How Theories of Speech Affect Research in Reading and Writing*

Alvin M. Liberman

When we remember that words are sounds merely, we shall conclude that the idea of representing those sounds by marks, so that whoever should, at any time after, see the marks would understand what sounds they meant, was a bold and ingenious conception, not likely to occur to one man of a million in the run of a thousand years...That it was difficult of conception and execution is apparent, as well by the foregoing reflections as by the fact that so many tribes of men have come down from Adam's time to ours without ever having possessed it.

—Abraham Lincoln

My aim is to promote two notions about the relation between reading/writing and speech: the right theory of speech is essential to a coherent account of reading/writing; and the conventional theory of speech is the wrong theory. To say, as I will, that the conventional theory has therefore made it hard for researchers to see how reading/writing differ from speech is not to deny progress in the field; indeed, I believe, to the contrary, that research of the last few decades has brought insights that are both new and important. I would only suggest that the researchers who are responsible for those insights have either worked from the right theory or else managed somehow to ignore the wrong one. Naturally, I believe about the 'right' theory, not that it is perfectly and forever true, only that it is, by comparison with its more conventional competitor, more nearly right, and more likely, therefore, to head the reading/writing researcher in the right direction.

I should say that the conventional theory I consider wrong and the unconventional theory I consider right are only about speech in the narrow sense, by which I mean the component of the broader language faculty that comprises the production and perception of consonants and vowels. Though one of the virtues of the unconventional theory is that it makes speech an organic part of language, instead of the biologically arbitrary appendage that the conventional theory portrays, it is nonetheless possible for our purposes to deal with speech in isolation. However, I reserve the right to suggest at a later point that the biologically based fit of speech to the other components of language is an important reason why appreciation of the alphabetic principle is hard to come by.

To advance the two notions that are the point of this paper, I will rely almost entirely on facts that are in plain sight, requiring only to be thought about if their implications for speech and reading are to be seen. Because I had not myself thought about those facts, I would therefore emphasize I was long ago taken in by the conventional view. Specifically, I was led by it to suppose that speech is an acoustic alphabet, with segments of sound as discrete as the letters that convey them, and that I could, therefore, contrive an acoustic alternative for use in a reading machine for the blind (Liberman, 1996). Only after I had failed miserably to produce that alternative, and had then done a lot of research to find out why, did I begin to see that I might have known better before I started had I simply gone beyond surface appearances to take
account of the somewhat deeper, if still visible, considerations I will invite you to think about. (A reading machine for the blind is now a reality, but it produces speech, not an acoustic alphabet.) I will claim that the conventional view fails the reading/writing researcher for much the same reason it failed me. If that failure has gone largely unnoticed, it is not because the conventional speech researchers have been unable or unwilling to understand what might seem plain, but only because they have not been concerned with the relevance of their theory to research on reading/writing or reading machines, and have therefore not had occasion to measure its implications against the hard realities of those enterprises.

In addition to the facts about language that are apparent to everyone, I will refer to just a few that have come from research on speech, but those are easily understood without a technical background in acoustic phonetics; moreover, they are not in dispute. All this is to say that the matter is not difficult at all, except, perhaps, in the telling.

The relevance of a theory of speech. As for the first notion—that a proper theory of speech is essential to an understanding of how people read—the most relevant consideration arises out of the deep biological gulf that separates the two processes. Speech, on the one side, is a product of biological evolution, standing as the most obvious, and arguably the most important, of our species typical behaviors. Reading/writing, on the other, did not evolve biologically, but rather developed (in some cultures) as a secondary response to that which evolution had already produced. A consequence is that we are biologically destined to speak, not to read or write. Accordingly, we are all good at speech, but disabled as readers and writers; the difference among us in reading/writing is simply that some are fairly easy to cure and some are not.

Viewing the matter from a slightly different angle, we see that, being at least as old as our species, speech has been around for 200,000 years or more, while the idea that it could be rendered alphabetically was born no more than 4000 years ago. Subtracting the latter number from the former, we conclude that it took our ancestors at least 196,000 years just to discover how to describe what it was they did when they spoke. Why did it take so long? Why was it so hard for our prealphabetic ancestors to make the momentous discovery, and why is it so hard for our preliterate children to understand it? Why has an alphabet been developed only once in all of human history? Surely, questions like those cry out for a theory of speech that explains in the same breath why an alphabetic description of speech is not immediately apparent to everyone, and why it should be almost wholly beyond the reach of some. Nothing less will do if we are to know how to teach children who are somehow ready to cope, while also helping those who are not.

Contrasting views of the biology of speech. There is a question that goes to the heart of the difference between the conventional and unconventional views of speech: does the specialization for language extend to the motor and perceptual processes underlying the consonants and vowels that speech and reading/writing use in common? The guiding assumption of the conventional view is that it does not (Crowder, 1983; Diehl & Kluender, 1986; Fujisaki & Kawashima, 1970; Kuhl, 1981; Lindblom, 1991; Massaro, 1987; Stevens & Blumstein, 1978; Sussman, 1989; Sussman, 1991; Warren, 1993). On that view, language simply appropriated modes of motor control and auditory perception that had already evolved in connection with nonlinguistic functions. Having been adopted by language for its purposes, those plain vanilla processes are now seen on the conventional view to work horizontally, serving linguistic and nonlinguistic behaviors alike.

According to the unconventional view, the specialization for language extends much farther, embracing even the very low level where the primary motor and perceptual representations of speech are to be found (Liberman & Mattingly, 1985; Liberman & Mattingly, 1989; Mattingly & Liberman, 1990; Mann & Liberman, 1983; Remez, Rubin, Berns, Pardo, & Lang, 1994; Whalen & Liberman, 1987). In other words, there are distinctly linguistic representations, not in the higher reaches of the cognitive machinery, but down among the structures of action and perception. Thus, language is seen as a vertically organized system in which linguistically specialized structures (and processes) are as central to phonetics as they are to syntax.

Among the more particular assumptions of the two views, perhaps the most fundamental concerns the nature of the ultimate constituents of speech (and, for that matter, language). On the conventional view, they are sounds. Just about everybody (including Lincoln, as the otherwise insightful epigraph makes clear) simply takes that for granted. And just about
everybody holds that the sounds of speech are serviceable as consonants and vowels to the extent that they evoke distinctive auditory percepts.

The unconventional theorist, on the other hand, takes the ultimate constituents to be, not sounds, but articulatory gestures. Thus, the consonant we write as ‘b’ is a closure of the vocal tract at the lips, ‘d’ a closure at the alveolar ridge, and so forth. This notion came originally from research on speech that revealed vast context-conditioned variations in the sound as a result of the coarticulation of seemingly invariant gestures. Among the unconventional investigators who take such gestures to be the primitives of the phonetic system, there is some question about exactly how they should be defined, but little of what I mean to say here turns on the answer. What is most important for our purposes are just two considerations: one is that the gesture-as-primitive view permits us to see a system in which the defining gestures can be overlapped and merged (i.e., coarticulated) so as to produce phonetic strings at the high rates that are, in fact, achieved; the other, that the phonetically relevant gestures were presumably selected (and refined) in evolution because they lent themselves to just those articulatory and coarticulatory maneuvers that were appropriate to their specifically phonetic function. Accordingly, they form a natural class, a phonetic modality, as it were, that has a linguistic purpose and no other.

As for speech production, the conventional view is that it is controlled by mechanisms of a general motor sort, mechanisms that are constrained to produce exactly the sounds that define the consonants and vowels. According to the unconventional view, on the other hand, the mechanisms of articulation and coarticulation are not instances of some more general mechanism of motor control, but rather the workings of a biological specialization—a phonetic module—that is no less distinctly linguistic than the specialized gestures it manages. The aim of its specialized gestures is not to achieve particular acoustic targets, but to represent consonants and vowels invariently in rapidly produced strings, allowing the resulting sounds to go wherever the acoustically complex effects of coarticulation happen to take them. That the articulation of consonants and vowels is, in fact, a biological specialization is plainly shown by the inability of nonhuman primates to learn to produce even the simplest syllables. They can't do it, not because they are not smart enough, or because they lack the appropriate pieces of anatomy, but because, being other than human, they are not endowed with a phonetic module.

Turning now to perception, we see, on the conventional view, only the most general processes of the auditory modality, which is to say that perception of consonants and vowels is supposed to be no different from perception of other sounds. All use the same mode of signal analysis, and evoke in the same perceptual register the same set of auditory primitives. Thus, the difference in perception between a consonant and some nonspeech sound is only in the particular mix of auditory primitives they comprise. They are made of the same perceptual stuff.

According to the unconventional theorist, on the other hand, the phonetic gestures are recovered in perception by the specialized phonetic module that controlled their production. Such a specialized process is necessary in order that proper account be taken of the specifically phonetic complications that coarticulation introduces into the relation between acoustic signal and the gestural message it conveys. Given that the message and the process that recovers it are both specific to phonetic communication, the resulting representation is specific to that kind of communication, too, which is to say that its modality is distinctly phonetic, not auditory.

There is one more important assumption of the conventional view, this one made necessary by the prior assumption that speech rests on motor and perceptual representations of some general sort. For it falls to anyone who holds that speech is supported in that way to explain how its initially nonphonetic representations are invested with phonetic significance, and so made appropriate for linguistic communication. The conventional explanation is that this is done at a cognitive stage, beyond action and perception, where the very ordinary motor and auditory representations are translated into units of a linguistic sort. There are various notions among the conventional theorists about exactly how that is done, but those seem to be distinctions without a difference, for they all come, necessarily, to the same thing: speaker and listener must, in effect, attach phonetic labels to their respective nonphonetic acts and percepts; neither party can experience phonetic representations at the level of action and perception, because phonetic representations are not supposed to exist there.
The unconventional theory needs no such assumption as the one just described, because it
takes the primary representations of speaker and listener to be immediately phonetic; they are
precognitive acts and percepts, not cognitive afterthoughts.

The implications for reading and writing. From what has so far been said about the two
theories, it is clear that they see the relation of speech to reading/writing in drastically different
ways. They must, of course, agree that reading/writing are not supported by a biological
specialization at the level of act or percept—that is, in production or perception of the letters of
the alphabet—and, accordingly, that the letters can take on linguistic significance only by virtue
of being named after the consonants and vowels to which they have been arbitrarily assigned.
Given that area of agreement between the theories, the critical difference hinges, then, on the
clear implication of the more conventional theory that what is true of the link between signal and
language in reading/writing must be true in speech, too: the primary acts and percepts of speech
can be no more linguistic than those of reading/writing, and no less arbitrarily connected to
language. Thus, the conventional view reduces the difference between speech and
reading/writing to a matter of making or hearing sounds, in the one case, and drawing or seeing
print, in the other.

On the unconventional theory, however, the difference between speech and reading/writing is
profound. In contrast to the letters of the alphabet, the gestural representations that are the
input to the phonetic module in production and its output in perception are, by their very nature,
pieces of language, not arbitrary stand-ins. Accordingly, speaker and listener are immediately
engaged in the language business in a way that writer and reader are not; the difference between
making or hearing sounds, on the one hand, and drawing or seeing print, on the other—the only
difference the conventional view allows—has precious little to do with the matter.

Consider, now, the implications of the contrasting views of speech for the questions we
should answer if we would understand the reading process and the difficulties that some have
with it.

Are writing/reading hard? The answer given by the conventional view of speech is: not
really; no harder, certainly, than speech. That is, of course, exactly what the avatars of Whole
Language take as their most fundamental premise (Goodman, 1986). Indeed, it may very well be
the conventional theory of speech that initially emboldened them to promote a proposition that is
so at odds with the most obvious facts, and so harmful to an understanding of how we ought to
teach children to read. But then if they had really thought hard about the implications of the
conventional theory, they would have been led to the even worse conclusion, if that were possible,
that reading and writing must be, not just as easy as speech, but significantly easier. For if, as
the conventional view would have it, the difference is only that between sound-tongue-ear for
speech and print-finger-eye for reading/writing, then reading/writing has the advantage on all
counts.

Taking, first, the nature of the signal, one quickly sees the superiority of print. Typically, the
printed characters are crisp and clear; the signal-to-noise ratio could hardly be better. The
speech signal, on the other hand, leaves much to be desired from a physical point of view, if only
because much of the acoustic information that is most important for phonologic purposes is least
prominent acoustically. As for the effectors—fingers versus tongue—the fingers win, and by a
wide margin. For the moving finger writes, and having written, moves on to play Bach's
Goldberg Variations or do brain surgery; in contrast, the moving tongue speaks, and having
spoken, lapses into inactivity, except as it is occasionally called on to lick the lips or help in
swallowing. Imagine fixing a stylus to your tongue and trying then to write your name. Turning
finally to the receptors—the ear vs. the eye—I simply note that, as a channel for transmission of
information, the eye is better than the ear by several orders of magnitude. How, then, are we to
understand why it is that speech is, by every conceivable measure, the easier. Indeed, if
linguistic communication were as the conventional view of speech says it is, then our concerns
would be the exact reverse of what they are: having taken it for granted that reading and writing
are dead easy, the members of this conference would be exchanging ideas about how to teach
would-be speakers to overcome the difficulties caused by the evident limitations of tongue and
ear, and what to do for those who can't manage. The unconventional view does not blink those shortcomings, but rather shows how speech, in a triumph of evolution over engineering, found ways around them. Special exertions by speaker and listener are not called for. What that means will become clearer below, where we consider the requirements of phonological communication and how they are met.

What is hard about writing and reading? Surely, we can't know how to teach a child to read or write except as we understand what he has to learn and why the learning might not be easy. But, as we have seen, the conventional view of speech tells us that reading/writing should be even easier than speech, which we already know to be quite easy, so the conventional view is not likely to be helpful at this very earliest stage of our inquiry. Let us, however, overlook that most unfortunate implication of the conventional view, and put our attention instead on what it reveals about the well-documented difficulty of grasping the alphabetic principle. For that purpose, we must digress a bit to consider the nature of phonological communication and the requirements it imposes.

Everyone understands that the function of the phonologic mode of communication is to generate an uncountably large number of words by variously combining and permuting the small number of meaningless segments we call consonants and vowels. That is the combinatorial principle that allows to language its property of openness or generativity, a property that is unique among natural communication systems; not surprisingly, then, it is the design feature that characterizes language at all levels. But if the principle is to work at the level of phonology, two requirements must be met. The more obvious is that the segments be commutable, which is to say discrete, invariant, and categorical. The possibly less obvious requirement derives from the fact that, if all utterances are to be formed by a small number of segments, then, inevitably, those segments will run to long strings, so it becomes essential that their production and perception be expeditious.

Now, on the conventional view, it is sounds and the ordinary auditory percepts they evoke that must have those critical properties, from which it follows that speech could only be an acoustic alphabet, offering a discrete, invariant, and categorical sound (and auditory percept) for each phonetic segment. Of course, the sounds would presumably be smoothed and connected at the places where they join, much as the shapes of cursive writing are, but, one way or another, there would have to be, for each segment, a commutable piece of sound. To produce such sounds, a speaker would necessarily make a discrete articulatory gesture for each one, in which case he could not produce a syllable like 'bag,' but only the three syllables, 'buh ah guh.' If speech had to be like that, it would come nowhere near meeting the requirements for commutability and speed, so a communication system that was generative at the level of word formation would not be possible.

Nor would things be much better if a means could somehow be found to deliver the alphabetic sounds more rapidly, for that would surely defeat the ear. The point is that speakers normally produce phonetic segments at rates that average about 10 or 12 per second and, for short periods, run up to 20 or 25. Now if each of those segments were represented by a discrete sound, as the conventional view says it must be, then rates that high would strain the temporal resolving power of the ear, and also overreach its ability to keep the order of the segments straight (Warren, 1993).

But even if we put all of the foregoing considerations aside, and assume that speech as portrayed by the conventional view could somehow be made to work, we are still not at all enlightened about why the alphabetic principle should not have been almost immediately apparent to those who lived before it was discovered, and why it is not equally apparent now to every normal child. All would have mastered a language that was conveyed, presumably, by an acoustic alphabet. Why, then, would they not already understand the alphabetic principle, and quickly learn to apply it in the visual modality simply by substituting the alphabetic letters for the correspondingly alphabetic sounds?

But if the conventional view leaves us puzzled about why it is hard to be aware that words come apart, it does suggest why some teachers might misunderstand how they are put together.
The misunderstanding manifests itself, and also begins to take its toll, when, having taught a child the 'spelling-to-sound rules,' the teacher urges him to 'blend'—that is, to form the alphabetic sounds, 'buh, ah, and guh,' for example, into the proper word 'bag.' I cannot presume to know what is in the mind of the teacher who tries to get the child to do that, but I suspect it is a resolution of the apparent conflict between what she believes about speech and what her ears tell her about how it sounds. With encouragement from the conventional view, she presumably believes that there are three sounds in 'bag,' and that these are represented by the letters 'b,' 'a,' and 'g.' However, I should think she would find it unsettling that she can't really hear three sounds, but rather something that is, from a purely auditory point of view, all of a piece. Perhaps, then, she supposes that the auditory appearances are deceiving, that the three sounds have been so thoroughly blended as to hide their individual identities. If so, then she is using the word 'blend' in its correct sense to mean a combination in which the constituent parts are indistinguishable; but she is imagining a most unfortunate contingency, for blending would cause language to lose its vital phonologic core, with the result that the combinatorial principle would no longer be available to produce vocabularies that are large and expandable. In any case, it is physically and physiologically impossible to produce a word by blending, or otherwise combining, the discrete sounds that are taken to be its individual phonetic constituents. So while 'blend' is the right word, it is the wrong idea.

I do not mean to suggest that implying to a child that 'bag' is a blend of three sounds is necessarily to court failure in reading and writing. It is rather to tell a white lie, and is better by far than characterizing the printed word as a picture, or advising the child to guess what the print says. Learning letter-to-sound correspondences and trying, on that basis, to 'sound out' words is likely, at least, to help bring the child to the correct understanding that words come apart, and that the alphabet has something to do with the parts. The error is in the belief that the parts are sounds. Most, but obviously not all, children who are taught that error manage somehow to rise above it, and so learn to read and write. Still, things would almost certainly go better if they were acquainted with the true state of affairs.

In contrast to the conventional theory, the unconventional account of these matters shows how phonological communication is possible, and, by the same token, why the alphabetic principle is hard to grasp. Remember that speakers are able to produce strings of phonetic segments at high rates, but only because the segments are gestures that are efficiently overlapped and merged. By that means, the speaker succeeds in producing phonologic structures that effectively 'spell' the words they convey. But there are several reasons why the illiterate or preliterate speaker nevertheless does not know how to spell, or even that words have a spelling. Perhaps the most obvious is that the phonetic module spells for the speaker. Once he has thought of the word, whatever that means, the phonetic module takes over, automatically selecting and coordinating the appropriate gestures. The speaker cannot know how the module did what it did, because it is true of all biological modules that their processes are not available to conscious inspection. On the other hand, the speaker can be aware of the representations the phonetic module deals with; but there is no reason he should be, because being inherently phonetic, the motor structures that are represented do not require translation, so they do not invite attention. And, finally, it is probably relevant that the mechanisms of articulation and coarticulation produce smoothed and context-sensitive movements at the surface, and so obscure the exact nature of the distal motor structures that are the actual phonetic units.

The relevant considerations are much the same in perception. There, coarticulation has allowed information about several successive segments to be conveyed simultaneously in the acoustic signal, and so relaxed the constraint on rate of perception imposed by the temporal resolving power of the ear. The listener can, therefore, keep up with the speaker, but only because his phonetic module is specialized to process the acoustic signal so as to extract the coarticulated gestures that produced its uniquely phonetic complications. A consequence is that the listener is likely to lack phonologic awareness for much the same reasons that keep a speaker in the dark. Though the signal is, in fact, parsed into its phonetic constituents, the listener is none the wiser, because the module runs on automatic in perception just as it does in production.
Deliberate, cognitive procedures are never necessary to do the job. Indeed, the job cannot be done
cognitively, because the complexities of the speech code are apparently too great, too special to
language, and too deep in our biology; certainly, no one has succeeded yet, though, given the
intense and long-continued efforts to build an automatic speech recognizer, we know it is not for
want of trying.

As for the representations that are the result of the module’s efforts, they are already
phonetic, as I’ve said so many times, hence perfectly appropriate for all further linguistic
processing. Therefore, the listener does not have to give them the attention they would require if,
like the letters of the alphabet, they had to be translated into pieces of language. Finally,
coarticulation has destroyed anything that remotely resembles a straightforward relation
between the segments of the phonetic message and such segments as can be found in the acoustic
signal. A consequence is that the consonants, at least, have never in the listener’s experience
been isolated and pointed to, as words, for example, commonly are. Surely, that is one reason
why preliterate children are more likely to be aware of words than of the phonologic segments
that form them. None of this is to say that listeners cannot be aware of the phonologic
constituents of words—indeed, if they could not, the use of alphabetic transcriptions would be
impossible—only that the unconventional view shows us why such awareness does not come for
free with mastery of speech.

What links writing to reading and speaking to listening? In linguistic communication, where
every sender is a receiver and every receiver a sender, the processes of production and perception
must somehow be linked. Mattingly and I have called this the ‘requirement for parity,’ and
wondered how it is met (Liberman, 1996; Liberman & Mattingly, 1989).

In reading/writing, parity cannot be said to rest on a primary biological base, but must rather
have been established by agreement. Somehow, those who developed an alphabet arrived at a
compact that specified which optical shapes were relevant to language, and which piece of
language each was relevant to. A result is that learning to read and write is largely a matter of
mastering the arbitrary terms of that compact, and, for all the reasons the unconventional view
has revealed, that is rather hard, and commonly requires help from a tutor.

In the matter of parity, as in all things, the conventional view of speech implies that as it is
in reading/writing, so must it be in speech, in which case learning speech has got to be just like
learning to read and write. So here again we find in the conventional view of speech full
justification for one of the fundamental, and fundamentally wrong, assumptions of Whole
Language—namely, that children should learn to read as they learned to speak, which is to say
that the educational process should be geared to provide conditions just like those under which
speech was acquired; children need not, and should not, be taught to analyze the language as
linguists do (Goodman, 1986).

The unconventional view, on the other hand, claims that parity in speech does not derive
from a compact of some kind, but rather reflects a fundamental aspect of our biology. For parity
is exactly what evolved: the necessary link between production and perception is given
immediately by the genetically determined phonetic module, which provides that the specifically
phonetic motor structure in the mind of the speaker is reproduced in the mind of the listener;
there is no need for the two parties to connect grossly dissimilar but equally nonphonetic acts
and percepts that were, like the letters of the alphabet, selected by earlier generations and then
arbitrarily assigned to phonetic categories. Thus, the phonetic module makes for a deep and
immediate intimacy between speaker and listener, an intimacy as necessary for linguistic
communication as that which sex affords is necessary for reproduction. An important difference,
of course, is that the one proceeds from parity, while it is disparity that lies at the root of the
other. But an equally important similarity is that both kinds of intimacy are the products of co­
evolution, since the two sides of the connection had, in each case, to develop in step, change for
change, else neither system could ever have become functional.

Though parity in speech is part of its underlying biology, it does not follow that speech is not
learned, only that it need not be taught. For surely, the necessary and sufficient conditions for
learning speech are but two: membership in the human race, and exposure to a mother tongue.
To get an idea of the nature of that kind of learning, it is helpful, I think, to see the phonetic module as one of a class of modules that have certain characteristics in common (Liberman, & Mattingly, 1989). One of those is plasticity over periods of time during which the module is shaped by environmental conditions. An example is the module for sound localization, which responds to interaural differences of time and intensity, using them as a basis for computing and then representing location in azimuth. Of course, those interaural differences change considerably as the head grows, and the distance between the ears increases, so the module must be continuously recalibrated. One might reasonably suppose that, in a somewhat similar way, the biologically coherent phonetic module is calibrated over several years by the phonetic environment in which it finds itself. In that case, the obvious effect of experience on speech would be to shape or hone a genetically determined system, not, as in the case of reading/writing, to provide the basis for acquiring arbitrary connections by processes of a cognitive sort. Given a normal environment, speech 'emerges,' in the terminology of Whole Language, but reading and writing most certainly do not. Thus, the unconventional view permits us to see that learning to speak and learning to read or write are fundamentally different processes.

Implications for reading disability. Accepting the conventional view of speech, the reading/writer researchers must believe, as I earlier said, that the phonologic segments are plainly displayed on the auditory surface, there for all to hear. Accordingly, such researchers have no way to see why it should be hard to be aware of those segments, and so to grasp the alphabetic principle. They can hardly be expected, then, to look to that difficulty for the causes of reading/writing disability, and indeed, they do not. Rather, they look where the conventional view most directly tells them to, which is at some aspect of the visual system. That system is the seemingly most promising target, because the substitution of eye for ear is virtually the only important difference between speech and reading/writing that the conventional view allows. So if reading proves to be hard, then it must be that some aspect of vision is at fault. Small wonder, then, that one or another aspect of visual function is the place where many of the theories of disability locate the problem (Geiger & Lettvin, 1988; Orton, 1937; Pavlides, 1985; Stein, 1988).

At the same time, the conventional view permits, if it does not actually encourage, the belief that the problem might be with the ear. That belief begins with the conventional assumption that speech is a string of brief sounds that follow each other in rapid succession. The problem, then, is that the auditory system of some children can't keep up. As a consequence, they have language problems, from which reading problems follow (Tallal, 1980). Now if speech were a string of acoustic segments, one for each phonologic segment, then it would be true that the relevant sounds would, indeed, be very brief, and would follow each other in rapid succession—so brief and in such rapid succession, that, as I earlier pointed out, sound segments would come along at rates between 10 and 25 per second. It is also true, as I said in the same context, that rates that high would strain the ability of everybody's auditory system, not just those of some unfortunate children. Fortunately, speech does not require people to do what their ears do poorly, which takes us now to the unconventional view and its radically different implications for how we might see disability in reading/writing.

Let us consider first the theory about disability just alluded to. On the unconventional view of speech, the known limitation on ability to perceive the order of brief sounds presented rapidly and in series is irrelevant to speech perception, because phonetic segments are not sounds, and speech is not a string of them. The unconventional view tells us that the true phonetic elements are gestures, and that their coarticulation smears the information for each one over a considerable stretch of the acoustic signal, overlapping it grossly with information for other segments. One important consequence is that ordinal position is marked, not by the temporal order of the sounds, but by their acoustic shapes. Thus, the syllables ba and ab, when pronounced briefly, have acoustic patterns in which information about consonant and vowel are completely overlapped. Accordingly, there are not two acoustic segments—one for each phoneme—hence no way to perceive which came first by paying attention to the way sounds succeed each other in time. Nevertheless, the listener infallibly knows which came first and which second because the acoustic shapes of the two syllables are very different. In fact, in that case, they are
exact mirror images. Given the services of the phonetic module, which are always at the disposal of the listener, the one shape is perceived as an opening gesture (consonant first, vowel second), the other as a closing gesture (vowel first, consonant second).

Having just seen what the unconventional view says is not true of speech, and therefore of a theory that locates the cause of reading/writing disability in the ear, I turn now to what it says is true, and how that gives us an entirely different slant on the probable causes of failure. We earlier saw how the unconventional view shows that learning to speak, however fluently, will not be sufficient to produce awareness of phonologic structure. Acting on precisely that consideration, Isabelle Liberman, Donald Shankweiler, Ignatius Mattingly, and their colleagues began the line of thought that led them to find that phonologic awareness is, in fact, largely absent in preliterate children (Liberman, 1973; Mattingly, 1972; Liberman, Shankweiler, Fischer, & Carter, 1974). Subsequent research by them and others amply confirmed that finding, while also establishing that the extent to which awareness is present counts as one of the best predictors of success in reading/writing (for reviews, see Blachman, 1989; Routh & Fox, 1984), and that training in awareness has generally happy consequences for those who get it (Bradley & Bryant, 1983; Content et al., 1986; Ball & Blachman, 1988; Lundberg, Frost, & Peterson, 1988; Olofsson & Lundberg, 1983; Vellutino & Scanlon, 1987).

Proceeding further with the implications of the unconventional view, researchers picked up on its assumption that there is a distinct phonological faculty—I have here called it a phonetic module—that is independent of cognition and, indeed, of all other-than-linguistic modes of production and perception. They found it reasonable to suppose that if such a faculty exists—though the conventional view provides no place for it—then it might work more or less well among otherwise normal children, with the result that there would be differences in the ease with which they would learn to read and write (Liberman, I. Y., & Liberman, A. M., 1990; Liberman, A. M., 1992; Liberman, I. Y., Shankweiler, D., & Liberman, A. M., 1989; Brady, S. A., 1991). Most obviously, the effect would be on the general quality or clarity of the phonologic representations, which would, in turn, affect the child's ability to become aware of them, and so to comprehend and apply the alphabetic principle. One would expect, too, an effect on the phonologic basis of the working memory that is an integral part of syntactic processing, and therefore on the child's ability to comprehend at the level of the sentence. Indeed, everything about language or reading/writing that depends on phonologic structures and processes would presumably be affected in some way. Exactly how, and with what consequences, are questions that motivate the research our colleagues are now actively pursuing. What is reasonably clear at this point is only that the leads afforded by the unconventional view are promising, and that the relevant research on the role of specifically phonologic processes is nearer its beginning than its end. It is very hard, I think, to see how we should ever have arrived at that beginning if researchers had remained true to the conventional view of speech. On the other hand, the hypothesis that phonologic factors deserve careful attention is now common enough that researchers may have lost sight of what the unconventional view of speech had to do with it.

REFERENCES


**FOOTNOTE**