text there was a choice between three words of the same class, only one of which was meaningful in the context.

On the basis of a qualitative analysis of the spelling errors obtained from the children the following two contrasting types of misspellings were made up: one type largely preserves the sound, but not the visual appearance of the target word; the other type largely preserves the appearance but not the sound. The omission of just one letter in a word was considered an appropriate means of making up this second type of error.

The text was based on passages from Jules Verne's *20,000 Leagues Under The Sea.* It was divided into four paragraphs, each of exactly 150 words. There were 15 target words in each paragraph, and in addition there were 9 words (3 nouns, 3 verbs, 3 function words) which were presented as a triple choice, in terms of the maze procedure. A few examples including both phonetic and nonphonetic misspellings may illustrate the technique.

There were four versions (conditions) for each of the four paragraphs.

1) Normally typed next (baseline condition).
2) Normal text with multiple choices (correct condition).
3) Text with multiple choices and the 15 misspelled target words that preserved sound, but not appearance (phonologically similar condition).

That night about eleven o'clock, I rceived a most unexpcted

village
sky from Captain Nemo.
visit

far
"Mr Aronnax, so plenty you have only vsited the ocean depths
and

sing
by daylight. Would you care to see them in the drkness of
talk

the nght? "

"Most willingly."

flower
"I warn you, the bird will be tirng. We shall have far
way

to walk and we must clmb a big mountain.

4) Text with multiple choices and the same 15 misspelled target words that now preserved appearance, but not sound (visually similar condition).

Each child received all four paragraphs, each in one of the four conditions in random order. The specific content of a particular section would thus have no systematic influence on a particular condition. The experimenter used a check list to mark any errors while the child read aloud, and also unobtrusively timed the reading.

Results and Discussion

The time scores were unaffected by the different types of misspellings. The average time taken to read the three paragraphs with multiple choices—with or without misspellings—was 67 secs (Sd=10) by good spellers, and 100 secs (Sd=28) by poor spellers, $p<.01$. The baseline condition (normal text without multiple choices) was read significantly faster by each of the two groups. Good spellers took 56 secs (Sd=8) while poor spellers took 70 secs (Sd=19) on average, $p<.05$. Thus, the poor spellers were consistently slower readers and this slowness needs further exploration.

It was somewhat disappointing that the misspellings were almost always recognized correctly. The present groups of children were well able to cope with both the embedded misspellings and the multiple choice procedure. On average only 4% (Sd=4) of the misspellings were not recognized correctly and there was no difference between the two groups. There was however a suggestion of a group x type of misspelling interaction. Only one of the five good spellers who had made any errors at all found phonetically similar misspellings harder than visually similar ones, compared to four out of the six poor spellers who had made errors, $p=.16$. Using t-tests, it was found that poor spellers were worse than good spellers (5% vs 1% errors) with misspellings that preserved sound, but were equal (3% vs 5%) with misspellings that preserved appearance. One might speculate that poor spellers recognize words better from their visual appearance, while good spellers recognize words better from their sound. Thus, poor spellers may have a "visual" approach to reading, i.e. they convert print into meaning, but do not normally convert print into sound. This hypothesis would also be consistent with the result of the first experiment. The majority of spelling errors that both groups of children had made on the Schonell dictation test were those that preserved the sound rather than the appearance of the word. There too it was found that the poor spellers were less good at reading these misspellings. It is likely that the possibility of making use of context in the present task aided the recognition of misspelled words to such a large extent that differences in difficulty were swamped.

Experiment III

The tentative finding concerning the recognition of misspelled words was put to a further test in a proofreading task. Misspelled words were easy to read. Would the errors therefore be easily missed while proofreading? Again, two types of misspellings were embedded in text: one type of misspelling did preserve the sound of the target word (sissors for scissors), the other did not (scarely for scarcely). Unlike the previous experiment both types of errors preserved the visual appearance of the target word. This was achieved by omitting just one letter of each target word. The main reason for this matching of the stimuli in terms of visual similarity was that the effect of sound could be studied without the possibly confounding effect of appearance of misspellings. Another reason was that misspellings that are visually close to the target word would be quite difficult to detect in a proofreading situation and hence a ceiling effect would be avoided. For the same reason also, the misspelled words were placed in highly redundant phrases. Such words might again be scrutinized less and hence misspellings would be noticed less. Five-word phrases were made up for this purpose, for example:

A deep and *horse* voice.
He sharpened knives and *sissors*.
The doctor helped the *patent*.
He *scarely* knew the man.

The letter omission produced another word in half the cases, a non-word in the other half. This meant that while proofreading, it was necessary to read for meaning: scanning for misspelled words alone would result in a large number of misses.

In the previous experiment the poor spellers tended to show little effect of sound compared to the good spellers. Hence, in the present task, it would be predicted that the poor spellers should not show any difference between these two types of misspellings. If they bypass conversion into sound, which this silent reading task certainly permits, they should never "notice" that some misspellings preserve sound and some don't. A prediction for the good spellers if they do *not* avoid conversion of print to sound would be that they *should* "notice" this difference. If good spellers convert print into sound then they should miss more phonologically correct misspellings, as these preserve the sound of the target word and should therefore be more readily accepted as correct. They should miss fewer of the other misspellings since as soon as they are converted into sound, it is clear that they are incorrect.

A list of 32 phrases, all randomly mixed, was presented. Eight phrases contained no errors; 12 contained a misspelled word that preserved the sound of the target word, half of them resulting in a word, half in a non-word; 12 contained a misspelled word that did not preserve the sound of the target word, half again resulting in a word, half in a non-word. The children were given some examples and then asked to go through the list as fast as possible, marking either a cross for any word in the phrase that was wrong in any way, or a tick for all that were correct.

Results and Discussion

There was no difference between errors that resulted in words and those that did not, indicating that reading for meaning was involved, not just proofreading for misspelled words. Also, as one might expect, poor spellers missed on the whole more misspellings than good spellers. The poor spellers missed both types of error equally often, those that preserved sound (33%) and those that did not (42%). This is consistent with the hypothesis that poor spellers did not convert print into sound. For good spellers on the other hand, a very strong effect of sound was obtained. However, the direction of this effect was contrary to prediction. Surprisingly, good spellers missed fewer of the errors when the errors sounded correct (16%) and missed more when the errors sounded wrong (31%), $p<.01$.

The large effect obtained regardless of whether or not it points towards a group difference, is difficult to explain. It does not fit in with the notion that print is converted to sound first. If so, any incorrect sounding misspelling would be noticed more rather than less, as was the case here. A replication of the present task using the prose passages and misspellings of the second experiment showed again that both good and poor spellers noticed more of the correct than of the incorrect sounding misspellings. This indicated that the effect was not an artifact due to a particular set of target words and context phrases.

One explanation is based on a critical examination of the types of misspellings. They did not differ only in terms of sound, i.e., whether or not they preserved the sound of the target word. This must be so since the two types of misspellings could only be derived by taking advantage of certain peculiarities of English orthography. As Chomsky and Halle (1968) have shown, English orthography represents linguistic knowledge on different levels. In particular there is a phonological level and a morphological level. The same sound can often be represented by different letters. Which letters are chosen is then decided on a morphological basis: e.g. "sign" could be spelled sighn, sine, syne, cyne, etc. If it relates to "signature" in meaning, then its spelling must be s-i-g-n. This example of a morphological rule shows that a misspelling that preserves sound by necessity must break such rules. Thus the reason that these misspellings are noticed more, probably means that broken morphological rules are readily detected.

The other type of misspelling that did not preserve sound had been contrived by omitting one letter at a more or less random position in the word. Thus, the omitted letter was not usually critical in terms of deeper level linguistic rules. This explanation is in need of further testing and can only be put forward here very tentatively. However, regardless of the interpretation of the present result, it throws considerable doubt on the notion that the silent reading/proofreading of either poor or good spellers is characterized by conversion of letters to sound. None of the experiments involving reading for meaning gave unequivocal evidence for the hypothesis that poor spellers are less skilled than good spellers at converting letters to sound. It appears that this may well be due to the fact that good spellers, too, when reading for meaning bypass the conversion of print to sound, and thus the hypothesis could not be tested.

There has been a controversy as to whether direct access from the printed word to its meaning without an intervening conversion into sound is in fact possible. This has been reviewed for example by Allport (1977), Coltheart, et al. (1977), and by Massaro (1977). All these authors conclude that, at a minimum, there is no reason to reject the possibility of "direct visual access" from print to meaning. Moreover, there are suggestions that "phonological coding prior to meaning" may not apply to normal reading of prose by any one: by skilled or unskilled, normal or disabled readers (Snowling & Frith, note 3). With meaningful words sounds may always be derived after their meaning has been identified.

One situation where we can be sure that letter to sound conversion must occur is in the reading of nonsense words. Written nonsense words cannot be decoded into meaning, but they can be decoded into sound. If poor spellers are less skilled at converting letters to sounds, then they should be poorer than the good spellers at reading nonsense words.

Experiment IV

Twenty nonsense words were made up on the basis of words in the Schonell Reading list that were well within the reading ability of all the children. Usually just one or two letters were changed, e.g. laucer for saucer, knobbedge for knowledge. Each word selected and hence each nonsense word tested one particular grapheme-to-phoneme correspondence rule, e.g., soft c before e as in laucer, silent k before n as in knobbedge. The child's response was scored as correct if this crucial correspondence rule was correctly applied, even if the pronunciation of the whole word was not as we had anticipated, i.e., analogous to the word it was derived from. Only this scoring was used, since in many nonsense words pronunciation must be considered ambiguous. We did not want to penalize for this inherent problem and hence only scored the critical grapheme-to-phoneme rule which at least in the opinion of two judges was unambiguous.

The good spellers scored 90% correct, the poor spellers 73%, $p<.05$. Ten of 12 given rules were mastered by at least 8 out of the 10 people in the group of good spellers. The poor spellers only mastered 7 rules. This result is interesting, as one might have expected perfect performance from both groups—who had all read the base words correctly. E.g., the word "antique" was read easily by all children, but the derivation "sentique" gave considerable problems. Therefore, it was not entirely knowledge of specific grapheme-phoneme rules which enabled the children to read the words previously. Indeed it is possible to hypothesise that the words had been read without use of such rules at all.

The experiment supports the hypothesis that the poor spellers are poorer at letter-sound conversion than the good spellers. However, the finding that they mastered some of the grapheme-phoneme rules shows that they are capable of using this skill to some extent. Would the two groups also differ on other tasks where phonological coding was required, for example, when judging whether word pairs rhyme or not?

Experiment V

In order to appreciate whether two words rhyme or not, it is necessary to compare the sound of the words. If the poor spellers were less skilled or simply less practiced at converting print into sound, then they should be slower at this task than the good spellers. In contrast to reading aloud, which may well have been disadvantageous to the poor spellers in Experiment II, the rhyming task required silent reading. The sound of words in this case refers to a mental image. If an impairment was again shown then it would seem that the problem is one of internal coding and not of, say, motor programs for articulation. The material was a list of pairs of words, half of which rhymed and half of which didn't, randomly mixed together. Each rhyming pair was visually dissimilar (e.g., might—kite) while each non-rhyming pair was relatively visually similar (e.g., peat—pear). This control ensured that a rhyming judgment could not be based on graphic similarity, but had to be based on sound similarity. The children were shown some examples and asked to put a tick next to every rhyming pair and a cross next to every non-rhyming one. They were given 60 seconds to judge as many word pairs as possible. The good spellers were able to judge 36.5 pairs correctly, the poor spellers 29.1 pairs, $p<.05$. Both groups made the same number of errors (4.1 and 3.8 respectively). This means that the poor spellers were performing about as accurately as the good spellers but were slower. Their slowness in the present experiment supports the hypothesis that the poor spellers are less adept at converting print into sound than the good spellers. Moreover, it suggests that this weakness applies even when only covert sound is involved. It also suggests that the hypothesis could be extended to state that conversion to sound is impaired, at whatever stage in the reading process this conversion was required. This would also account for slower performance when reading aloud, as in Experiment II. Thus one would expect poor spellers to read more slowly than good spellers when reading aloud but not when reading silently.

Therefore, it was now necessary to test the prediction that silent reading of meaningful text would be as efficient and fast in the poor spellers as in the good spellers. If it is true that reading for meaning does not necessarily rely on letter to sound conversion, but can be described as a letter to meaning conversion, and if the poor spellers possess this skill, then they should do as well as the good spellers. Therefore, the following experiment was carried out.

Experiment VI

This task was based on a technique described by Baron (1973). The children were given a list of 10 phrases that they had to read silently as quickly as possible judging whether or not the phrases made sense. There were five of each type randomly mixed. The children were given a practice list and were asked to place a tick next to each phrase that made sense, and a cross next to each phrase that did not. The critical words in the phrases that did not make sense were all visually similar to the target word, in order to ensure a certain degree of difficulty. Examples are:

please wait for me	✓
don't forget to write	✓
bigger that a horse	×
the time war near	×

As predicted, both groups did this task equally well (23.5 sec for good spellers vs 24.2 sec for poor spellers) with equal accuracy (11% and 16% errors on average). Thus the poor spellers are indeed as fast as the good spellers when reading silently. Their apparent slowness in Experiment II may thus well have been due to the requirement of reading out loud.

The present task has been used by Baron (1973) and Baron and McKillop (1975) in an ingenious way so as to impose a particular strategy: they presented a list of phrases, all of which contained a wrong word, and hence all phrases were nonsense. However, half the phrases still sounded as if they made sense. The subjects were now instructed to tick all the phrases that sounded as if they made sense, and cross all those that did not. Below are some examples:

by her a present	✓
she was there mother	✓
a gloss of wine	×
the rain his stopped	×

The hypothesis of a weakness in sound conversion in the poor spellers would predict that they should be impaired in this task. The children were again given examples and instructed to work as fast as possible. The result was in line with the prediction: the poor spellers took 28.8 seconds to complete the task while the good spellers took 19.9 seconds, $p<.05$. The interaction of groups x conditions, based on a t-test on difference scores for both tasks was significant on the .05 level. Hence we can say that poor spellers were less efficient when conversion to sound was required compared to their normal reading strategy. Curiously, the good spellers made only 2% errors in the second condition, while the poor spellers made a similar amount as before (14%). Thus, one may wish to conclude that the good spellers were particularly efficient at this task when instructed to convert print into sound, but did not spontaneously choose this strategy.

Conclusion

The findings of the present series of experiments point to the fact that the two groups of 12-year-olds of similar reading levels but of dissimilar spelling levels compared here, differed on one specific aspect of reading. The poor spellers consistently showed weaker ability in converting letters to sounds at whatever stage in the reading process this conversion applied, whether sound was overt or covert. This weakness is decidedly not a lack of this ability, rather a lack of preference for, or an avoidance of, this aspect of reading. The other aspect of reading, the conversion of letters into meaning, was, however, much used and skillfully applied by this group. This was the case in reading meaningful words alone and in context. This shows that reading can go a long way with a preferred strategy of direct access to meaning.

The good spellers have been shown to be able to convert letters into sound very skillfully, when this was necessary for the task. This was the case with nonsense words, the rhyming judgment, the judgment of whether phrases sounded as if they made sense, and the recognition of grossly distorted words that sounded correct. On all these tasks the good spellers were superior to the poor spellers. On tasks that did not necessarily require converting letters into sounds, the two groups were similar. Thus, it seems that the good spellers are equally adept at both aspects of the reading process.

The present experiments have demonstrated that good reading ability can indeed coexist with poor spelling ability. Does this discrepancy now seem less paradoxical? In some respects the paradox has perhaps lost some of its initial implausibility. A separation of input and output processes in visible language must be postulated. It seems very likely that a similar separation would apply to spoken language. Speech perception may exceed speech production (Dodd, 1975). For the special group of children who show spelling problems without reading problems, a minor deficit in terms of input processes has been demonstrated. The question remains how the minor difficulties in the input processes can be related to the major difficulties in the output processes. The aim of the present paper was not to explore these relationships, but to demonstrate that two kinds of skilled reading must be distinguished: conversion to meaning and conversion to sound. The evidence collected here shows that normal young readers, whenever possible, prefer to go directly from print to meaning.

Reference Notes

1. Frith, U.
Spelling difficulties with and without reading difficulties. In: *Cognitive Processes in Spelling,* U. Frith (Ed.). Y.Y.: Academic Press. To appear.

2. Frith, U.
Spelling difficulties: An annotation. *Journal of Child Psychology and Psychiatry.* In press.

3. Snowling, M. & Frith, U.
Is text that only sounds correct harder to read than text that only looks correct? Submitted for publication.

References

Allport, A.
Letter and word recognition. In: P. A. Kolers, M. E. Wrolstad & H. Bouma (Eds.), *Processing Visible Language.* Conference held at Eindhoven, September 1977. To appear.

Baron, J.
Phonetic stage not necessary for reading. *Quarterly Journal of Experimental Psychology,* *25,* 241-246. 1973.

Baron, J. & McKillop, B. J.
Individual differences in speed of phonetic analysis, visual analysis, and reading. *Acta Psychologica,* *39,* 91-96. 1975.

Chomsky, N. & Halle, M.
The Sound Pattern of English. New York: Haper and Row. 1968.

Coltheart, M., Davelaar, E., Jonasson, J. T. & Besner, D.
Access to the internal lexicon. In: S. Dornic (Ed.), *Attention and Performance VI.* Hillsdale, N.J.: Erlbaum. 1977.

Dodd, B.
Children's understanding of their own phonological forms. *Quarterly Journal of Experimental Psychology,* 27, 165-172. 1975.

Guthrie, J. T.
Reading comprehension and syntactic responses in good and poor readers. *Journal of Educational Psychology,* *64,* 241-247. 1973.

Massaro, D. W.
Reading and listening. In: P. A. Kolers, M. E. Wrolstad & H. Bouma (Eds.), *Processing Visible Language.* Conference held at Eindhoven, September 1977. To appear.

Nelson, H. E. & Warrington, K. E.
Developmental spelling retardation and its relation to other cognitive abilities. *British Journal of Psychology,* *65,* 265-274. 1974.

Schonell, F. J.
Backwardness in the Basic Subjects. London: Oliver & Boyd. 1942.

I am grateful to Jocelyn Robson who assisted in carrying out all the experiments reported.

Using Spelling-Sound Correspondences Without Trying to Learn Them

Visible Language, XII 1 (Winter 1978), pp. 55-70.
Jonathan Baron's address: Department of Psychology, University of Pennsylvania, Philadelphia, PA 19174.
0022-2224/78/0100-0055$02.00/0

Adult subjects learned spoken responses to nonsense words written in an artificial alphabet. Correspondences between letters and phonemes were hidden by the use of right-to-left correspondences. Even though subjects did not notice the existence of correspondences, they were able to decode new nonsense words in the same alphabet. In a second experiment, nonsense words written with hidden correspondences were read more quickly than nonsense words without correspondences. A third experiment suggested that this effect was due to the fact that similar words had similar responses. In general, the results suggest that correspondences can be used without the use of special strategies for learning correspondences, but when this occurs, people use examples rather than knowledge of the correspondences themselves.

No matter how a child learns to read, much of his experience typically consists of reading aloud. Producing vocal responses to printed words undoubtedly promotes learning of associations between printed and spoken words. The question addressed in this paper is whether anything else is learned as a consequence of this learning. In particular, does the learning of associations between printed and spoken words enable a beginning reader to decode words he has never seen before in print?

The practical importance of our question is this: if a child somehow gets to the point at which he can read aloud, does he need further training in finding and using spelling-sound correspondences? Or, alternatively, will practice at reading aloud inevitably lead to learning to use correspondences that do exist (Venezky, 1970)? For example, if a child learns to read by the whole-word method, it might be inevitable that he will ultimately learn to use the correspondences with further practice at reading, even reading of fairly simple materials (on the assumption that these materials contain most of the correspondences he will need to know). If such learning occurs, reliance on practice alone might be a useful alternative method for teaching children who seem unable to learn to use correspondences from direct instruction. Of course, it is also possible that such learning from practice works for some children but not others.

It is useful to distinguish among four different ways in which a person might come to use spelling-sound correspondences as a result of experience at learning associations between printed and spoken words. I shall call these ways *implicit abstraction, explicit abstraction, generalization,* and *analogy.*

In *implicit abstraction* the learner comes to represent the actual rules of correspondence, possibly in the form of associations between letters and phonemes or between groups of letters and groups of phonemes. Such learning occurs in the absence of special strategies of any sort; the learning is assumed to be the result of the operation of basic learning mechanisms. Some theories of learning would seem to imply that such learning would occur.

In particular, some versions of stimulus-response theory might hold that learning is inevitable when a response is intentionally produced in the presence of an attended stimulus. When a person produces a spoken word in response to a printed word, all the conditions are present for learning an association between a part of the printed word and a part of the spoken word. The part of the spoken word is indeed an intentionally produced response. And the part of the printed word is apparently attended to. By this mechanism, associations would be formed between every letter in a word and every phoneme in the response. For example, if the stimulus were *bit* and the response /bit/, associations would be formed between the *b* and the /b/, the *b* and the /i/, the *b* and the /t/, the *i* and the /b/, etc. Only some of these associations would be those we are interested in having a child learn. However, if a child learned a large set of words, it seems likely that some of these associations would be strengthened by repetition and others would not. For example, if the next word a child learned were *bam,* the association between the *b* and the /b/ would be strengthened, but the association between the *b* and the /i/ would not.

This learning is called implicit because it is not the result of any explicit effort to learn what is ultimately learned. Technically, implicit learning is a kind of incidental learning; a person is given the task of learning (or practicing) associations between printed and spoken words, and, as a result, he incidentally learns associations between letters and sounds. The study of implicit learning has received a large boost from the recent work of Reber (e.g., 1976; see also Brooks, 1977, for another view). Reber examined the learning of the rules of artificial grammars governing sequences of letters, measuring learning by the ability to classify new sequences according to whether or not they conform to the grammar. In one experiment, subjects instructed simply to memorize grammatical sequences learned more of the grammar than did other subjects instructed to try to discover the rules of the grammar while they were memorizing. Reber (personal communication) has suggested that there might be special mechanisms for implicit learning and that these mechanisms might be used not only in such things as learning a first language but also in learning spelling-sound correspondences from practice at reading. Reber does not identify these mechanisms with those of stimulus-response learning of part-part associations, but for present purposes, such learning seems roughly equivalent to true implicit learning.

In *explicit abstraction,* the learner uses special strategies for discovering rules at the time that the examples of the rules (correspondences) are presented. This is probably what most of us would do if we were asked to memorize the pronunciations of a number of Greek words (assuming we don't know Greek). While we were memorizing the words, we would also be generating and testing hypotheses about the sounds of individual letters such as rho and gamma. As adult readers, this would be particularly easy because we already have a way of representing phonemes abstractly, that is, a way of naming phonemes. As we discover the sounds of the Greek letters, we can formulate rules of the form, "The thing that looks like a P sounds like an R." It is not at all clear that young children have a way of representing phonemes before they learn to read, so we cannot assume that this explicit mechanism is open to them.

Previous work on the learning of whole words in artificial alphabets is consistent with the view that explicit abstraction is the only mechanism available for learning letter-sound correspondences. Bishop (1964) taught a group of college students associations between printed and spoken nonsense words using Turkish letters. Bishop then tested subjects on their ability to use correspondences by asking each subject to pronounce new words according to the same correspondences. Eight of the twenty subjects did not consciously discover any correspondences, and these subjects performed at chance levels on the transfer test. The remaining twelve subjects were able to pronounce some of the transfer-test items, and these subjects were also able to report explicit spelling-sound correspondences they had discovered.

Jeffrey and Samuels (1967) did a similar experiment with six-year-old children, using artificial letters as stimuli and consonant-vowel English words as responses. Without additional training in "phonic blending," none of the children was able to decode any new words on the basis of correspondences present in the original set. The experiments to be reported here show that transfer of correspondences *can* occur in the absence of explicit abstraction. Our experiments thus stand in apparently direct contradiction to the earlier studies, and we are forced to face the question of why transfer without explicit abstraction occurs in some situations but not others.

One consequence of the existence of letter-sound correspondences is that words that look similar will have pronunciations that sound similar (Brooks, in press; Baron, 1977). In most cases, three-letter words with two letters in common (e.g., pat—pan) will have pronunciations with two phonemes in common; words that have no letters in common will have pronunciations with no phonemes in common, etc. Such a state of affairs could easily facilitate transfer of learning on the basis of well-known mechanisms of stimulus and response *generalization* gradients (Osgood, 1949). In general, it has been found that there is a great deal of positive transfer between learning one stimulus-response association and learning another when the two stimuli are similar and the two responses are similar. This finding is presumably the result of basic learning mechanisms rather than the result of application of any special strategies; the same principle ought to apply to pigeons as well as children.

Such a mechanism might even operate to allow some decoding of new words presented for the first time (Brooks, in press). For example, suppose a child knew the words *bat, rat,* and *ran,* and we give him the new word *ban.* Because of the similarity between *ban* and each of the three original words, the responses to these words will be activated by stimulus generalization. Possibly one of the known responses will then be produced (as often happens with young children presented with such a novel word). However, another possibility is that none of the three original responses will be produced, because these responses will inhibit one another through competition. In this case, we might imagine that some sort of "response averaging" mechanism would take over. A new response would be constructed by taking pieces from the old responses and putting them together (as occurs in Spoonerisms and other tips of the slongue). Such a response would have a good chance of being correct—or at least of being close enough so that the actual word could be recognized with the help of context in a real reading situation. Note that this mechanism of stimulus generalization is like implicit abstraction in that both are a necessary consequence of learning. These mechanisms differ, however, in what needs to be present in memory for transfer to occur. For implicit abstraction, rules themselves must be in memory, for stimulus generalization, only examples.

The final possible mechanism for decoding novel words on the basis of experience with old ones is *analogy* (Baron, 1977). For example, a person seeing BOROUGH for the first time might think of THOROUGH. Analogy is like generalization except that the use of responses to old stimuli is not automatic. Instead, it is assumed that the subject actively tries to recall words that are similar to a new word he is trying to learn or to decode, and then uses additional strategies to modify the pronunciations of these old words to derive the pronunciation of the new word. Note that there are two points at which special strategies are used: first, the recall of similar words; and second, the modification of the pronunciations of these words to derive the pronunciation of the new word.[1]

These four mechanisms for transfer—explicit abstraction, implicit abstraction, generalization, and analogy—may be distinguished in two ways, according to what is learned and according to whether strategies are used. For implicit and explicit abstraction, the rules (correspondences) themselves are learned from exposure to words that follow the rules. Theoretically, it would be possible to forget all the examples and still remember the rules themselves (as seems to occur, for example, in some cases of artificial-grammar learning; see Smith, 1966). For the mechanisms of generalization and analogy, transfer is based on the memory of examples alone. There need be no representation of the correspondences themselves in order for these mechanisms to operate. As for the other distinction, implicit abstraction and generalization do not require any special strategies for learning or transferring rules. They rely on basic mechanisms that are presumably present in all people and that do not have to be learned. On the other hand, explicit abstraction requires the use of hypothesis-testing and representational strategies at the time the examples are presented, while analogy requires the use of retrieval strategies and strategies for modifying pronunciations at the time the transfer-test items are presented.[2]

The present experiments are designed to find out whether other mechanisms besides explicit abstraction, such as analogy or generalization, are available to college students for transferring correspondences on the basis of experience with examples following those correspondences. Depending on what those other mechanisms turn out to be, such studies may shed light on the original question of whether children can learn to use correspondences from practice at reading whole words. In general, these experiments use the artificial-alphabet technique first used by Bishop (1964) to address these questions. However, we make an effort to hide the existence of correspondences from the subjects, so that explicit abstraction will not be used.

Experiment I

In the first experiment twelve college students learned to read a set of nonsense words printed in an artificial alphabet. The subjects were told that there were no correspondences between letters and sounds and that the best strategy was simply to try to memorize associations between whole stimuli and whole responses. In fact, the subjects were deceived. Six of the twelve stimuli had right-to-left correspondences, with the rightmost artificial letter corresponding to the first phoneme, etc. The other six stimuli were derived from a similar set of stimulus-response pairs, but the assignment of responses to stimuli was rearranged so that there were no consistent letter-sound correspondences.

Method. The stimuli and responses used are shown in Table I. In the table the stimuli from the two sets are shown separately, but in the experiment, all twelve stimuli used for a given subject were mixed together randomly (by shuffling the cards on which the stimuli were printed before each trial). This made the correspondences in the correspondence set even less obvious than would otherwise have been the case. In addition, the letters used in the two sets were chosen in pairs; each letter in the correspondence set had a corresponding letter in the no-correspondence set to which it was similar. This was to discourage the subject from looking for regular correspondences by comparing responses to similar stimuli. The stimulus most similar to a given stimulus was in fact from the other set. Assignment of stimuli to sets was balanced across subjects.

Each subject learned the responses to the twelve stimuli, printed on cards, by a self-paced anticipation procedure in which the subject was free to answer or not, as he chose. The correct answer was spoken by the experimenter if the subject did not give it. This procedure continued for 25 trials (runs through the list). Only one subject succeeded in learning all twelve responses. After completing the 25 learning trials, subjects were asked to try to decode transfer words (see Table I). These stimuli were presented in a random order. After the first presentation of the transfer words, the subjects were told that some of the stimuli had had correspondences. The subjects were then shown the transfer words again and asked to try to decode them. Finally, the subjects were told that the correspondences were right-to-left and were asked to try to decode the transfer items a third time. After this, the subjects were interviewed in detail about the way they learned the responses to the original stimuli, about whether they noticed any correspondences, etc.

Results. The main result was that the subjects made fewer errors on the items in the correspondence set than on the items in the no-correspondence set during the learning trials. For the correspondence set, 67% of the responses were correct, for the no-correspondence set, 56%, $p<.01$. The interviews after the experiment revealed that two of the subjects suspected that there might be correspondences (although they could not discover any); one subject actually discovered the correspondences in the middle position (which would, after all, be present even if the subject were looking for left-to-right correspondences), and two other subjects knew how to read Hebrew (which has right-to-left correspondences). When these five subjects were eliminated from the analysis, there was still a significant difference (52.3% vs. 45.6%, $p<.01$) between the conditions across the remaining seven subjects. In sum, the presence of correspondences facilitated the learning of those items containing correspondences even though the subjects were unaware of the existence of the correspondences and thus (presumably) did not employ specific strategies for learning the correspondences.

Eight subjects were used for the analysis of the transfer items.

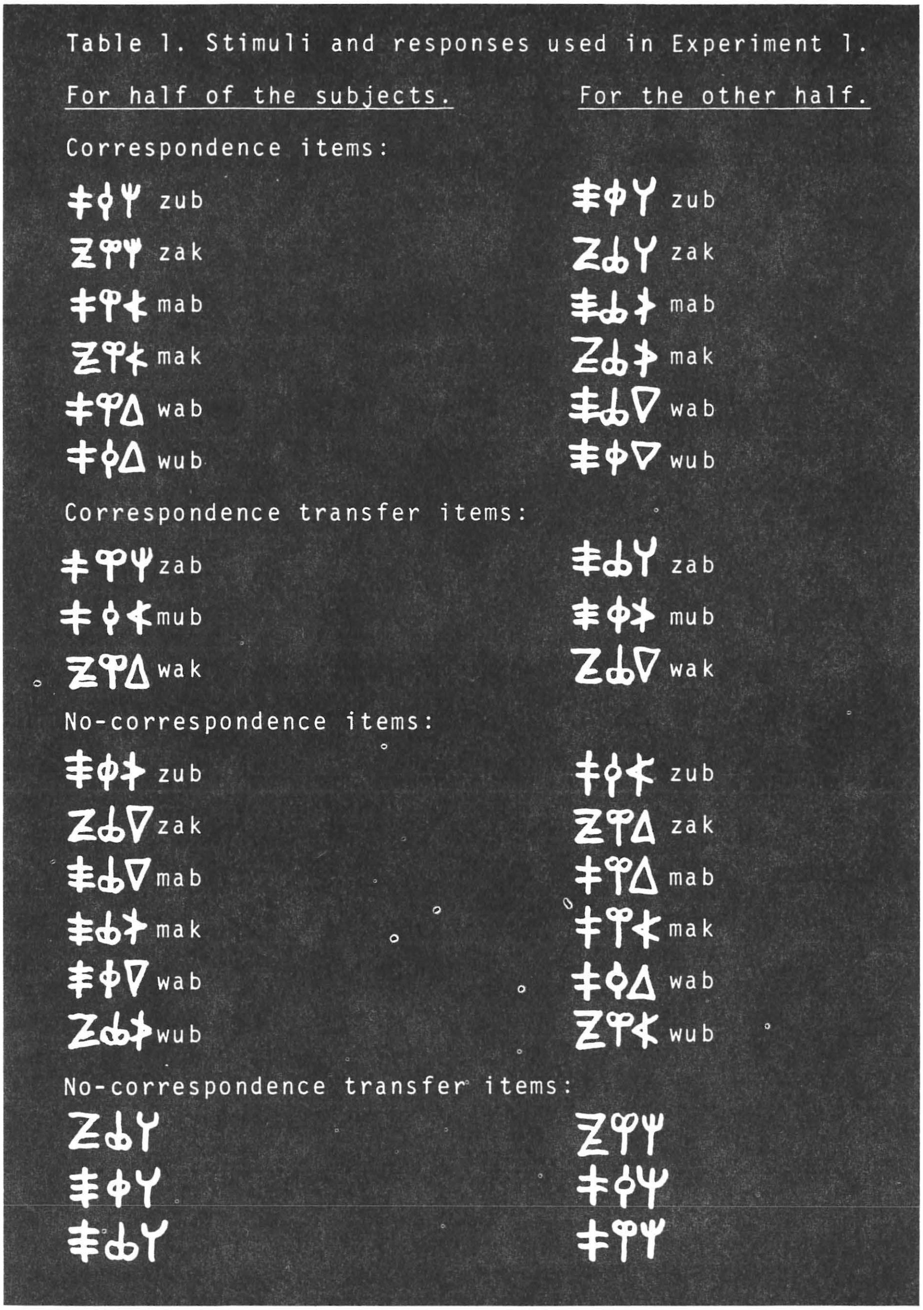


Table 1. Stimuli and responses used in Experiment 1.

For half of the subjects. | **For the other half.**

Correspondence items:

For half of the subjects	For the other half
zub	zub
zak	zak
mab	mab
mak	mak
wab	wab
wub	wub

Correspondence transfer items:

For half of the subjects	For the other half
zab	zab
mub	mub
wak	wak

No-correspondence items:

For half of the subjects	For the other half
zub	zub
zak	zak
mab	mab
mak	mak
wab	wab
wub	wub

No-correspondence transfer items:

The three subjects who noticed correspondences or looked for them during learning were excluded. The other excluded subject was the only one who spelled out his responses instead of pronouncing them, and only 31% of the letters he used were letters used in the spellings of actual responses, in contrast to 81% for the next lowest subject. To find out whether transfer occurred, only the correspondence trials were analysed. The number of phonemes in the correct position was determined for each test trial. The means for the three trials, respectively, were 4.75, 4.75, and 4.38, out of a possible 9.00 on each trial. To find out whether these scores could be due to some sort of guessing, it was necessary to calculate the scores expected by chance. To do this, we assumed (conservatively) that if the subject had guessed, he would have produced the same three responses as he did produce, but would have assigned them randomly to the three stimulus items. These expected scores were calculated for each trial for each subject. Their means were 3.91, 4.22, and 4.10 for three trials, respectively. The difference between the actual scores and the expected scores was not quite significant for any of the three trials alone, but it was significant across subjects by a *t* test for all three trials combined ($p<.03$) and for the first two trials combined ($p<.025$). In sum, the results support the claim that transfer to new items does occur even when the rules transferred were not thought to exist at the time of learning. Further, the subject need not know that the rules go from right to left for such transfer to occur; this is shown by the significant result for the first two test trials.[3]

From the results of this experiment, there seems to be little doubt that subjects can transfer spelling-sound correspondences even though the subjects are unaware of these correspondences at the time of learning associations between printed and spoken words. It thus seems—in spite of the earlier results of Bishop (1964), described above—that other mechanisms besides explicit abstraction are available for transferring correspondences on the basis of experience with whole words.

Clearly, the results can be explained by some implicit-abstraction mechanism (except for the one subject who did not appear to transfer). It is also worth noting that the conditions were present here for either generalization or analogy to operate as well. For example, all of the results could be accounted for by the analogy mechanism. In the correspondence condition, stimuli that have two letters in common always have two phonemes in common. During learning, a subject might be unable to recall the response to some stimulus but might be able to recall the response to a similar stimulus. He may then use the response to the second stimulus as a cue to remember the response to the first; a response sharing two phonemes might be sufficient to remind the subject of the correct response to the stimulus at hand. Likewise, during transfer, the test items might remind the subjects of stimuli presented during training. In fact, each of the three test stimuli had two letters in common with each of two stimuli presented during training. If the subject could recall these training stimuli and their respective responses, he could correctly infer two of the phonemes of the correct response to the test item simply by producing the phonemes common to the two responses he could recall. If he were clever, he might even infer that the remaining phoneme must be different from either of the two training items, and he could be correct on all three phonemes in the test item. No subject appeared to be this clever. Completely correct answers were given on only 12.5% of test trials.

Experiment II

Experiment II does not attempt to find out which of the various mechanisms is available, although its results may bear on this issue. Rather, it attempts to show that spelling-sound correspondences can be useful in fluent responding to printed words as well as in learning the responses to these words for the first time. Also, it shows that correspondences can facilitate memory even in a situation in which all learning is incidental, even the learning of the associations between printed and spoken words.

Method. In this experiment, the subject was given what he thought was only a reaction-time task. He was given an index card with a set of four stimuli and four corresponding responses; he was also given a column of 24 stimuli. Each stimulus on the list of 24 was one of the stimuli on the card. His task was to look at each stimulus on the list of 24, look at the card to find the stimulus and its associated response, produce the response as quickly as possible, and move on to the next stimulus on the list until all 24 stimuli had been responded to. He was not encouraged to memorize the four responses. The two measures of interest were the time required to go through the list and the memory of the responses to the stimuli presented at the end of the experiment.

Some of the stimuli and responses used in this experiment are shown in Table II. Each of the index cards contained a set of stimuli and responses. Half of the subjects used sets A, C1, B, and D2; the other half used sets A, C2, B, and D1. When set A is used in combination with set C1, there are consistent rules relating letters and phonemes. This is also true for set B in combination with set D1. However, when sets A and C2, or sets B and D2, are used together, there are no consistent correspondences. Each letter can correspond to two phonemes instead of one. It is important to note that the correspondences could not be noticed within the members of a single set. The correspondences were consistent only across sets. Thus, the subject had no need to learn the correspondences simply to deal with the card in his hand. It is also worthy of note that half of the correspondences were consistent only when viewed from left to right, and half only when viewed from right

Table 2. Stimuli and responses for Experiment 2, by list.

	List A		B
Γ∇	bah	≃∞	faa
ΣΛ	meh	≡∿	vih
Φ<	gaw	∂Ψ	zuh
T→	koo	§Ω	tay

	C1	C2		D1	D2
ΓΛ	beh	kaw	≃∿	fih	tuh
Σ∇	mah	goo	≡∞	vaa	zay
Φ→	kaw	beh	∂Ω	tuh	fih
T<	goo	mah	§Ψ	zay	vaa

to left. The basic design of this experiment is thus analogous to that of the first experiment, except that learning is incidental.

Twenty college-age subjects were tested, 10 with A and C1 as the correspondence condition, 10 with B and D1. For half of the subjects in each group of 10, the assignments of stimuli to the right-to-left versus left-to-right conditions were switched (by switching the first response on the B1 and D1 card with the second, and the third response with the fourth). Thus, the assignment of stimuli to the correspondence versus no-correspondence conditions, and to the forward versus backward conditions within the correspondence condition, was completely balanced across subjects.

The subjects went through 16 lists (24 items each, four different lists run through four times) with each index card. The order of presentation of conditions insured that sets in the correspondence condition would occur in different blocks of trials. This would minimize the chances of the subjects noticing the correspondences. The subject first alternated between the C and D cards, starting with the C card (i.e., C1 or C2), for a total of eight runs with each card. Then the same was done with the A and B cards. Then the whole procedure was repeated before the memory test. In the memory test, the stimulus items were presented in a list in the fixed order, A, B, C, D. Since there was complete counterbalancing, the order of testing cannot account for any results.

Results. In the memory test, for incidental learning of the responses, the responses in the left-to-right correspondence condition were recalled correctly (in entirety) 37.5% of the time, the right-to-left correspondence items were recalled 35% of the time, and the control no-correspondence items were recalled 22.5% of the time, $p<.05$. This indicates that the consistent correspondences did promote incidental learning. Of interest is the lack of difference (nowhere close to statistical significance) between the left-to-right and right-to-left conditions. This suggests that the effect of the correspondences was not a result of explicit abstraction of the correspondences. Surely, explicit abstraction mechanisms would have been more likely to discover left-to-right rules than the reverse.

The difference between correspondence and no-correspondence conditions was apparent in the times for going through the lists as well as for the memory scores. This effect was tested for sets A and B only, since the C and D conditions were presented first, before the effects of correspondences could be manifest. The mean time for each item in the list was 1.16 sec for the correspondence condition and 1.22 sec for the no-correspondence condition, $p<.05$. The availability of consistent letter-sound correspondences thus improved speed of responding as well as incidental learning of the responses.

Experiment III

So far, the experiments have indicated fairly strongly that spelling-sound correspondences may be transferred without the use of explicit abstraction. The first experiment showed transfer of correspondences of which the subjects were unaware. In the second experiment, the transfer measured was actually transfer of learning (or performance) between sets using the same correspondences, rather than transfer of rules to new instances. However, it is necessary to distinguish among the three remaining mechanisms.[4] One critical test concerns the effect of similarity relations between stimuli and between responses. Recall that both the analogy and generalization mechanisms rely primarily on the fact that similar stimuli have similar responses, whether or not spelling-sound correspondences are preserved. Experiment III is an attempt to make use of this property of these two mechanisms to find out whether true implicit learning—the only mechanism that does not rely on similarity relations—occurs. This experiment is as much like Experiment II as possible, except that both correspondence and no-correspondence conditions are identical in the extent to which similar stimuli are associated with similar responses. In particular, every pair of stimuli with a letter in common now has a corresponding pair of responses with exactly one phoneme in common.

Method. The method is identical to that of Experiment II, except for the responses assigned to the stimuli. Some of these responses are shown in Table III. The remaining sets are identical except that the first and second responses to the C1 and D1 sets are switched, as are the third and fourth responses, as in Experiment II.

Table 3. Stimuli and responses for Experiment 3, by list.

	List A		B
Γ∇	zee	≃∞	foo
ΣΛ	taw	≡∿	kay
Φ<	zee	∂Ψ	foo
T→	taw	§Ω	kay

	C1	C2		D1	D2
ΓΛ	zaw	zaw	≃∿	koo	koo
Σ∇	tee	zaw	≡∞	fay	koo
Φ→	tee	tee	∂Ω	fay	fay
T<	zaw	tee	§Ψ	koo	fay

One way of looking at the design is to examine what the subject would learn if he tried to learn correspondences from left to right or from right to left. Consider sets A and C1. For the first two items in each set, a right-to-left correspondence would work perfectly across the two sets, and for the second two items, a left-to-right rule would work. For this reason, set A was considered to be in the correspondence condition when it occurred along with set C1. Now consider sets A and C2, a combination of conditions received by other subjects. Here, consistent correspondence would not work. There is no way of keeping consistent left-to-right or right-to-left correspondences in lists A and C2 combined. The sets A and C2, when used together, thus make up the control condition. If subjects are learning letter-sound correspondence, they should do better on A and C1 together than on A and C2.

Another feature of this design was that transfer of letter-sound correspondences between lists could be used to read some letters but not others. For example, there is no way the subject who has lists A and C2 can derive the response EE to the first item in A, while there are potentially two ways in which he might derive Z, (1) if he takes the correspondences to go from left to right, (2) if from right to left. If the subject does really learn letter-sound correspondences, then we might expect him to do particularly poorly on these particular correspondences for which no other correspondence could possibly exist. If the subject relies on overall similarity relations, however, there is no reason to expect such a difference.

Results. The data were quite clear. There was absolutely no effect of anything. The mean time per item here was 1.25 sec for the

Osgood, C. E.
The similarity paradox in human learning: A resolution. *Psychological Review,* 1949, *56,* 132-143.

Reber, A. S.
Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory,* 1976, *2,* 88-94.

Smith, K. H.
Grammatical intrusions in the recall of structured letter pairs: Mediated transfer or position learning? *Journal of Experimental Psychology,* 1966, *72,* 580-588.

Treiman, R., & Baron, J.
Separable perception of phonemes: Development and relation to reading. Submitted for publication.

Venezky, R. L.
The Structure of English Orthography. The Hague: Mouton, 1970.

1 We can imagine intermediate cases between analogy and generalization, in which one of these strategies is used but not another. For example, the retrieval of memories of similar words might be the result of basic mechanisms but the modification of the responses to these words might require special strategies.

2 A further complication is the possibility that a person can present items to himself, whether by recalling examples he has learned or by making up transfer items. However, we can treat these self-generated items as equivalent to externally-presented stimuli for purposes of our classification scheme.

3 Typically, subjects give either old responses (mab, zak) or responses with the same phonemes (zam, wuk). Sometimes, new phonemes were inserted. Subjects could have achieved higher scores than they did by matching probabilities of phonemes in positions, but the procedure for estimating scores expected by guessing would also be influenced by such matching, so we can conclude that it did not account for the results. Finally, it may be of interest that the positive results were due almost entirely to correct performance on the vowel (77% obtained vs. 56% predicted, $p<.015$). Performance on the first phoneme in the response was only slightly above chance, while performance on the last phoneme was slightly below chance.

4 We might think that the analogy mechanism would operate too slowly to be useful in Experiment II. But we have no right to make such an assumption. How slowly is too slowly?

Discussions with Lee Brooks have been helpful at every stage of this work. He and Dom Massaro also provided helpful comments on the paper. It was supported by grants from the National Science Foundation (BMS 74-23158) and the Spencer Foundation (through the University of Pennsylvania).

References

Baron, J.
Mechanisms for pronouncing printed words: Use and acquisition. In D. LaBerge and S. J. Samuels (Eds.), *Basic Processes in Reading: Perception and Comprehension.* Hillsdale, N.J.: L. E. Erlbaum Associates, 1977.

Baron, J.
Orthographic and word-specific mechanisms in children's reading of words. Submitted for publication.

Baron, J., & Strawson, C.
Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance,* 1976, *2,* 386-393.

Bishop, C. H.
Transfer effects of word and letter training in reading. *Journal of Verbal Learning and Verbal Behavior,* 1964, *3,* 215-221.

Brooks, L. R.
Non-analytic concept formation and memory for instances. In E. Rosch (Ed.), *Cognition and concepts.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

Brooks, L. R.
Non-analytic correspondences and pattern in word identification. In J. Requin (Ed.), *Attention and Performance VII.* Hillsdale, N.J.: Lawrence Erlbaum Associates, in press.

Chall, J.
Learning to read: The great debate. New York: McGraw Hill, 1967.

Jeffrey, W. E., & Samuels, S. J.
Effect of method of reading training on initial learning and transfer. *Journal of Verbal Learning and Verbal Behavior,* 1967, *6,* 354-358.

(Baron and Strawson, 1976; Baron, submitted). Such differences might arise in at least two different ways. It might be that people who are good at using correspondences are those who are good at finding analogies in their memories. If so, practice at this skill might be useful. Or, on the other hand, it may be that the retrieval of similar words is automatic (through generalization), once the stimulus word is perceived in the right way. In this case, teaching people to "parse" a word correctly may help them generalize along appropriate dimensions. The third difficulty is that the transfer effects we found are small. If these effects are indicative of what we might expect outside the laboratory (although there is no reason to assume they are), explicit instruction might be faster.

The fourth argument against relying too heavily on the present results is that these results might tell us about mechanisms available to college students only; these mechanisms for transferring by analogy or generalization might not be available to young children (as the results of Jeffrey and Samuels, 1967, suggest). Children might be different for two reasons. First, they might be less prone to apply intentional strategies such as those involved in using analogies. Second, children might perceive printed or spoken words differently. If, for example, a child perceives spoken words more as wholes and less as strings of separate phonemes, (Treiman and Baron, submitted), the mechanisms of generalization and analogy, both of which require preservation of some phonemes and replacement of others, may not work as well.

correspondence condition and 1.21 sec for the control condition. This effect is slightly opposite to what we predicted. For these mean times, the difference in magnitude of transfer between the Experiments II and III was significant, $p<.025$. For memory, the mean scores were 28.9% for the correspondence condition and 27.5% for the control condition. This difference of 1.3% is to be compared to the difference of 13.8% in the last experiment, although the difference between experiments did not quite reach statistical significance. When only sets A and B were examined, memory was correct 37.5% of the time for the correspondence condition and 42.5% for the control condition, again, an effect slightly opposite to our prediction.

Finally, we can compare the individual phonemes for which the subject had two sources of transfer (such as "Z" in the first item in set A, for subjects who used list C2) with the letters for which he had no sources (such as "EE" in this same item). The Z could be inferred either from an association between the Z in ZAW and the first letter of the first item of list C2 or the second letter of the second item. There is no association from which EE could be inferred. For the letters with two sources of transfer, the subjects produced the correct phoneme 55.0% of the time, and for the letters with no sources, 52.5%. These figures are almost identical. Again, there is absolutely no evidence for transfer based on associations between letters and phonemes.

Discussion

We must acknowledge that these results may be restricted to these somewhat artificial situations. In addition, Experiments II and III might not be quite as comparable as we intended; there were, for example, even fewer responses to be learned in Experiment III than in Experiment II, and this may have affected the processes used. Still, the overall results of these experiments do point to analogy and generalization as the most likely mechanisms for transferring spelling-sound correspondences in the absence of knowledge of the existence of the correspondences. When similarity relations are controlled, as in Experiment III, no incidental learning based on transfer of the spelling-sound correspondences is found.

If it is true that analogy and generalization are important in using spelling-sound correspondences, we might want to consider the implications of this fact for teaching children to decode new words. One might think from the present findings that there is little to be done. It might appear that since rules may be abstracted from learning to read whole words, this is all that is required. However, there are four arguments against this. First, many studies have (Chall, 1967) shown that the whole-word method is often inferior to other methods in the classroom, and by our argument, it ought to be superior in all cases. Second, there do appear to be individual differences in the extent to which people can use spelling-sound correspondences

James G. Martin, Richard H. Meltzer, and Carol B. Mills

Visual Rhythms: Dynamic Text Display for Learning to Read a Second Language

Visible Language, XII 1 (Winter 1978), pp. 71-80
James Martin's address: Department of Psychology, University of Maryland, College Park, MD 20742.
0022-2224/78/0100-0071$02.00/0

A method is described in which sentences are presented in a dynamic visual display. A television monitor is used to present simultaneously the visual and auditory versions of a sentence, with each of its successive visual and auditory syllables yoked in parallel; the onset of each visual syllable is synchronized with the onset of each syllable as it is heard through the auditory channel. The result is a sentence which "grows" left-to-right across the screen, one syllable at a time, in "visual rhythm." In an experiment, the subjects were three groups of secondary-school students learning Spanish as a second language. In training sessions, the rhythmic group saw the sentences in "visual rhythm," the unrhythmic group saw the same sentences but in static visual display, and the control group had no exposure to either visual-auditory display. Before and after training, all groups provided pre-test and post-test measures of oral reading fluency. The dependent measure was pre-test to post-test relative change in judged reading fluency. The results favored the rhythmic group.

Introduction

This paper proposes a novel method for visual presentation of sentences and other language materials as a way of facilitating the teaching of reading. The essential idea is that a sentence is printed in a dynamic display that changes through time. The method makes use of a TV monitor to present simultaneously the visual and auditory versions of a sentence, with each of its successive visual and auditory syllables yoked in parallel. This is accomplished by synchronizing the visual onset of each syllable with the onset of each syllable as it is heard through the auditory channel (the syllable rather than the word is the more appropriate temporal unit of speech). The result is a sentence which "grows" left-to-right across the screen, one syllable at a time, in "visual rhythm." To illustrate, say the sentence, "The boy is in the boat," and imagine that each syllable (the, boy, is, . . .) appears on the screen just as you speak it, so that the complete sentence is on display when the last syllable is spoken. The sentence may be presented visually either by itself or in combination with the auditory channel. The display functions as a link between a sentence as it is displayed visually in space and as it is displayed auditorily in time. It presents dynamic information about language in the visual mode, and it displays this dynamic information in what is probably the simplest and more relevant form for most purposes, readable text.

Our first experiment using the method (Martin and Meltzer, 1976) was an attempt to assess the reactions of young readers to the novel visual display, and to assess its possible effects on samples of their oral reading. The subjects were children in a summer remedial reading program who were to be second, third, or fourth graders in the fall. They were divided into two training groups. In several training sessions, the Rhythmic group was exposed to sentences appearing syllable-by-syllable in "visual rhythm" as described above. In training sessions for the Unrhythmic group, the same sentences were presented in a static display, that is, the complete sentence appeared all at once rather than syllable-by-syllable. In either condition the sentence was presented sometimes on both visual and auditory channels, sometimes on the visual channel alone, and on some trials the child attempted to read the sentence out loud. Each child had three training sessions of approximately 10 minutes spaced over a two-week period. At the beginning of the first training session, the child read aloud several sentences which were tape-recorded, and the same sentences were read aloud by each child and tape-recorded during the last session; together these sentences provided a pre-test and post-test measure of reading fluency for each child.

After training sessions were completed, each of the recorded pre-test and post-test readings of a given sentence by a given child were paired together, but in random order with respect to whether the pre-test or post-test reading was the first member of the pair. Finally, all utterance pairs were assembled on a new audiotape. Six judges listened to all utterance pairs and indicated which item of the pair sounded more fluent. The results of the experiment were in favor of the Rhythmic group. A higher percentage of post-test readings of a pair were judged to be more fluent in the case of the Rhythmic group than in the case of the Unrhythmic group.

The results suggested that the method might have some promise, but much more research, of course, was needed. Of the many alternatives which might be taken as the next step, we tried the method with another relevant and accessible population, second-language learners. In the experiment reported in this paper, the subjects were secondary-school girls who were taking a course in beginning Spanish, and as in the earlier experiment, the effects of training sessions were assessed by changes in pre-test and post-test oral reading fluency. There was a Rhythmic group and an Unrhythmic group differing in terms of intervening audiovisual training materials, and an additional Control group who provided pre-test and post-test readings without intervening experimental training with the visual-auditory display.

Method

Materials

Training sentences. Materials for the 39 sentences used in the experiment were based on the textbook used by the subjects for their course in beginning Spanish, and were chosen to supplement their classroom work. Sentences were selected or adapted from the text using the following criteria: (a) they contained less than 37 characters and spaces and (b) they contained only English punctuation and characters. These criteria were used because of limitations of the computer screen to be used in preparing the training materials. Two examples of sentences are: *Juan pone la ropa en la maleta. Alvedor de la casa hay una finca.*

The sentences, which ranged from 4 to 16 syllables in length, were read with normal intonation and rate by a native female Spanish (Cuban) speaker, who was the teacher of the Spanish class. Sentences were recorded in a sound-treated room on an Ampex AG-500 tape deck.

Preparation of training materials. The time-voltage waveform of each sentence was digitized and displayed on a PDP-12 computer screen in an oscilloscope-type display so that syllable onsets could be measured. Using this timing information, the rhythmic version of the sentences could be prepared. For this version a TV display was generated by the computer such that each syllable appeared visually on the screen in synchrony with the auditory syllable. Each syllable appeared to the right of the preceding syllables until the entire sentence was displayed, after which the sentence remained on display for an additional 500 msec. A space appeared between each letter for greater clarity; triple spaces appeared between words.

Each rhythmic sentence had an unrhythmic counterpart. In this version the only difference was that all syllables in the entire sentence appeared on the screen simultaneously[2] and remained on the screen for the same length of time as the total time of the rhythmic sentence counterpart (visual first syllable onset to visual sentence offset). This total time ranged from about 1.2 to 3.0 sec.

Two parallel video tapes were constructed, one containing the rhythmic and the other containing the unrhythmic sentence versions. Each sentence was repeated five times sequentially on the tape with about 5 sec between repetitions and 6 sec between each set of 5 repetitions.

Pre-test and Post-test reading materials. These materials consisted of 16 sentences ranging from 7 to 24 syllables in length. Half of these were selected from the sentences displayed during training sessions (old), half were new. The new sentences contained familiar vocabulary but were otherwise unrelated to the training materials. The old and new sentences were mixed together and typed on a single sheet of paper.

Subjects and Procedure

The subjects were 45 girls enrolled in the second semester of a beginning Spanish course at their parochial secondary school, of which 14, 15, and 16 were in the Rhythmic, Unrhythmic, and Control groups, respectively. Seventeen additional girls participated in some portion of the experiment but were eliminated because they missed a training session or one of the reading fluency tests. Assignment of the subjects into the three groups was determined by scheduling convenience, but there were no significant differences in mean grades for the first semester of Spanish, and their teacher regarded the groups as alike in ability. All the girls were 9th or 10th graders with the exception of one girl who was in the 12th grade.

Prior to the training sessions, each subject was taken individually into a classroom and asked to read the 16 test sentences into a microphone for tape recording; these were the pre-test readings for the experiment. Each subject participated in five training sessions, each of which was 15 to 20 minutes long, in groups of from 5 to 16 subjects. The training sessions were held two times a week during regular Spanish class periods and spanned a 2-1/2 week interval.

The procedure during training session 1 for subjects in the Rhythmic and Unrhythmic groups was as follows. First the subjects arranged their chairs around the TV monitor (19" screen) so that everyone in the group could clearly see and hear the sentence. Then each sentence was presented for five repetitions and responded to (according to their instructions) as follows. On the first two repetitions they watched and listened. On the third and fourth repetitions, they attempted to say the sentence aloud in time with the speaker. On the fifth repetition they again watched and listened. This routine was repeated for 21 different sentences.

The instructions were changed slightly during training sessions 2 and 3 since the sentences had been used during the first session and were now familiar. On the first repetition of each sentence the subjects watched and listened; on the next four repetitions they spoke the sentences aloud. During training sessions 4 and 5 a different set of 18 sentences was used. The instructions for training session 4 were like those of training session 1, since the 18 sentences were unfamiliar. The instructions for training session 5, after the sentences had become familiar, were like those of training session 2.

The Control group had no training sessions with the visual-auditory display. Instead, all of the materials from the old sentences in various contexts were made familiar to this group as part of their normal classwork routine. The results from this group provide a baseline for comparing changes over the time of the experiment.

After the completion of all training sessions, the 16 test sentences were read by each subject and tape-recorded for use as a post-test measure under identical conditions with the pre-test reading.

Listener Judgments of Reading Fluency

To reduce the size of the judgment task for the judges, only half of the 16 pre-test and post-test sentence pairs per subject were used. Four of these were old sentences, i.e., used during training; four were new sentences, ranging from 7 to 14 syllables. For each of the 45 subjects, each of the eight sentences recorded during the pre-test was paired with the identical sentence recorded during post-test, making a total of 358 pairs (360 minus 2 pairs from one subject which were missing because of inadvertent post-test recording mishap). Half of the 358 pairs were placed in the pre-test and post-test order in which they had been read, half were the reverse. All pairs were assembled on a new audiotape with the following constraints. Pairs were arranged in blocks of 45, of which approximately half (either 22 or 23) were reversed pairs. Each block contained one pair read by each of the 45 subjects. Each of the eight different test sentences read by at least four subjects was included in each block.

There were four judges, of which two were college-level Spanish instructors and two were college students whose native language was Spanish. The judges were informed of the randomized nature of the tape and were asked to make the fluency judgments on the basis of how smoothly sentences were spoken as a whole, and not on the basis of the pronunciation of individual words. They followed a typed manuscript of the sentences while listening to the sentence pairs and judged which utterance of the pair was more fluent. They listened to each pair as often as necessary. After choosing one of the pair as more fluent than the other, each judge then gave a confidence rating for that choice on a five-point scale.

Results

A judgment that the post-test reading of a given sentence pair by a given subject is more fluent indicates improvement over the time of the experiment, and percentages and mean confidence ratings across subjects permit a comparison of relative improvement among the three training groups in the reading of both old and new sentences. Table I gives the summary data. The results are quite similar for direction of judgment (percentages) and mean confidence ratings. The result of interest is the relatively greater improvement by the Rhythmic group in reading aloud the new sentences during the post-test.

The percentages were tested by summing the number of pair judgments in favor of the post-test reading for each subject and each judge. With four sentence pairs this value could range from zero to four. These sums were submitted to analysis of variance with sentence type (old/new) as a within-subjects factor and training condition (Rhythmic/Unrhythmic/Control) and judge as between-subjects factors. The following results were significant: sentence type,

Table I

Percent Judgments and Mean Confidence Ratings
(Ratings in Parentheses)

Sentences	Group: Rhythmic	Unrhythmic	Control	All
Old	79.5%	80.0%	64.5%	74.7%
	(2.58)	(2.54)	(1.36)	(2.16)
New	73.1%	59.2%	63.7%	65.3%
	(1.71)	(0.80)	(o.97)	(1.16)
All	76.3%	69.6%	64.1%	70.0%
	(2.15)	(1.67)	(1.17)	(1.66)

F (1,168)=16.80, p<.01; training condition, F (2,168)=5.85, p<.01; and the interaction of sentence type and training condition, F (2,168)= 6.16, p<.01.

Mean confidence ratings were tested by assigning negative sign to judgments in favor of pre-test rather than post-test readings and averaging across the four sentences; these values could range from −5 to +5 with positive values in favor of the post-test readings. Analysis of variance yielded the following significant factors: sentence type, F (1,168)=27.30, p<.01; training condition, F (2,168)=7.93, p<.01; judge, F (3,168)=4.34, p<.01; and interaction between sentence type and training condition, F (2,168)=4.14, p<.025.

Separate analyses of the Rhythmic vs Unrhythmic conditions alone yielded significant differences also, as did an analysis using sentences rather than subjects as the sampling unit.

Discussion

The relatively greater improvement by Rhythmic and Unrhythmic groups over the Control group can be attributed to practice in pronouncing sentences aloud with the help of the visual-auditory display, and relatively greater improvement (Rhythmic and Unrhythmic groups) in old sentences over new sentences can be attributed to practice with these particular sentences. The difference favoring the Rhythmic group over the Unrhythmic group in the case of new sentences suggests greater transfer to fresh sentences after practice with the rhythmic display. This result generalizes the earlier result (Martin and Meltzer, 1976) in favor of the visual rhythm over the static display.

Speculating on the reasons for differential transfer effects from the rhythmic and static display, little can be said on the basis of observations of individual behavior during the training sessions since the subjects were run in groups. But it is plausible to suggest that since the rhythmic visual display duplicates timing information in the auditory channel, it helped the subject to know when she was successfully shadowing the auditory channel and when she was lagging behind. The dynamic display also tends to convey a sense of rapid speaking rate, and this is another reason why it may be an effective pacer for oral reading practice, as it appears to be. Of course these considerations predict greater improvement for not only new but old

Sounds and symbols are paired dynamically so that many-to-one and one-to-many correspondences between visual and auditory modes may be noted. In pluralization, for instance, the "s" in "caps" and in "cabs" is seen as the same in both cases in visual display, but is heard on the auditory channel as /s/ in the former but /z/ in the latter. When "cake" is pluralized we see and hear the syllable ending in /s/. But when "horse" is pluralized, we hear it as two syllables, the latter of which is /Iz/, and we also simultaneously *see* it as two syllables since they appear on the screen temporally separated. Examples could be multiplied. While some aspects of orthography are handled more naturally than others, the display might facilitate the concurrent part and whole learning in a rather efficient way. The display is quite flexible in its technical aspects; its usefulness can be evaluated in other ways than by oral fluency, and it can be used with various populations, including the deaf who are known to have particularly difficult problems in learning to read (e.g., Odom, 1974). It is possible that any population with reading difficulties might profit from experience with the dynamic display.

In conclusion, our work with the visual rhythms display is an attempt to bring some of our ideas about speech perception (Martin, 1972; Meltzer, Martin, Mills, Imhoff and Zohar, 1976) to bear on the reading process by focussing on dynamic aspects of the stimulus. It is not clear that this work can be used to choose among sentences also, in a comparison between Rhythmic and Unrhythmic groups. But old sentences were seen and spoken repeatedly (10-15 times each) during training sessions by both groups. This much familiarity with the sentences could easily provide a ceiling on the advantages of a rhythmic over a static display. Additionally, one might even hypothesize a negative transfer component to old sentences in the Rhythmic group which could have partially cancelled positive transfer effects, since for this group these sentences were in dynamic display during training but static display during post-test.

The studies we have done using oral fluency tend to emphasize holistic aspects of learning to read, but it is possible that the visual rhythms display will be useful for the learning of part aspects also, for instance, the learning of sound-symbol correspondences. Learning to process written language is often assumed to be more difficult than learning to process spoken language (Kavanagh and Mattingly, 1972). One likely bottleneck is that the beginning reader must first learn to decode the parts (letters, words) of a sentence rapidly, or else the earlier items in a sentence are lost from memory before they can be integrated (Gough, 1972). In other words, parts must be mastered before wholes can be grasped, in learning to read a sentence. A potential advantage of the visual rhythms display is that part and whole aspects of learning to read may procede concurrently. Consider the learning of sound-symbol correspondences.

competing models of the skilled reader (Massaro, 1975), or that it can illuminate much of what takes place during any of the various kinds of skilled reading (Gibson and Levin, 1975). But the outlook for the study of learning to read seems more promising. As Massaro characterized Gough's (1972) view of the situation: "When a child begins to learn to read, he already knows how to listen. If somehow the printed symbols on the page could be made to speak, all that the young reader would have to do is listen" (Massaro, 1975, p. 260). We can almost agree with this, without having to believe that it comes anywhere near to the whole story.

1 This research was supported in part by a grant from the Army Research Institute for the Behavioral and Social Sciences and by a grant from the Bio-Medical Sciences Support Committee, University of Maryland, to the Center for Language and Cognition. Computing services were provided by the Computer Science Center of the University of Maryland. Thanks are expressed to Maria Torres, Clara Porcel, and the rest of the faculty of Regina High School, Hyattsville, Maryland, for their advice and their help in providing subjects for the experiment. Robin Jeffries was the research assistant for this project. Requests for reprints should be sent to James G. Martin, Psychology Department, University of Maryland, College Park, MD 20742.

2 This display was not quite simultaneous; computer hardware limitations required that each letter in the display appear sequentially, in this case at 8 msec intervals. Hence while letters within, say, a three-letter syllable appeared to be simultaneous in the Rhythmic version, the sequencing was slightly noticeable in the Unrhythmic version, since a sentence containing; e.g., 18 letters and spaces took 144 msec before the display was complete. In this case, the effects were as though the letters appeared across the screen in a "burst."

3 The criteria for judgments although vague are apparently not biased in favor of any particular group. However, future research should include separate judgments of timing, inflection, word pronunciation, etc.

References

Gibson, E. J., & Levin, H.
The Psychology of Reading.
Cambridge, Mass.: MIT Press, 1975.

Gough, P. B.
One second of reading. In J. F. Kavanagh and I. G. Mattingly (Eds.), *Language by ear and eye, the relationships between speech and reading.*
Cambridge: MIT Press, 1972.

Kavanagh, J. F., & Mattingly, I. G. (Eds.),
Language by Ear and by Eye.
Cambridge: MIT Press, 1972.

Martin, J. G.
Rhythmic (hierarchical) versus serial structure in speech and other behavior.
Psychological Review,
79, 487-509, 1972.

Martin, J. G., & Meltzer, R. H.
Visual rhythms: Report on a method for facilitating the teaching of reading.
Journal of Reading Behavior,
1976, *8,* 153-160.

Massaro, D. W.
Primary and secondary recognition in reading. In D. Massaro (Ed.), *Understanding Language: An information-processing analysis of speech perception, reading and psycholinguistics.*
New York: Academic Press, 1975.

Meltzer, R. H.; Martin, J. G.; Mills, C. B.; Imhoff, D., & Zohar, D.
Reaction time to temporally-displaced phoneme targets in continuous speech.
Journal of Experimental Psychology,
1976, *2,* 277-290.

Odom, P. B.
Some speculations concerning why the deaf cannot read well. Paper presented to SRCD Conference, New Orleans, August, 1974.

Speech Analysis During Sentence Processing: Reading and Listening

The present paper is concerned with the role of speech recoding during reading. Specifically, it examines information processing with respect to where reading and listening might come to share common mechanisms during comprehension. The paper is divided into four sections. The first section contains a review of evidence related to the issue of whether speech recoding is necessary prior to lexical access. The weight is against this view. The second section of the paper explores an alternative view—namely that speech recoding occurs in working memory, where word units are held in a speech form until comprehension of phrases or sentences occurs. Section III describes an experiment which shows that disrupting wording information in memory does not lead to semantic comprehension failure. These results suggest that reading does not occur by converting visual signals into a speech code until comprehension occurs. Finally, the general discussion centers on models of visual language processing.

The relationship between auditory and visual forms of language processing has received recent attention in both short-term memory and reading research. In memory, theorists were confronted with the problem of explaining modality differences in recall from short-term memory. In reading, interest lay in where in information processing auditory and visual forms of language come to share common comprehension mechanisms. The main focus of considerable research in both fields was on the question of whether visually presented language is first converted to a speech code before accessing meaning or being understood. That is, does reading simply consist of an additional grapheme to phoneme conversion stage added to the mechanisms used for speech perception? Such a view of reading seems reasonable, since beginning readers are already fluent speakers who could presumably bring their linguistic knowledge to bear on the problem of learning to read.

The main focus of the present paper is on where speech recoding occurs during reading. The first section of the paper examines evidence to support the view that visually presented language is converted to a speech form in order to gain access to meaning representations held in memory. This discussion is followed by a summary of

Visible Language, XII 1 (Winter 1978), pp. 81-101.
Author's address: Department of Psychology, McMaster University, Hamilton, Ontario L8S 4K1, Canada.
0022-2224/78/0100-0081$02.00/0

more recent findings which suggest that while speech recoding is often observed during visual word identification, it is an optional not a necessary stage prior to lexical access. The second section of the paper deals with an alternative view of where speech recoding of visual language might occur. The idea here is that while individual words can be understood without speech recoding, a speech code is used to keep the words of phrases or whole sentences available in working memory until parsing and integration mechanisms have abstracted the sentence gist. Reading and listening then become identical processes after the working memory stage of analysis. The third section of the paper contains an experiment which examines the relationship between speech recoding and meaning analyses during reading. Its purpose is to demonstrate that meaning analyses during reading can occur without speech recoding in working memory. Rather, speech recoding provides a more detailed record of the surface structure of sentences and may be useful in performing some reading tasks. With evidence on speech recoding during reading, the final discussion centers on models of reading and how these are constrained.

The evidence discussed comes from a variety of visual processing tasks, and no attempt has been made to categorize these according to their similarity to "real" reading. Also, speech recoding has been used as a general term encompassing all forms of speech based codes (acoustic, articulatory, auditory imagery, etc.). Since little evidence is available on which of these is the most appropriate level of analysis, use of a neutral term avoids prejudging this issue.

Speech Recoding for Lexical Access

In studies of short-term memory and of word recognition, some evidence is available which supports the view that speech analysis precedes meaning access, even for visually presented language. Some early short-term memory theorists expressed this view by postulating a phonemic short-term memory which preceded long-term semantic memory (e.g., Baddeley, 1966a,b). The idea is that language is first analyzed in terms of its speech characteristics before making contact with the meaning representations held in long-term memory. It is necessary to convert visually presented language to a phonemic form to ensure storage in short-term memory. Evidence supporting this view came mainly from studies of modality differences in retention. As early as 1960, Sperling had demonstrated that the recall of visual information from a spatial display was markedly reduced only 500 msec after its presentation. This observation led to the conclusion that visually encoded language fades rapidly from memory. It was also observed that errors made in the short-term recall of visually presented letters were remarkably similar to the perceptual errors made in identifying auditory letters presented in noise (Conrad, 1964). Together, these findings suggest that since visual representations are rapidly lost, items are encoded phonemically to be held in short-term memory.

This view of short-term memory as a speech store was further reinforced by the finding that only acoustic similarity caused a recall decrement in short-term memory, while only semantic similarity affected long-term recall (Baddeley, 1966a,b). Also, final list items are better retained when the lists are presented auditorily rather than visually (Murdock, 1967). An argument was made that while auditory items were compatible with the phonemic code of short-term memory, visual items required additional processing to access this store. Finally, when speech recoding was suppressed by asking subjects to repeat irrelevant speech during list presentation, serial recall was markedly reduced for visually presented lists but not for auditorily presented lists (Levy, 1971; Peterson & Johnson, 1971). Here again it appears that visually presented material is dependent on translation to a phonemic form to ensure short-term storage. Auditory processing, on the other hand, appears to directly access short-term memory. Since the accepted view was that short-term memory preceded long-term memory, these modality differences were consistent with the view that visually presented language is first converted to a speech code in order to access meaning.

While list learning research sometimes appears to be quite irrelevant to "real" language comprehension, a set of ideas and observations which closely parallel those developed above can be found in the word identification and reading literature. Since recent reviews (Barron, note 1; Bradshaw, 1975; Massaro, 1975) have summarized the available data in some detail, only the main findings need to be summarized here. Considerable evidence has been amassed which shows a syllable effect in visual recognition tasks. That is, two- (or multi-) syllable words are named more slowly than single syllable words (Eriksen, Pollack & Montague, 1970); same/different judgments are slower for multisyllabic than single syllable items (Klapp, 1971); and letter detection is more accurate in one than two syllable words (Spoehr & Smith, 1973). Since the syllable effect is obtained for words equated for visual length, the effect can be taken to indicate translation into a phonemic form during the visual recognition process. However, since the syllable effect does not occur in all visual recognition situations, the phonemic translation stage may be an optional not a necessary stage of processing (Green & Shallice, 1976; Johnson, 1975; Klapp, Anderson & Berrian, 1973). Also, the syllable effect may be related to orthographic regularity rather than phonemic encoding (Massaro, 1975).

A further source of evidence cited in favor of speech recoding during visual processing is that in lexical decision tasks it takes longer to decide that a letter string is *not* a word when that string sounds like a real word (e.g., brume) than when it does not (Meyer & Ruddy, note 2; Rubenstein, Lewis & Rubenstein, 1971). This is presumably because the letter string has been converted to a phonemic form that can then be confused with the real word. However, when Forster & Chambers (1973) compared naming times and lexical decision times for words, non-words, and unfamiliar words, they found that naming times were shorter for words than non-words, and for high frequency than low frequency words. These findings suggest that lexical access is achieved before naming, as the existence of a lexical entry or word frequency should not influence naming if it precedes lexical access. Further, while naming and lexical decision times were correlated for words, they were not correlated for nonwords, indicating that word naming did not occur via the same grapheme-phoneme translation route used for pronouncing nonwords. While lexical decisions sometimes may appear to involve phonemic recoding, it is not altogether clear that this encoding *precedes* lexical access. The Meyer & Ruddy (note 2) and the Rubenstein, et al. (1971) experiments do not, therefore, provide clearcut evidence that visual signals are converted to a speech format, with comprehension following the speech perception route.

A similar interpretative limitation applies to the short-term memory research reviewed earlier. While the evidence described suggests that visual language is sometimes converted to a speech code during processing, it cannot be taken to support the stronger claim that such recoding is either necessary or prior to meaning analysis. A striking demonstration that these stronger claims would be unwarranted came from an experiment by Kroll, Parks, Parkinson, Beiber & Johnson (1970). These investigators presented single memory letters, either visually or auditorily, during an ongoing auditory shadowing task (used to prevent speech translation). The results showed a marked superiority in retention of visual over auditory memory letters, even 30 seconds after presentation. Kroll, et al. (1970) concluded that preventing speech translation led subjects to maintain the items visually, thus uncovering a capacity normally hidden by the tendency to name visually presented language.

On the basis of subsequent work, Kroll (1975) argued that visually presented language can be encoded and rehearsed visually without recourse to an auditory translation step. However, he also noted that subjects often form word associates to the memory letters, thus converting the memory set to a semantic code. This strategy is particularly prevalent when the shadowing task is difficult. A series of experiments in my laboratory (Levy & Brevik, note 3) confirmed this use of a semantic strategy in the shadowing paradigm. Subjects reported thinking of words that began with the memory letter as aids to their retention, a strategy that proved easier for visual than for auditory memory letters. Controlling this strategy, by asking subjects to generate words to the shadow letters as well as the memory letters, led to a loss of the visual memory advantage, suggesting that the visual to semantic strategy is quite important in maintaining the high levels of visual recall during shadowing. The important point, of course, is that visual to semantic conversions occurred without an intervening phonemic step, since visual like auditory recall would have been hindered by shadowing if speech recoding had occurred. Thus while earlier work in short-term memory suggested that visually presented items were translated to a phonemic form for further processing, it appears that under other circumstances this phonemic stage can be bypassed. In general, the strict phonemic short-term memory to semantic long-term memory view has been abandoned in favor of more flexible processing views, as evidence accumulated that was inconsistent with this simple model (e.g., Craik & Lockhart, 1972; Baddeley & Hitch, 1974).

Similarly in the reading literature, when demonstrations more directly addressed the issue of the *necessity* of phonemic recoding on route to meaning, the results were consistently negative. Bower (1970) asked speakers of Greek to read passages which contained misspellings that were pronounced exactly the same as the correct spellings. This was accomplished by interchanging vowels that were pronounced identically but spelled differently. The Greek readers were considerably slowed down by this visual distortion, suggesting that their normal reading must be via some route disrupted by the visual change. Obviously the grapheme to phoneme route was still available and undistorted (though perhaps it was less familiar) indicating that it was not the route used during rapid reading. A similar study using English words and speakers has been reported by Theios & Muise (1976), and again misspellings which maintained correct phonemic translations led to slower reading times.

In a similar vein, Baron (1973) demonstrated that subjects had no more difficulty in deciding that a phrase was nonsense when it sounded sensible, than when it didn't. That is, they could classify the phrase, *tie the not,* as nonsense, as quickly as the phrase, *I am kill.* One might have expected the phonemic correctness of the first phrase to slow down rejection time if phonemic translation had occurred. Also, Kleiman (1975) found little effect of preventing phonemic translation (by using an auditory subsidiary task) when subjects made synonymity judgments about visually presented word pairs. Since the synonymity judgments required access to meaning, this result suggests that meaning access does not depend on phonemic translation. Finally, Green and Shallice (1976) demonstrated that misspelling delayed meaning decisions more than phonological ones. If reading consisted of a visual to auditory to semantic route, then visual errors (misspellings) should influence the first but not the second stage, and misspelling should affect phonological and meaning judgments equally. Green and Shallice argued that semantic judgments were made directly from the visual display without intervening phonemic recoding.

A further source of evidence suggesting that comprehension can occur through routes other than speech recoding is the observation that phonemic dyslexic patients make errors during readings that are of a visual and/or semantic nature, but rarely make phonemic errors (Marshall & Newcombe, 1966; Patterson & Marcel, 1977; Shallice & Warrington, 1975). Further, these patients are unable to read aloud orthographically regular nonwords which they can repeat when presented auditorily, and they are not slowed down in rejecting nonwords in a lexical decision task when the nonwords sound like real words (Patterson & Marcel, 1977). These findings suggest that grapheme to phoneme conversion is not possible for these patients, yet they can read and make judgments of wordness from visual displays. Clearly a phonemic translation stage is not necessary for such visual tasks. Further, Allport (1976) has produced similar results using a masking procedure with normal readers. He noted that when tachistoscopically presented words were masked so that they could not be correctly identified, the errors were sometimes semantically related to the target word even when the error and target words shared no visual or phonemic features. The argument here was that meaning appears to have been analysed with no evidence of a phonemic mediation stage, since phonemic errors would also have been expected.

Looking at all of this evidence together, it seems reasonable to conclude that while speech recoding may occur, visually presented words can be understood without first being encoded into a phonemic form. That is, speech recoding is not necessary *prior* to lexical access.

Speech Recoding in Working Memory

While the current weight of evidence is against the view that speech recoding is needed for lexical access, some recent work suggests that it does play an important role in comprehension. The claim made is that speech recoding acts as a holding device to keep word units available in memory until phrase or sentence comprehension occurs (e.g., Conrad, 1972; Norman, 1972). Kleiman (1975) pointed out that while speech recoding is not needed for lexical access of single items, it does seem to be important in the comprehension of complete sentences. Kleiman suggested that words are converted to a speech form to be held in working (or short-term) memory until the sentence message could be grasped. The idea was that while a visual route is capable of handling a few single items, its capacity is overtaxed by the requirement of holding phrases or sentences until these units could be integrated to yield a semantic interpretation of the whole string. Speech recoding, then, acts as an overflow device to aid comprehension of long or difficult passages.

Some evidence can be marshalled in support of this view. Hardyck & Petrinovich (1970) trained subjects to control their subvocal activity during reading by using a feedback signal to indicate when EMG activity went above an acceptable baseline level. They found that during easy reading, college students were able to suppress their subvocal behavior with no cost to comprehension. However, when reading more difficult passages these students were unable to suppress their subvocalizations, and attempts to do so led to a comprehension loss. Hardyck & Petrinovich suggested that phonemic mediation may aid comprehension during difficult reading, when perhaps the memory load is greater.

Further evidence for the working memory position came from studies of vocal activity during reading. Kleiman (1975) asked subjects to shadow auditorily presented letters while making judgments about pairs of words presented in a visual display. The shadowing task was used to prevent speech recoding of the visual items. Three types of judgments about word pairs were made: did the two words of each pair rhyme (phonemic), look similar (graphemic) or mean the same thing (synonymity)? Kleiman found a larger decrement due to shadowing when subjects made phonemic than when they made either graphemic or synonymity judgments. From this he concluded that analyzing the meanings of word pairs does not require speech recoding, since synonymity judgments require meaning access but are little affected by shadowing. In contrast, a subsequent experiment demonstrated that judgments about the semantic acceptability of entire sentences were quite adversely affected by shadowing. Kleiman interpreted this decrement as showing that sentence acceptability judgments, unlike simple synonymity decisions, requires the simultaneous evaluation and integration of several items, and speech recoding is required to keep the word units available for this more extensive processing. From these data Kleiman assigned speech conversion of visual signals to the working memory rather than the lexical access stage of analysis.

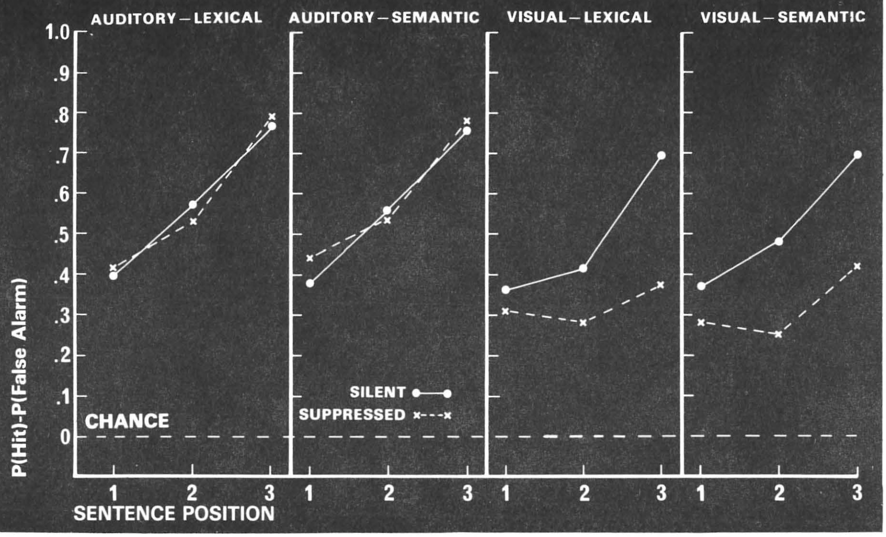
According to Kleiman's model, comprehension occurs via a direct visual to meaning route when the information load is small, but when visual capacity is taxed speech recoding occurs to provide temporary holding until parsing and integration processes comprehend the incoming messages. These ideas are reasonably compatible with the data reviewed in this paper. The short-term memory research suggested that in memory tasks where word holding is probably encouraged, subjects convert visually presented language to a phonemic form for immediate storage. The word recognition studies, on the other hand, probably involve information loads too small to tax the visual analysis route, thus not envoking a speech recoding need. The point to be taken here is that while other forms of processing may be possible for visually presented language, the speech recoding route does operate as an aid to *fluent* reading when the information load is heavy or when learning is involved. Similar arguments regarding the role of a memory requirement in explaining speech recoding effects during reading have been made by Baron (1976).

It is in this memory framework that my own experiments on reading and listening best fit. The ideas behind the work stem rather directly from my earlier observation in short-term memory studies that suppressing speech recoding (by asking subjects to repeat an irrelevant syllable during list presentation) adversely affected the serial recall of visually presented lists, while not affecting recall of auditory lists (Levy, 1971). The results appear to suggest that visually presented words are more dependent on active speech recoding than are auditorily presented words, perhaps because auditory items are already in a suitable speech code. In order to generalize these findings to reading, a change detection paradigm (Sachs, 1967) was used. Since two earlier reports (Levy, 1975, 1977) describe in detail some experiments on comparisons of memory for visually and auditorily presented sentences, only the basic findings will be summarized here.

To address the distinction between a reading deficit and a general language comprehension deficit, we followed the route described by Gleitman and Rozin (1977). These authors defined fluent reading as "the skill of extracting meaning from print to the same degree that one extracts it from the sound stream." The idea was that if a message can be processed auditorily, then any deficit in visual processing of the same message must be a reading, not a general linguistic problem. Therefore, we used modality manipulations to evaluate the visual specificity of the effect to be described. The paradigm used was a simple one. Subjects read or heard sets of three unrelated sentences, all of the form: article, adjective, noun, verb, article, adjective, noun. The sentences were presented serially, at a rate of one sentence/three seconds. Subjects either read (or listened) silently, or they counted from one to ten, quickly and continuously during the sentence presentations (called suppression). A test sentence followed each set of three sentences and subjects indicated whether the test sentence was identical or changed from one of the presented sentences. In fact there were two types of changes, lexical and semantic, but subjects were not required to indicate the type of change made. For lexically changed sentences, a synonym was substituted for one of the nouns, thus changing the wording while leaving the meaning unaltered. For example, if the original sentence was, *The attractive man liked the passing girl,* a lexical change might be, *The attractive fellow liked the passing girl.* A semantic change consisted of switching the subject and object nouns, thus maintaining the original wording and syntactic structure of the sentence while altering its meaning. For the above example, the semantic test would be, *The attractive girl liked the passing man.*

While with unpractised subjects a small inconsistent decrement due to counting was observed for auditorily presented sentences (Levy, 1975), the modality speci-

Fig. 1 Lexical and semantic detection performance in four modality-vocalization conditions (from Levy, 1977).



ficity of the suppression decrement became quite striking when practised subjects were used. As Figure 1 (from Levy, 1977) clearly shows, both lexical and semantic detection were adversely affected by the counting requirement when the sentences were presented visually, but not when they were presented auditorily. The data are presented as P(hit) minus P(false alarm) to correct for response bias (analyses using d' scores yielded the same results). From these data it was concluded that visual processing is more dependent on translation to a speech code than is auditory processing, perhaps because the auditorily presented sentences were already in a speech format appropriate for comprehension.

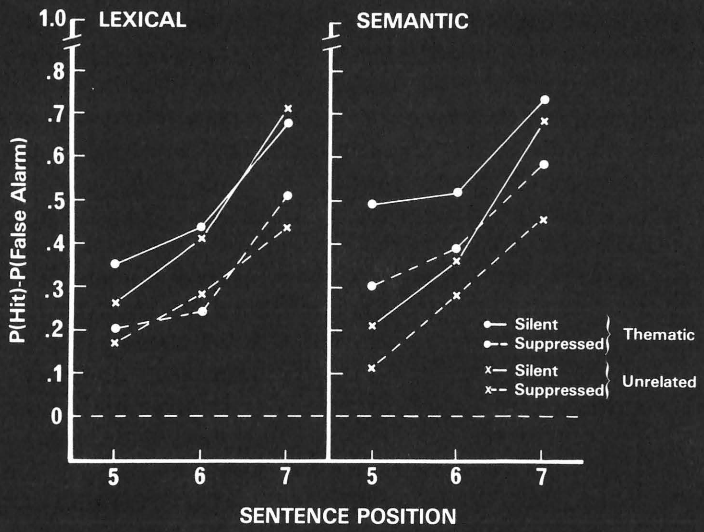
On attempting to relate further the speech suppression effect to comprehension, however, a problem became apparent. Our earlier studies had used sets of unrelated sentences, perhaps leading subjects to adopt a literal rather than semantic reading strategy. The apparent speech dependency may then have been caused by subjects not reading for meaning. This view was reinforced by the similarity of effects in the lexical and semantic measures. The question addressed, therefore, was whether comprehension of a passage's meaning was affected by the suppression task. We first wanted independent evidence that the semantic measure uniquely reflected variations in meaningfulness. Once this had been established, we could then ask whether suppressing speech responding was less detrimental for "meaningful" as opposed to more literal forms of reading.

To test these notions we constructed a number of seven-sentence passages that addressed a central theme. While the sentences were all of the form described earlier, the passages did form a story. The subjects were instructed to try to relate the sentence meanings since this would help them to remember the whole set. We also provided titles for each story, making sure that none of the title words also appeared in the tested sentences. One group of subjects read half of these stories silently and half while counting. Another group of subjects read unrelated sets, half silently and half while counting. The unrelated sets had tested sentences identical to those of the thematic passages, but the six context sentences of each passage were randomly distributed across unrelated sets so that no meaningful relationships existed within sets. Lexical and semantic tests for the final three sentences in the sets were given (additional but unanalysed tests of earlier sentences ensured that all seven sentences were read).

The left-hand panel of Figure 2 represents the detection of wording changes when the meaning remains consistent between presented and tested versions. While suppression of speech responding had a large detrimental effect on word memory for both thematic and unrelated material, there was no difference in retention for the thematic and unrelated passages. Since thematicity is a semantic manipulation, one would not expect such a variable to influence the choice between synonyms. The right-hand panel presents the detection of changes in meaning when wording and syntax were constant. The picture is quite different. Now there is a large effect of thematicity with semantic changes being better detected in thematic than in unrelated passages. Also, there is a large decrement due to suppression. Importantly, there was no interaction between these two variables ($F < 1$). That is, when subjects read for meaning as indicated by the facilitating effect of thematicity, they still were dependent on speech recoding as indicated by the suppression decrement. Performance reflected independent contributions of both meaning and speech analyses.

What does this tell us about the role of speech recoding in reading? The main point is that analysis of meaning appears to proceed in spite of speech disruption. The effect of thematicity was as large in the suppressed condition as it was in the silent condition. However, detection of a semantic change was also affected by suppression. The problem, then, is to determine those aspects of processing that are affected by suppression and those that are independently aided by thematicity. One possible explanation arises from the nature of the semantic detection task itself. Since lexical, semantic, and identical tests occurred randomly in the sequence of trials, subjects were always required to maintain wording information. Thus the suppression decrement may be due to poorer surface structure information following counting, while the thematicity effect may be due to enhanced meaning recall for thematic passages. Since the semantic detection task could have been solved from either surface structure information (noting that the word order dif-

Fig. 2 Lexical and semantic detection performance for thematic and unrelated passages (from Levy, 1977).



fered), or from deeper semantic knowledge (gist recall), both sources of information may have contributed to overall performance in this task. That is, thematicity may have aided recognition of a general idea in the passage, while surface structure memory aided in determining such details as, it was the girl who kissed the boy and not the reverse relationship.

The experiment to be reported here attempted to test this "surface structure versus meaning" explanation of the independent suppression and thematic effects. If the suppression decrement found in the subject-object reversal task was due to loss of word order information, then a semantic task which depended solely on gist, with wording information being uninformative, should fail to show the suppression decrement. That is, for tasks that tap only semantic information and that cannot be solved from the surface structure information, a suppression decrement should fail to occur if meaning analyses can proceed independently of the speech encoding.

The experiment to be described resembled our earlier thematic study with subjects reading either thematically related or unrelated sentences. One group of subjects performed the lexical-semantic detection task in an attempt to replicate the independent thematic and suppression effects (Levy, 1977, experiment III). A second group of subjects read the same passages, but their task was to detect paraphrases. For the latter subjects no test sentences were identical in wording to the presented sentences, and their task was to detect changes which altered sentence meaning, as opposed to changes which maintained sentence meaning. We hypothesized that since wording was altered on every trial and therefore could not aid in paraphrase detection, subjects would not maintain wording information, and therefore a suppression decrement would not be observed in this task. For the lexical-semantic subjects we expected independent effects of thematicity and of suppression, but for paraphrase subjects only a thematic effect was expected.

Experiment

Method

Materials. The paragraphs used by Levy (1977, experiment III) were used here, except that in half of them the tested sentence was moved to positions two, three, and four. This change allowed us to evaluate memory over a longer interval so that the memorial duration of the suppression and thematic effects could be observed. Context sentences were then revised to maintain the paragraph's continuity and thematicity. Each paragraph therefore contained one to-be-tested sentence surrounded by six context sentences which formed a story. Sentences were of the form: article, adjective, noun, verb, article, adjective, noun, although some minor modifications were allowed so that the paragraphs were more linguistically natural. Each paragraph was given a title that reflected its main idea. Table I contains examples of the paragraphs and both types of tests used. Matched unrelated sets were formed by keeping the tested sentences in the positions they held in the paragraphs, but randomly distributing the six context sentences across sets. No thematic relationship existed among the seven sentences of an unrelated set and no titles were given.

Two sets of thematic and a matched two sets of unrelated passages were constructed with each set containing 96 critical experimental passages plus 24 filler passages. The filler passages were used only for the lexical-semantic detection subjects and contained changes in a sentence verb or adjective. Their purpose was to ensure that lexical-semantic subjects attended to all of the sentence words, not just to the nouns. The 96 critical passages were identical for both lexical-semantic and paraphrase subjects. The same sentence per passage was tested for both groups, though the tested versions differed. All subjects participated in two experimental sessions, with material sets balanced across sessions for both groups of subjects.

Design. Three main variables were studied for both paraphrase and lexical-semantic subject groups. For both task groups half of the subjects read thematically related passages and the other half read sets of unrelated sentences. Thematicity was a between-subjects variable. Sentence position and silent reading vs counting were varied within-subjects. Sentences two to seven were tested equally often and randomly, balanced across all other conditions. Also for each subject, half of the passages were read silently and half while counting (from one to ten quickly and continuously). The order of silent and suppressed (counting) blocks was

Table I

Examples of paragraphs and tests used in Experiments I and II.

Bad Manners	A young painter flirted with passing girls.
	The flippant artist amused the startled ladies.*
	But art galleries aren't for such activities.
	The exhibit guard grasped the artist's shoulder.
	His gruff voice contained a clear message.
	Flirtatious behavior is unbecoming to an artist.
	The playful youth resumed a demure posture.
Lexical Change:	The flippant artist amused the startled women.
Semantic Change:	The flippant ladies amused the startled artist.
Paraphrase - Yes:	The witty artist amused the startled women.
Paraphrase - No:	The obnoxious artist amused the startled critics.
A Hospital Visit	The elderly chap entered the bustling hospital.
	Taking the elevator he found the room.
	The lifeless patient nodded a silent hello.
	The bedside clock ticked the long hours.
	The visitor's chatting cheered the sick patient.
	The two men recounted their past adventures.
	The dying man saddened the visiting relative.*
Lexical Change:	The dying fellow saddened the visiting relative.
Semantic Change:	The dying relative saddened the visiting man.
Paraphrase - Yes:	The dying fellow distressed the visiting relative.
Paraphrase - No:	The dying doctor thanked the visiting relative.

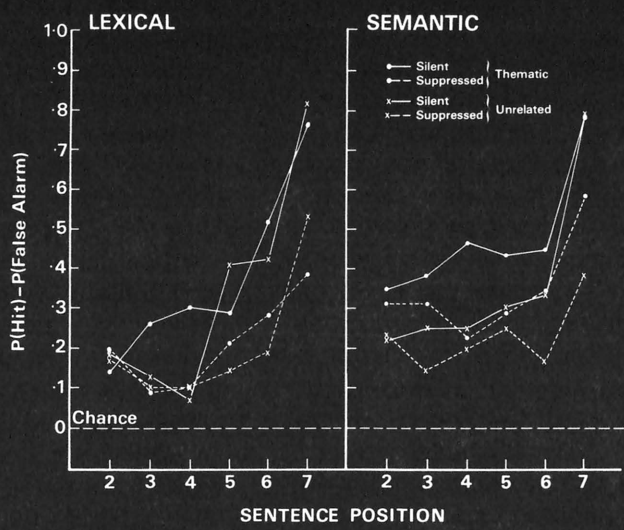
counterbalanced across subjects, such that all materials and test types were tested equally often with counting and with silent reading.

Tasks. While the two main subject groups read the same passages under the same vocalization condition, their task demands were quite different. *Lexical-semantic* subjects attempted to detect changes that occurred in the tested sentences. Within each session half of the tested sentences were in fact identical and half were changed. Again, both *lexical* changes, which consisted of a synonym substitution for one of the nouns (thus maintaining meaning), and *semantic* changes, where the subject and object nouns were interchanged (thus altering meaning, but not wording or structure) were used. Subjects responded "changed" in both cases without attempting to identify change type. Also half of the filler passages were tested as changed and half as identical to maintain the constant probability of changed and identical tests. Test types were balanced across subjects such that all sentences were tested equally often in their identical and changed versions, and changed tests were equally often lexical and semantic. Test sequences were identical for thematic and unrelated groups.

Paraphrase subjects decided whether changes made in the test sentence maintained or altered that sentence's meaning. The same two words were always changed for the paraphrase and altered versions. For meaning preserving changes, synonyms were substituted for the two changed words thus altering the wording but not the meaning of the sentence. For meaning altered versions the two-word substitutes changed both the wording and meaning of the sentence. However the altered version remained a sensible sentence and could plausibly (though not as well as the original) have occurred in the paragraph. Table I contains examples of these two types of changes. Subjects simply indicated whether the test sentence was a paraphrase of the original. Across each set of 96 experimental passages, the adjectives, nouns, and verbs were changed with approximately equal frequency, but unpredictably, thus forcing subjects to attend to all words in the sentence. Test sequences for thematic and unrelated passages were identical. Within each session and across subjects, all passages were tested equally often in their paraphrased and altered versions, and equally often with silent and suppressed reading.

Procedure. The sentences were typed one per card and presented at a rate of one card every two seconds. Subjects were given practice in counting as rapidly as possible and were corrected if their rate slowed during the session. A practice sheet containing examples of the test types was given to ensure that subjects understood the use of the responses. All subjects receiving thematic passages were instructed to try to relate the sentence meanings so the entire passage would be more memorable. The subjects receiving unrelated passages were told to remember all seven sentences. Test sentences were also typed on cards, and

Fig. 3 Lexical and semantic detection performance for thematic and unrelated passages.



subjects were forced to respond within ten seconds on every trial.

Subjects. Thirty-two undergraduate volunteers served for two sessions each in the lexical-semantic detection task. Another sixteen volunteers participated for two sessions each in the paraphrase detection task. Each subject was paid $6 for completing both sessions.

Results and Discussion

Lexical-Semantic Detection. Figure 3 presents the hit minus false alarm values for the main conditions. Each data point in Figure 3 represents the mean for sixteen subjects, where the hit probability is based on four observations, and the false alarm probability on eight observations for each subject. Hits represent the correct, and false alarms the incorrect, occurrences of "changed" responses. This measure corrects for response bias (Sachs, 1974). The results replicate those of Levy (1977, experiment III) and provide additional information about the memory duration of the effects involved. As is apparent from Figure 3, the effects of thematicity again occur with semantic detection, with no effect on the synonym discriminations of the lexical task. Again, both measures show a substantial suppression decrement. These observations were confirmed by analysis of variance. These results can best be summarized as showing that memory improves from early to late sentence positions; silent reading is superior to reading while counting, at least for the final sentences; semantic detection is superior to lexical detection for early sentence positions and thematicity aids only semantic not lexical discriminations. Importantly, analysis of the semantic measure indicated that the effects of thematicity and of suppression were completely independent, as indicated by an $F < 1$ for their interaction term. Thus, as in Levy (1977, experiment III) both thematicity and suppression of speech responding made independent contributions to semantic detection performance.

The interaction of sentence position with suppression but not with thematicity was also informative. While the effects of semantic relatedness appear to extend throughout long-term memory, the suppression decrement dissipates after three or four sentences. The loss of the suppression decrement for lexical detection could be attributed to a floor problem, but the disappearance in semantic detection, where at least in the thematic condition performance is still reasonably good, is not open to this interpretation. One explanation of this dissipation is that suppression is indeed related to word memory, and memory for wording dissipates after short intervals (Sachs, 1967; Begg, 1971). If this is true, the suppression effect is itself limited by the duration of word memory.

Paraphrase Detection. Figure 4 presents the paraphrase detection data (each data point represents the mean for eight subjects, where the hit and false alarm estimates are each based on eight observations per subject). As the figure indicates, suppression had no consistent effect on this meaning measure ($p<.05$) while thematicity of the seven-sentence passages enhanced performance ($p<.025$).

Taken together, the lexical-semantic and paraphrase results of the present experiment offer good support to the "surface structure versus gist" explanation of the independent thematicity and suppression effects observed in subject-object reversal tasks. When the task can be solved from surface structure information (e.g., subject-object reversals) then a suppression effect occurs that is independent of the effects of meaningfulness. When the task cannot be solved from surface information (e.g., word order cues), but is completely dependent on meaning information (e.g., the paraphrase task), a suppression decrement does not occur. Thus the presence or absence of a suppression decrement is related to the usefulness of exact wording information in performing the task. A second finding that supports the word memory interpretation of the suppression decrement is that the effect dissipates after a few sentences. It has also been demonstrated that memory for sentence wording dissipates after short intervals, again consistent with the view that suppression and word memory are related.

The present findings converge on the view that suppressing speech responding harms memory for sentence wording. This information can be useful in some reading tasks. However, speech responding does not affect meaning analyses, which proceed independently in the lexical-semantic task, and which are unaffected by suppression in the paraphrase situation. Comprehension does not appear to depend on converting visual words into a speech code.

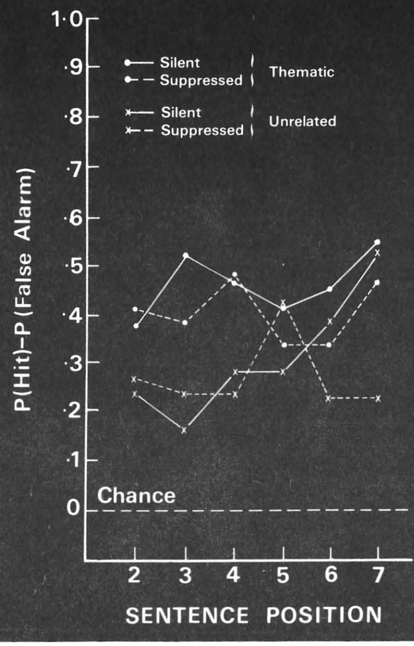


Fig. 4 Paraphrase detection performance for thematic and unrelated passages.

General Discussion

The literature reviewed in the first section of this paper suggests that while subjects often convert visual language into a speech code during reading tasks, the process is *not* a necessary stage prior to lexical access. An alternative view of where speech recoding occurs during reading, namely in working memory to aid sentence comprehension, was explored in the next section of the paper. Kleiman (1975) proposed a model of working memory in which comprehension of phrases or sentences required the words to be retained in a speech code in working memory, until parsing and integration mechanisms processed the sentence message. Since auditorily presented language was already in a speech code, the point of convergence for listening and reading appeared to be at the speech recoding stage in working memory. That is, once in working memory all language is encoded in a common speech format upon which the comprehension mechanisms work.

Both Kleiman's results and my own earlier studies fit quite well into this working memory framework. Some aspects of the present data also support the position. Suppressing speech responding by asking subjects to count led to decrements for both lexical and semantic detection during reading. Also since the suppression decrement was unreliable after about three sentences, some relatively short-term process is implicated. If the suppression decrement is related to word memory, it should be a relatively short-term phenomenon, since memory for sentence wording dissipates after about seven seconds (Begg, 1971). These data are consistent with the working memory view.

It is in relating the suppression effect to comprehension and memory for meaning that a problem arises. According to Kleiman's view, speech encoding is used to keep words available in memory until comprehension occurs. By this view, if suppression interferes with word retention, it should also interfere with comprehension, since the words are needed by the parsing and integration processes which lead to sentence comprehension. Further, in Kleiman's model thematic processes were said to influence a stage of processing later yet in the sequence. It is difficult to see how such a model could explain a thematic effect that is independent of disruption in the earlier speech stage. Since the speech code in working memory provides the data base on which comprehension processes work, disruption of the data base must also disrupt the semantic analyses. Evidence that word memory was seriously hampered by suppression in the present studies is available from the lexical and semantic change detection task. However, in the paraphrase tasks speech suppression did not inhibit semantic analyses during reading, even though word processing must have been disrupted. These findings are difficult to reconcile with Kleiman's model.

The apparent discrepancy between the present results and those of the earlier work of Kleiman (1975) and of Levy (1975, 1977, experiments I and II), where sentence processing was disrupted by speech suppression, may well be due to the tasks they used. Kleiman's sentence acceptability task required subjects to hold all sentence words in memory until the final critical word was presented, thus forcing a dependency on word holding mechanisms. The task also involves single unrelated sentences, providing no opportunity for thematic processes to affect performance. Similarly the unrelated sentences used in my earlier studies probably discouraged thematic processing. Also, the semantic detection task produced very subtle semantic alterations, and could be solved by simply noting the word order changes in the surface structure. These conditions probably discouraged the semantic processing underlying the thematic effect observed here and may have produced an artificial dependency between speech recoding and comprehension.

The sequential working memory model proposed by Kleiman seems unable to handle the independent thematicity and suppression effects observed during reading. To handle the present results a model requires processors for meaning and wording that can act independently rather than sequentially, but with a mechanism available for combining the output from these processors. For example, the working memory model proposed by Baddeley and Hitch (1974) could handle these independent effects. The articulatory loop is a speech system which could be suppressed by counting, while the independent executive processor assessed meaning and was aided by thematicity of a passage. A limitation of this approach is that the articulatory loop holds only three words or a second and a half of speech, estimates too small to account for the suppression decrement observed in our experiments. Also, since the nature of the executive processor is unspecified, it is unclear how thematicity actually aids processing and how the information in the executive processor and the articulatory loop combine to jointly affect performance. An interesting alternative to the working memory approach is Rumelhart's (1977) interactive model, where processing occurs simultaneously and interactively at many levels of encoding (e.g. visual feature, letter, syntax, semantic etc.), with each level attempting to understand the input within its level of responsibility, but using "hints" from other levels to aid its processing. Comprehension is the cumulative result of information from all analyses.

While the present data do not conclusively support any one model of language processing, they do constrain the class of models that is acceptable. A system which attempts to map reading onto listening comprehension mechanisms at some early stage of analysis seems unlikely. Rather, a processing system which allows independent analyses of signal and semantic levels of information, with a specified means of combining these sources of information, seems preferable. The interactive approach offers a fruitful framework for further reading research and has also been advocated as a model of listening (Marslen-Wilson, note 4). Rather than viewing reading as parasitic on listening, these two activities may share a common set of interactive processors.

1 This paper was prepared while the author was on sabbatical leave, visiting the MRC Applied Psychology Unit, Cambridge, England. I thank Unit members for stimulating discussions which influenced views expressed in the paper, and for secretarial and technical assistance in preparing the manuscript. The research was supported by Grant A7657 from the National Research Council of Canada. I also gratefully acknowledge the assistance of Catherine Mamer, Debbie Richter and Mike Withey, who contributed to data collection and analyses. Finally, the critical comments of D. Shankweiler, F.I.M. Craik, and D. Massaro on the drafts of this manuscript were greatly appreciated.

Reference Notes

1. Barron, R. W.
Access to the lexical meaning of written words. Unpublished manuscript, 1976 (available from the author).

2. Levy, B. A., & Brevik, M.
Auditory and visual processing during subsidiary shadowing activity. McMaster Technical Report, No. 61, 1974.

3. Meyer, D. E., & Ruddy, M. G.
Lexical-memory retrieval based on graphemic and phonemic representations of printed words. Paper presented at the meeting of the Psychonomic Society, St. Louis, 1973.

4. Marslen-Wilson, W.
Processing interactions and lexical access during word recognition in continuous speech. Paper read at the Experimental Psychology Society Meeting, Oxford, 1977.

References

Allport, D. A.
On knowing the meaning of words we are unable to report: The effects of visual masking. In S. Dornic (Ed.), *Attention and Performance IV,* London: Academic Press, 1976.

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. (a).

Baddeley, A. D.
The influence of acoustic and semantic similarity on longterm memory for word sequences. *Quarterly Journal of Experimental Psychology,* 1966, *18,* 302-309. (b).

Baddeley, A. D., & Hitch, G.
Working memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation.* Vol. 8, New York: Academic Press, 1974.

Baron, J.
Phonemic stage not necessary for reading. *Quarterly Journal of Experimental Psychology,* 1973, *25,* 241-246.

Baron, J.
Mechanisms for pronouncing printed words: Use and acquisition. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension,* Potomac, Maryland: L. E. Erlbaum, 1976.

Begg, I.
Recognition memory for sentence meaning and wording. *Journal of Verbal Learning and Verbal Behavior,* 1971, *10,* 176-181.

Bower, T. G. R.
Reading by eye. In H. Levin & J. P. Williams (Eds.), *Basic Studies on reading,* New York: Basic Books, 1970.

Bradshaw, J. L.
Three interrelated problems in reading. A review.
Memory and Cognition,
1975, *3,* 123-134.

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

Conrad, R.
Speech and reading. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationships between speech and reading.*
Cambridge, Mass.: MIT Press, 1972.

Craik, F. I. M., & Lockhart, R. S.
Levels of processing. A framework for memory research.
Journal of Verbal Learning and Verbal Behavior,
1972, *11,* 671-684.

Ericksen, C. W.; Pollack, M. D., & Montague, W. E.
Implicit speech: Mechanisms in perceptual coding.
Journal of Experimental Psychology,
1970, *84,* 502-507.

Forster, K. I., & Chambers, S. M.
Lexical access and naming time.
Journal of Verbal Learning and Verbal Behavior,
1973, *12,* 627-635.

Gleitman, L. R., & Rozin, P.
The structure and acquisition of reading, I: Relations between orthographics and the structure of language. In A. Reber & D. Scarborough (Eds.),
Toward a psychology of reading.
Hillsdale, N.J., L. E. Erlbaum, 1977.

Green, D. W., & Shallice, T.
Direct visual access in reading for meaning.
Memory and Cognition,
1976, *4,* 753-758.

Hardyck, C. D., & Petrinovich, L. F.
Subvocal speech and comprehension level as a function of the difficulty level of reading materials.
Journal of Verbal Learning and Verbal Behavior,
1970, *9,* 647-652.

Johnson, N. F.
On the function of letters in word identification: Some data and a preliminary model.
Journal of Verbal Learning and Verbal Behavior,
1975, *14,* 17-29.

Klapp, S. T.
Implicit speech inferred from response latencies in same-different decisions.
Journal of Experimental Psychology,
1971, *91,* 262-267.

Klapp, S. T.; Anderson, W. G., & Berrian, R. W.
Implicit speech in reading, reconsidered.
Journal of Experimental Psychology,
1973, *100,* 368-374.

Kleiman, G. M.
Speech recoding in reading.
Journal of Verbal Learning and Verbal Behavior,
1975, *14,* 323-339.

Kroll, N. E. A.; Parks, T. E.; Parkinson, S. R.; Beiber, S. L., & Johnson, A. L.
Short-term memory while shadowing: Recall of visually and aurally presented letters.
Journal of Experimental Psychology,
1970, *85,* 220-224.

Kroll, N. E. A.
Visual short-term memory. In D. Deutsch & J. A. Deutsch (Eds.),
Short-term memory.
New York: Academic Press, 1975.

Levy, B. A.
Role of articulation in auditory and visual short-term memory.
Journal of Verbal learning and Verbal Behavior,
1971, *10,* 123-132.

Levy, B. A.
Vocalization and suppression effects in sentence memory.
Journal of Verbal Learning and Verbal Behavior,
1975 *14,* 304-316.

Levy, B. A.
Reading: Speech and meaning processes.
Journal of Verbal Learning and Verbal Behavior,
1977, *16,* 623-638.

Marshall, J. C., & Newcombe, F.
Syntactic and semantic errors in paralexia.
Neuropsychologia,
1966, *4,* 169-176.

Massaro, D. W.
Understanding language: An information-processing analysis of speech perception, reading and psycholinguistics.
New York: Academic Press, 1975.

Murdock, B. B.
Auditory and visual stores in short-term memory.
Acta Psychologica,
1967, *27,* 316-327.

Norman, D. A.
The role of memory in the understanding of language. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationship between speech and reading.*
Cambridge: Mass: MIT Press, 1972.

Patterson, K. E. & Marcel, A. J.
Aphasia, dyslexia and the phonological coding of written words.
Quarterly Journal of Experimental Psychology,
1977, in press.

Peterson, L. R., & Johnson, S. T.
Some effects of minimizing articulation on short-term retention.
Journal of Verbal Learning and Verbal Behavior,
1971, *10,* 346-357.

Rubenstein, H.; Lewis, S. S., & Rubenstein, M. A.
Evidence for phonemic recoding in visual word recognition.
Journal of Verbal Learning and Verbal Behavior,
1971, *10,* 647-657.

Rumelhart, D. E.
Toward an interactive model of reading.
Technical Report No. 56.
Center for Human Information Processing, University of California, San Diego, 1976.

Sachs, J. S.
Recognition memory for syntactic and semantic aspects of connected discourse.
Perception and Psychophysics,
1967, *2,* 437-442.

Sachs, J. S.
Memory in reading and listening to discourse.
Memory and Cognition,
1974, *2,* 95-100.

Shallice, T. & Warrington, E. K.
Word recognition in a phonemic dyslexic patient.
Quarterly Journal of Experimental Psychology,
1975, *27,* 187-199.

Spoehr, K. T., & Smith, E. E.
The role of syllables in perceptual processing.
Cognitive Psychology,
1973, *5,* 71-89.

Sperling, G.
The information available in brief visual presentation.
Psychological Monographs,
1960, *74,* (11; Whole No. 498).

Theios, J., & Muise, J. G.
The word identification process in reading. In N. J. Castellan, Jr., & D. Pisoni (Eds.)
Cognitive Theory, Volume II,
Potomac, Maryland: L. E. Erlbaum, 1976.

Biography

Dominic W. Massaro is professor of psychology and associated faculty member of the Research and Development Center for Cognitive Learning at the University of Wisconsin (Madison, WI 53706). His research and teaching interests include human information processing, speech perception, reading, and the perception of art and music. Dr. Massaro is author of *Experimental Psychology and Information Processing* and has edited and contributed chapters to *Understanding Language: An Information-Processing Analysis of Speech Perception, Reading, and Psycholinguistics.*

Richard E. Pastore is an associate professor of psychology at the State University of New York at Binghamton (Binghamton, NY 13901). His major research interests are in psychoacoustics, signal detection theory, selective attention, masking, and phoneme perception. He earned a Ph.D. from Purdue University and did post-doctoral training at Central Institute for the Deaf in St. Louis. He currently is on sabbatical leave at Haskins Laboratories in New Haven, Connecticut.

Uta Frith is a research psychologist at the Medical Research Council's Developmental Psychology Unit (London, WCIH OAN, England). Having received a traditional classical education in Germany and having read art history and psychology at Saarbrücken University, she came to London to train as a clinical psychologist at the Institute of Psychiatry. She gained a Ph.D. in the area of childhood autism. Her main research interests are abnormalities of cognitive development. She is at present editing a book on cognitive processes in spelling.

Jonathan Baron is an associate professor of psychology at the University of Pennsylvania (Philadelphia, PA 19174). He received a Ph.D. in psychology from the University of Michigan in 1970. His interests include reading, cognitive development, and educational psychology. June Hodge is a reading specialist in the West Deptford, New Jersey (08066), school system. She received a B.A. in psychology from the University of Pennsylvania and a B.S. and an M.S. from the University of Pennsylvania School of Education.

James G. Martin, a Minnesota Ph.D., is professor of psychology at the University of Maryland (College Park, MD 20742). Richard H. Meltzer, a University of California (San Diego) Ph.D., taught at Maryland and is now with General Electric, Inc. Carol B. Mills is a 1976 Maryland Ph.D. now on a fellowship at Gallaudet College (Washington, DC 20002).

Betty Ann Levy is an associate professor of psychology at McMaster University (Hamilton, Ontario, Canada), where she has taught since 1970. She has been visiting scholar at the University of Sussex, 1969-1976, and at MRC Applied Psychology Unit, Cambridge, 1977. Dr. Levy's earlier publications were concerned with encoding processes in short-term memory, but more recently her interest has been in information-processing models of reading and listening.

Acknowledgements

The Editor gratefully acknowledges the assistance of the people listed below in reviewing the papers in this special issue. Their reviews were highly helpful to both the editor and the authors.

1 Jonathon Baron
University of Pennsylvania
Philadelphia, PA 19104

2 James Cutting
Wesleyan University
Middletown, CT 06457

3 Glenn Kleiman
University of Illinois
Urbana, IL 41801

4 Neal Kroll
University of California
Davis, CA 95616

5 Kenneth Paap
New Mexico State University
Las Cruces, NM 88003

6 Martin Singer
University of Denver
Denver, CO 80208

7 Glen Taylor
University of Wisconsin
Madison, WI 53706

8 Alex Wilkinson
University of Wisconsin
Madison, WI 53706

Résume des Articles

La lecture et l'audition figurées à l'aide d'une maquette à trois dimensions *par Dominic W. Massaro*

Le traitement du langage consiste à former un sens à partir d'un signal, tel qu'un texte imprimé ou un énoncé verbal. L'objet de la maquette est de montrer non pas seulement ce que le lecteur ou l'auditeur doit savoir pour comprendre la nature du langage, mais la manière même dont ces données sont traitées. Le traitement du langage est considéré comme une succession d'opérations internes qui s'effectuent entre le moment du stimulus et l'élaboration du sens. A chaque stade les opérations prennent un certain temps et affectent l'information d'une manière ou d'une autre, en vue du stade suivant. Dans la maquette proposée il y a un organe *classement*, pour figurer la nature de l'information à un stade donné et un autre organe, *fonction*, pour figurer les opérations de traitement proprement dites. Cette maquette est proposée à des fins heuristiques, et pour incorporer les données et les théories de toute étude sur la fonction langage.

Les phonèmes et les symboles alphanumériques: les parallélismes éventuels entre les systèmes humains de communication, *par Richard E. Pastore*

Les symboles alphanumériques et les phonèmes peuvent être considérés comme des codes dans deux systèmes humains de communication. Si ces systèmes sont conçus en vue d'obtenir certains résultats, on peur s'attendre à retrouver des caractéristiques communes dans les éléments physiques de la représentation de ces codes et de leur perception. Ces caractéristiques sont examinées en vue d'établir leur importance par rapport à tous les systèmes de communication en général et par rapport à leur manifestation dans le langage auditif et visuel. Dans cette perspective générale de tous les systèmes de communication on peut s'attendre à trouver un grand nombre de parallèles dans la perception des symboles alphanumériques et des phonèmes. L'auteur examine quelques uns de ces parallèles en s'appuyant sur ce que nous savons sur la perception chez l'homme.

Du texte au sens et du texte au son, ou comment lisent ceux qui ne savent pas l'orthographe, *par Uta Frith*

Deux groupes d'enfants. Age: douze ans. Même niveau d'intelligence. Même capacité de lecture. Les uns savent l'orthographe; les autres, non. De même force pour lire les mots et les phrases, mais inégaux à d'autres exercices, par ex. la lecture à haute voix et la lecture des mots dénués de sens. A l'examen, il est apparu que ceux qui ne savent pas l'orthographe attrappent bien le sens, mais sont empêchés lorsqu'il est question de phonétiser. Tandis que ceux qui savent l'orthographe sont également bons pour comprendre et pour phonétiser.

La perception des correspondances entre l'orthographe et l'audition *par Jonathan Baron et June Hodge*

Quelques adultes ont appris des mots dénués de sens écrits dans un alphabet truqué. Les correspondances entre lettres et phonèmes étaient cachées par inversion de droite à gauche. Bien qu'ils n'aient pas remarqué ces correspondances, les participants sont parvenus à déchiffrer d'autres mots formés au moyen du même alphabet. Dans une expérience suivante des mots aussi dénués de sens mais formés à l'aide de correspondances ont été lus plus rapidement que d'autres mots tout aussi dénués de sens mais formés sans aucune correspondance. Une troisième expérience révéla que ce résultat était attribuable au fait que des mots semblables produisent des effets semblables. Au total, il semble bien que des correspondances peuvent bien être utilisées sans recours à des stratégies particulières pour les découvrir et que les gens se fabriquent des modèles plutôt que de rechercher les correspondances elles-mêmes.

Les rythmes visuels: la mobilité dans la présentation des textes pour l'étude des langues *par James G. Martin, Richard H. Meltzer, et Carol B. Mills*

La méthode proposée consiste en une présentation mobile des textes. Un écran de télévision sert à présenter simultanément la version visuelle et la version auditive d'une même phrase, de telle sorte qu'elles forment des séquences parallèles. L'apparition de chaque syllable sur l'écran est accompagnée par l'émission de la prononciation. On voit donc une phrase "énoncée" de gauche à droite sur l'écran, syllable par syllable selon un rythme visuel. On a fait une expérience pour laquelle on avait formé trois groupes d' élèves du secondaire qui avaient choisi l'espagnol comme deuxième langue. Au cours des leçons, ceux du groupe "rhyme" voyaient défiler les phrases selon un rhyme visuel; ceux du groupe "statique" voyaient la même phrase mais en image fixe; ceux du troisième groupe ne voyaient ni l'une ni l'autre. Avant comme après le cours ils avaient tous fait la démonstration qu'ils savaient parfaitement lire à haute voix. Ce qu'il s'agissait de mesurer c'était les progrès réalisés dans la lecture orale: les élèves du cours rythme eurent l'avantage.

Les opérations du langage: lire et écouter *par Betty Ann Levy*

L'auteur s'intéresse aux opérations de recodage du langage au cours de la lecture. Et plus spécialement aux opérations au cours desquelles la lecture et l'audition pourraient partager des mécanismes communs pour la compréhension. L'article est partagé en quatre sections. Dans la première l'auteur examine de point de savoir si un recodage paraît indispensable pour atteindre le niveau lexicologique. Dans la deuxième, il envisage une autre hypothèse, selon laquelle ce recodage se situerait dans la mémoire active, celle où les mots sont conservés sous leur forme d'énoncé jusqu'au moment où le sens de la phrase apprait. Dans la troisième, il décrit une expérience qui tend à montrer que de déranger l'ordre des mots dans la mémoire ne compromet pas nécessairement la compréhension sémantique. Les résultats obtenus tendent à montrer que les opérations de la lecture ne consistent pas à convertir les signaux visuels en un autre code, celui du langage oral, jusqu'à ce que les sens apparaisse. Finalement l'auteur aborde la question plus générale des modèles figurant les opérations du langage.

Kurzfassung der Beiträge

Ein Stufenmodell des Lesens und Zuhörens
von Dominic W. Massaro

Sprachverarbeitung ist die Abstraktion der Bedeutung von einem physischen Zeichen wie einem gedrucktem Text oder einer Folge von Sprachlauten. Das Ziel eines Informationsverarbeitungsmodells ist es, zu beschreiben, *wie* Sprache verarbeitet wird, nicht einfach, was der Leser oder Zuhörer wissen muß, um Sprache zu verstehen. Sprachverarbeitung wird betrachtet als eine Folge von inneren Verarbeitungsstufen oder Operationen, die zwischen dem Sprachreiz und der Bedeutung auftreten. Die Operationen einer bestimmten Stufe brauchen ihre Zeit und transformieren die Information der nächsten Stufe der Verarbeitung zugänglich machen. In dem vorliegenden Modell beschreibt die *Speicherungskomponente* die Art der Information auf einer bestimmten Stufe, wogegen die *funktionale Komponente* die Operationen auf einer Stufe der Verarbeitung beschreibt. Das Informationsverarbeitungsmodell wird heuristisch angewendet, um Daten und Theorie einer Reihe von Sprachverarbeitungs-Studien einzubeziehen.

Phoneme und Alphanumerische Zeichen: Mögliche Komponenten paralleler menschlicher Kommunikationssysteme *von Richard E. Pastore*

Alphanumerische Zeichen und Phoneme können als Informationsverschlüsselungssysteme betrachtet werden, die in menschlichen Kommunikationssystemen benutzt werden. Wenn solche Kommunikationssysteme so geplant sind, daß sie effektiv sein können, dann sollten wir erwarten, bestimmte Merkmale zu finden, die sich in der Natur der physischen Repräsentation dieser Informationsverschlüsselungen und in ihrer Wahrnehmung niederschlagen. Diese Merkmale werden diskutiert hinsichtlich ihrer Bedeutung für Kommunikationssysteme allgemein und ihrer Verwirklichung in der hörbaren und sichtbaren menschlichen Sprache. Unter dem Blickwinkel solcher Kommunikationssysteme gesehen, sollten wir erwarten, viele Parallelen zwischen der Wahrnehmung von alphanumerischen Zeichen und von Phonemen zu finden. Dieser Aufsatz untersucht einige dieser Parallelen, wobei er sich auf unsere Kenntnis der menschlichen Wahrnehmung bezieht.

Von der Schrift zur Bedeutung, und von der Schrift zum Klang, oder: Wie man lesen kann, ohne buchstabieren zu können *von Uta Frith*

Zwei Gruppen von Zwölfjährigen normaler Intelligenz und Lesefertigkeit wurden verglichen. Eine Gruppe bestand aus Schülern, die sicher in der Rechtschreibung waren, die andere aus solchen, die darin schlecht waren. Beide Gruppen waren gleich gut beim Lesen einzelner Wörter und Sätze. Sie unterschieden sich jedoch in anderen Leseaufgaben, besonders bei sinnlosen Wörtern und anderen Aufgaben, bei denen Schrift in Laute umgesetzt werden muß. Die Unterschiede deuten darauf hin, daß die schlechten Rechtschreiber erfolgreich Schrift direkt in Bedeutung umsetzen konnten, aber weniger gut Schrift in Laute. Im Gegensatz dazu zeigten gute Rechtschreiber Beherrschung beider Aspekte des Lesens, indem sie Schrift in Bedeutung und in Laute umsetzten.

Die Bedeutung von Schreibweise-Klang-Entsprechungen ohne den Versuch, sie zu lernen *von Jonathan Baron und June Hodge*

Erwachsene Versuchspersonen lernten gesprochene Reaktionen auf sinnlose Worte, die in einem künstlichen Alphabet geschrieben waren. Entsprechungen zwischen Buchstaben und Phonemen waren hinter Rechts-Links-Entsprechungen verborgen. Obwohl die Versuchspersonen das Vorhandensein von Entsprechungen nicht bemerkten, waren sie doch in der Lage, neue sinnlose Wörter in dem gleichen Alphabet zu entschlüsseln. In einem zweiten Experiment wurden sinnlose Wörter, die mit verborgenen Entsprechungen geschrieben waren, schneller gelesen als sinnlose Wörter ohne solche Entsprechungen. Ein drittes Experiment legte den Schluß nahe, daß dieser Effekt darauf zurückzuführen ist, daß ähnliche Wort ähnliche Reaktionen hatten. Allgemein geht aus den Ergebnissen hervor, daß Entsprechungen benutzt werden können, ohne die Benutzung besonderer Strategien zum Erlernen der Entsprechungen. Aber wenn das auftritt, benutzen die Leute eher Beispiele als das Wissen um die Entsprechung selbst.

Visueller Rhythmus: Dynamische Textdarbietung zum Erlernen einer zweiten Sprache *von James G. Martin, Richard H. Meltner, & Carol B. Mills*

Es wird eine Methode beschrieben, mit der Sätze dynamisch visuell dargeboten werden. Ein Fernseh-Bildschirm wird verwendet, um gleichzeitig die visuelle und die akustische Version eines Satzes darzubieten, wobei jede seiner aufeinanderfolgenden visuellen und akustischen Silben parallel miteinander verbunden werden; die Darbietung jeder visuellen Silbe ist synchronisiert mit der Darbietung jeder Silbe, die durch den akustischen Kanal gehört wird. Das Ergebnis ist ein Satz, der von links nach rechts "wächst" über den Bildschirm, Silbe für Silbe, in "visuellem Rhythmus". In einem Experiment dienten als Versuchspersonen drei Gruppen von Mittelschülern, die Spanisch als zweite Sprache lernten. In den Übungssitzungen sah die rhythmische Gruppe die Sätze in "visuellem Rhythmus", die unrhythmische Grupp dagegen in statischer Darbietung, und die Kontrollgruppe war keinerlei audio-visuellen Darbietungen ausgesetzt. Vor und nach der Übung wurden von allen Gruppen Vor- und Nachtestmaße für die Flüssigkeit ihres lauten Vorlesens erhoben. Abhängige Variable war die relative Veränderung der beurteilten Leseflüssigkeit von vorher zu nachher. Die Ergebnisse waren in der rhythmischen Gruppe am besten.

Sprachanalyse während der Verarbeitung von Sätzen *von Betty Ann Levy*

Die vorliegende Arbeit befaßt sich mit der Rolle der Sprach-Umkodierung während des Lesens. Insbesondere untersucht sie die Informationsverarbeitung hinsichtlich des Punktes, wo Lesen und Zuhören während des Verständnisses möglicherweise gemeinsame Mechanismen benutzen. Der Aufsatz ist in vier Teile unterteilt: Der erste Abschnitt enthält eine Übersicht der Ergebnisse hinsichtlich des Problems, ob Sprach-Umkodierung vor dem lexikalischen Zugang notwendig ist. Die meisten Daten sprechen gegen diesen Standpunkt. Der zweite Abschnitt des Aufsatzes untersucht einen alternativen Gesichtspunkt—nämlich, daß Sprach-Umkodierung im Arbeitsgedächtnis auftritt, wo Worteinheiten in Sprachform behalten werden, bis ein Verständnis der Satzteile oder Sätze auftritt. Abschnitt III beschreibt ein Experiment, das zeigt, daß eine Unterbrechung der Formulierungsinformation nicht zum Versagen des semantischen Verständnisses führt. Diese Ergebnisse legen den Schluß nahe, daß das Lesen nicht darin besteht, visuelle Signale in sprachliche Verschlüsselung umzusetzen, solange kein Verständnis auftritt. Schließlich konzentriert sich die allgemeine Diskussion auf Modelle der visuellen Sprachverarbeitung.

Resumen de los Aríctulos

Un modelo de etapas del leer y escuchar, por Dominic W. Massaro

El proceso del lenguaje es la abstracción de significado de una senal física tal como un texto impreso o la serie de sonidos de lenguaje. El fin de un modelo de proceso de información es de describir cómo se procesa el lenguaje y no simplemente lo que el lector o el oyente debe saber para entender el lenguaje. Se considera el proceso del lenguaje como una serie de etapas de proceso interno u operaciones que ocurren entre el estímulo del lenguaje y el significado. Las operaciones de una etapa en particular toman tiempo y transforman la información en algun sentido haciendo que la información transformada sea disponible para la siguiente etapa del proceso. En el presente modelo el componente almacenamiento describe la naturaleza de la información en una etapa particular del proceso mientras el componente funcional describe las operaciones de cualauier etapa del proceso. El modelo de proceso de información es usado para incorporar datos y teoría de una variedad de estudios de proceso del lenguaje.

Fonemas y caracteres alfanuméricos: posibles componentes de sistemas paralelos de comunicación humana, por Richard E. Pastore

Se puede considerar a los caracteres alfanuméricos y fonemas como códigos de información usados por sistemas de comunicación humanos. Si dichos sistemas de comunicación fueron disenados para ser efectivos, entonces deberíamos esperar de encontrar ciertas características que deberían manifestarse en la naturaleza de las representaciones físicas de los códigos de información y en la percepción de dichos códigos. Se discuten estas características en términos de su importancia en los sistemas de comunicación en general y sus manifestaciones en lenguaje humano oíble y visible. Al contemplar desde la perspectiva de tales sistemas de comunicación, debemos esperar encontrar muchos paralelos en la percepción de caracteres alfanuméricos y fonemas. Este artículo examina algunos de estos paralelos extrayendo de nuestro conocimiento de la percepcion humana.

De la imprenta al significado y de la imprenta al sonido, o cómo leer sin saber deletrear, por Uta Frith

Se compararon dos grupos de 12 años con inteligencia normal y en edad de lectura. Un grupo consistía de buenos deletreadores y otro de pobres deletradores. Los dos grupos eran igualmente buenos en leer palabras solas y oraciones. Sin embargo diferían en otras tareas de lectura, notablemente con palabras tontas y otras tareas que involucraban la conversión de la imprenta en sonido. Las diferencias indicaron que los deletreadores deficientes eran peritos en ir de la imprenta directamente al significado, pero fallaban al convertir imprenta en sonido. En contraste, los buenos deletreadores mostraron maestría en ambos aspectos de la lectura, convirtiendo imprenta en significado e imprenta en sonido.

El uso de correspondencias de deletreo-sonido sin tratar de aprenderlas, por Jonathan Baron y June Hodge

Sujetos adultos aprendieron respuestas orales para tontas palabras escritas en un alfabeto artificial. Correspondencias entre letras y fonemas estuvieron escondidas por el uso de correspondencias de derecha-a-izquierda. Aun cuando sujetos no notaron la existencia de correspondencias, fueron capaces de decodificar nuevas palabras tontas en el mismo alfabeto. En un segundo experimento, palabras tontas escritas con correspondencias ocultas fueron leídas mas rapido que palabras tontas sin correspondencias. Un tercer experimento sugería que este efecto se debia al hecho que palabras similares tenian respuestas similares. En general, los resultados sugieren que correspondencias pueden usarse sin utilizar catrategias especiales para aprender correspondencias, pero cuando esto ocurre, la gente usa ejemplos en vez del conocimiento de las correspondencias en si.

Ritmos visuales: muestra dinámica del texto para aprender a leer un segundo idioma, por James G. Martin, Richard H. Meltzer, y Carol B. Mills

Se describe un método por el cual se presentan oraciones en una muestra visual dinámica. Se usa un monitor de televisión para presentar simultáneamente las versiones visuales y audibles de una oración, con cada una de sus sucesivas sílabas visuales y audibles uncidas paralelamente; el comienzo impetuoso de cade silaba visual es sincronizado con el comienzo de cada sílaba como se escucha a traves del canal auditivo. El resultado es una oración que "crece" a través de la pantalla de izquierda-a-derecha, cada sílaba por separado, en "ritmo visual." En un experimento, los sujetos eran tres grupos de estudiantes de escuela secundaria que estaban aprendiendo el español como segunda lengua. En las sesiones de entrenamiento, el grupo rítmico vió las oraciones en "ritmo visual"; el grupo no-rítmico vió las mismas oraciones pero en una muestra visual estática, y el grupo control no se expuso a ninguna muestra visual-auditiva. Antes y despues del entrenamiento, todos los grupos otorgaron medidas de pre-test y post-test de lectura oral fluída. La medida dependiente era el cambio relativo de pretest a post-test en la juzgada lectura fluída. Los resultados favorecieron al grupo rítmico.

Analisis del lenguaje en el proceso de la oración: el leer y el escuchar, por Betty Ann Levy

El presente articulo se refiere al papel de la recodificación del lenguaje durante la lectura. Específicamente, examina el proceso de la información con respecto a dónde el leer y el escuchar pueden venir a participar de mecanismos comunes durante la comprensión. El artículo se divide en cuatro secciones. La primera sección contiene un examen de la evidencia relacionada al punto en cuestion si el recódigo del lenguaje es necesario con anterioridad al acceso léxico. El peso de la evidencia está en contra de este punto de vista. La segunda sección del artículo explora un punto de vista alternativo—a saber—que la recodificación del lenguaje ocurre en la memoria activa donde unidades de palabras se mantienen en una forma de lenguaje hasta que la comprensión de frases y oraciones tiene lugar. La seción III describe un experimento que muestra que la información interrumpida de palabras en la memoria no conduce al fracaso de comprensión semántica. Estos resultados sugieren que la lectura no se da al convertir señales visuales a un código de lenguaje hasta que la comprensión tiene lugar. Finalmente, la discusión general se centra en modelos de proceso de lenguaje visual.

Merald E. Wrolstad, Ph.D., Editor and Publisher.

General Information

Visible Language is concerned with research and ideas that help define the unique role and properties of written language. It is a basic premise of the Journal that writing/reading form a distinct system of language expression which must be defined and developed on its own terms. Published quarterly since 1967, Visible Language has no formal organizational affiliation. All communications should be addressed to

Visible Language
Box 1972 CMA
Cleveland, OH 44106 USA
Telephone 216/421-7340

Subscription Rates

	One Year	*Two Years*	*Three Years*
Individual subscription	$15.00	$28.00	$39.00
Institutional subscription	25.00	47.00	66.00

Foreign subscribers: add $1.00 for postage to each year's subscription

All orders must be prepaid. To be honored free of charge, claims for missing issues must be made immediately upon the receipt of the next published issue.

Reprints and Back Copies

A folder listing the contents of all past Journal issues is available on request. Individual reprints are not available. A limited quantity of all back numbers is available at a per issue cost of $3.00 to individuals and $5.00 to institutions.

Letters to the Editor

Comments on articles, and letters that have appeared in the Journal are welcome and should be addressed to the Editor. The Editor will also relay to the author questions or comments on any article. Your response—and the author's comment in reply—will not be published without your permission and your approval of any editing.

Manuscripts

Manuscripts, inquiries about research articles, and other contributions to the Journal should be addressed to the Editor. An Author's Guide for the organization, preparation, and submission of manuscripts is available and includes special instructions for designers in preparing research reports. Manuscripts should be accompanied by an abstract typed on a separate sheet of paper.

Advertising

Detailed information for advertisers is available upon request.

typography and cover design
by peter megert