History of Psychology
2015, Vol. 18, No. 1, 78–99

1093-4510/15/$12.00 http://dx.doi.org/10.1037/a0038299

SEEKING A READING MACHINE FOR THE BLIND AND DISCOVERING THE SPEECH CODE

Donald Shankweiler and Carol A. Fowler
University of Connecticut and
Haskins Laboratories, New Haven, Connecticut

A machine that can read printed material to the blind became a priority at the end of World War II with the appointment of a U.S. Government committee to instigate research on sensory aids to improve the lot of blinded veterans. The committee chose Haskins Laboratories to lead a multisite research program. Initially, Haskins researchers overestimated the capacities of users to learn an acoustic code based on the letters of a text, resulting in unsuitable designs. Progress was slow because the researchers clung to a mistaken view that speech is a sound alphabet and because of persisting gaps in man-machine technology. The tortuous route to a practical reading machine transformed the scientific understanding of speech perception and reading at Haskins Labs and elsewhere, leading to novel lines of basic research and new technologies. Research at Haskins Laboratories made valuable contributions in clarifying the physical basis of speech. Researchers recognized that coarticulatory overlap eliminated the possibility of alphabet-like discrete acoustic segments in speech. This work advanced the study of speech perception and contributed to our understanding of the relation of speech perception to production. Basic findings on speech enabled the development of speech synthesis, part science and part technology, essential for development of a reading machine, which has found many applications. Findings on the nature of speech further stimulated a new understanding of word recognition in reading across languages and scripts and contributed to our understanding of reading development and reading disabilities.

Keywords: reading machine envisioned, interplay of science and technology, discovering the speech code, synthesizing speech from print

Since the 19th century, a movement has existed to help blind people better their educational opportunities and enable them to achieve greater independence (Zahl, 1950). Inability to read is obviously one of the greatest obstacles to education of the blind, which has led to devising various remedies. This article traces a late chapter in the history of attempts to exploit optic and electrical technology to develop a surrogate reading device. Machines that can in some fashion enable blind people to read printed material date from the early years of the 20th century. These were optical scanning devices, based on the photocell; they produced sounds that varied as different letters from a line of print came under the scanner. This approach to a reading machine was eventually abandoned because the devices proved difficult to learn to use and permitted only very slow rates of word recognition. A committee of scientists was commissioned by the defense science directorate of the U.S. Government in 1944 to seek effective sensory devices for the blind and to revive the reading machine goal. The scientists shared the belief that developments in technology spurred by the war effort would make tractable the problems of design and implementation that had

This article was published Online First December 22, 2014.

Donald Shankweiler and Carol A. Fowler, Department of Psychology, University of Connecticut, and Haskins Laboratories, New Haven, Connecticut.

The authors are grateful for the advice of several colleagues on earlier drafts: Brian Byrne, Leonard Katz, John Lankford, Arthur Abramson, Brian Leahy, Laura Palumbo, Philip Rubin, Michael Sokal, Michael Studdert-Kennedy, and Douglas Whalen. We thank Yvonne Manning for remaking the figures.

Correspondence concerning this article should be addressed to Donald Shankweiler, or Carol Fowler, Department of Psychology, University of Connecticut, 406 Babbidge Road Unit 1020, Storrs, CT 06269-1020. E-mail: donald.shankweiler@uconn.edu, carol.fowler@uconn.edu

frustrated the earlier attempts. The committee chose Haskins Laboratories, a small, nonprofit laboratory then in New York City, to develop a multisite research program to devise a more satisfactory reading machine.

This article is a case study in the intertwining of science and engineering technology under the stimulus of a specific and elusive goal. It recounts how the project was conceived in 1944 as a problem in application of known principles and available technology. It explains why efforts in the first years led to failure and how the project was transformed leading to unanticipated developments in science and its applications over the next 40 years. The difficulties in developing reading machine prototypes that humans could learn to use revealed large regions of ignorance of human perceptual capabilities and had broad repercussions for cognitive science. Encountering these difficulties served as a catalyst at Haskins Laboratories for asking and answering questions about the nature of speech, how it evolved, and how it functions so well for communication of language.

Following Thomas Kuhn's (1962) proposal regarding the structure of revolutions in scientific thought, some writers on the nature and varieties of knowledge in engineering design (e.g., Constant, 1980; Vincenti, 1990) have invoked a distinction between "normal design," following accepted practices in design of an artifact in which application of established scientific and engineering principles suffices to address design problems, and "radical design" (the terms are Vincenti's) where the design problem requires a thoroughgoing reformulation, transcending or overturning established principles and theory. Repeated failures in creating a successful design for a sought-after artifact can be a symptom of a problem requiring radical design (Petrosky, 2006). During the first years of the Reading Machine Project, Haskins Labs researchers considered that a modernized and improved version of an existing device, the Optophone (see), would do the job (Cooper, 1950). But after some years of failed efforts along these lines, it became clear to the researchers that a satisfactory reading machine for the blind would require new fundamental knowledge, knowledge of speech processes and how to simulate them, that was not yet on the horizon (Cooper, Gaitenby, & Nye, 1984). They then came to appreciate that the reading machine goal imposed a radical design problem.

As it evolved, the reading machine history has proved to be one of continuing cross talk between technology and science. Layton (1971) argues that technologies and sciences are distinct domains and that that findings in each domain most typically feed further research in its own domain. Even so, he notes (Layton, 1974) that there can be cross-fertilization between technology and science at all levels. The history of research to develop a reading machine provides one example of this less common, but far from unique, interdomain cross fertilization, in this case from technology to science. A related example of a technological, engineering innovation that enabled scientific discovery is the sound spectrograph developed at Bell Telephone Labs in the early 1940s (described). The spectrograph greatly extended the possibilities for investigating the acoustics of speech and other mechanical, sound-producing environmental events, and for determining the relations of specific features of acoustic signals to their mechanical sources. Likewise, the Haskins Labs Pattern Playback, that enabled reconversion of spectrograms to sound (discussed), made possible a new experimental approach to speech perception and speech synthesis. Another pertinent example is electromyography, a case of biomedical engineering technology that further contributed to fundamental developments in speech science.[1]

Finally, the history of the reading machine underlines the social aspect of technology. The impetus for technology is usually a social need. Technological innovation is constrained by the needs and desires of potential consumers, and the success of its products is determined by how well they match consumer capacities and how well they meet consumer expectations. For most of the history of reading machine research, there were persisting gaps between what the technology offered and the needs of blind users.

[1] Electromyography is a procedure that uses electrodes inserted into muscles or positioned on skin surfaces over them to detect electrical activity in contracting muscles.

Haskins Laboratories Oversees Research on Sensory Devices and Conducts Research on Reading Machines

A few words about Haskins Laboratories and the circumstances that led to its taking on the Reading Machine project: The Laboratories had been founded in 1935 by Caryl P. Haskins, a recent Harvard PhD (biophysics) and a pioneer in the new field of radiation biology. From the beginning he was joined by Franklin S. Cooper, a physics PhD (MIT), who was also an electrical engineer. These young scientists had met at the research laboratories of the General Electric Company in Schenectady, New York, where both were employed briefly in the mid 1930s. Wishing to continue their then pioneering research on the effects of radiation on the physiology and growth of animal and plant tissues after the project was discontinued at General Electric, Haskins and Cooper formed a plan to join forces. The time, during the Great Depression years, afforded few alternatives for pursuing novel scientific work of this kind. Haskins drew on a modest inheritance to fund a small private nonprofit laboratory based in Schenectady. The Laboratories survived, and by the end of the decade, modest additional funds became available from private sources (including funds provided by Alfred Loomis, a Wall Street financier and founder of Loomis Laboratory in Tuxedo Park, New York, and Philip Pillsbury of the Pillsbury Co., Minneapolis; Nye, 2006, p. 104) allowing the Laboratories to relocate to New York City. An information pamphlet about Haskins Laboratories (Haskins Laboratories, 1953) reflecting the philosophy of its first 18 years of operation states that it was created “for basic research and research training in certain pioneer areas which involve several scientific disciplines . . .” with the potential to become accepted “growing points” of research knowledge, then to leave them for others to cultivate, and move on to identify and advance new areas. Haskins and Cooper were joined by a few other young scientists who were attracted by this philosophy. Two who would become well known for path-breaking research on metabolic processes of microorganisms were Seymour Hutner and Luigi Provasoli.

At the end of 1940, with World War II looming, Haskins and Cooper moved to Washington to work for the National Defense Research Committee (NDRC), a group of prominent researchers assembled by Vannevar Bush, who was appointed by President Roosevelt to direct and coordinate scientific activities in support of the war. Later the NDRC was subsumed in the Office of Scientific Research and Development (OSRD, with Bush as overall director. Haskins became deputy director of the NDRC (under James Conant), and Frank Cooper became liaison officer to the Committee (Hauger, 1995; Nye, 2006).

In mid-1944, the OSRD began to turn its attention from the war effort to preparing for its aftermath. One object was to anticipate the problems of reintegration into civilian life of returning veterans who were blinded by war wounds and to consider how some of the new developments in science and technology could be marshaled to ameliorate these problems. OSRD director Bush appointed a Committee on Sensory Devices (CSD), to be headed by George Corner, Dean of Medicine at the University of Pennsylvania to guide research on sensory aids at several laboratories which would receive federal funds (Hauger, 1995; Nye, 2006).[2]

Bush suggested Haskins Laboratories to conduct and oversee research on guidance devices for the blind and to undertake the development of a reading aid that could make ordinary printed material widely available to blind people, a need not met by braille or recorded books. Hence, a reading machine was a priority even though these other means existed to give access to written material. Braille relies on specially prepared materials scanned by touch and is difficult to master. Moreover, braille books are bulky and difficult to disseminate. Recorded books, though they require no new skills, also have limited offerings. Neither alternative

[2] Other CSD members included (from Hauger, 1995): Henry Barton, a physicist serving as Director of the American Institute of Physics and Vice Chairman of the Division of Physical Sciences, National Research Council; Anton Carlson, a retired physiologist from the University of Chicago; Wallace Fenn, a professor at the Professional School of Medicine and Dentistry, University of Rochester; Stacy Guild, Director of the Otological Research Laboratory at Johns Hopkins University; and Karl Lashley, Director of the Yerkes Laboratories for Primate Biology in Orange Park, Florida.

meets the needs of students or any blind person who wishes to have unlimited access to print.

Bush recommended the Laboratories for this role at least in part because it fit his ideal of a "small, compact working group, operating on a nonprofit, noncommercial basis" (instructions from Vannevar Bush to George Corner, January 7, 1944, quoted in Hauger, 1995, p. 22.[3] In mid-1944, Caryl Haskins and Franklin Cooper resumed work in New York, although Haskins himself remained based primarily in Washington.[4]

The research at the Laboratories on guidance devices and the reading machine for the blind became the primary responsibility of two of Caryl Haskins' partners at the Laboratories, Franklin Cooper and Paul Zahl. They recruited others as the research got underway. It was clear that new work on sensory aids for the blind would be concerned largely with man-machine interactions, requiring expertise in experimental psychology. A recruit from 1944, who perhaps more than any other individual was to leave his personal stamp on the research program stemming from the reading machine, was Alvin M. Liberman, a young psychologist, with a newly minted PhD from Yale.

The First Phase of Work at Haskins Labs on the Reading Machine: Dashed Hopes for Sound Alphabets[5]

Cooper and colleagues understood at the outset that a reading machine designed to cope with existing printed materials would have two essential components, an input device, a photoelectric scanner, to detect optical patterns in print and a coordinated output device to convert those patterns to others in another perceptible modality. Early on Cooper (1950) suggested that conceivable machines, whether or not they were feasible at the time, could be classified along two dimensions. On the input side, the possibilities were direct translation or recognition machines. Direct translation machines would detect optical properties of printed letters without recognizing letter identity as such and use the letter properties to control an output signal. Recognition machines would identify letters and use the letters identified to drive the output. The difference is in whether it is the job of the listener or instead the machine to identify the letters of the text. On the output side, the patterns produced, if they were acoustic, might be nonspeech patterns or, for recognition machines, might be letter names, letter sounds, or spoken words. In the early years, only direct translation machines were feasible. Cooper (1950) noted that a recognition machine would require a suitable optical character recognizer, which was beyond the engineering capabilities of that time.

Cooper's plan was not to engineer a complete reading machine. Rather, it was to concentrate on the output component. This part of the device interfaces with the human user and was the aspect of the problem that Haskins researchers considered themselves best suited to solve. In Cooper's recollections in 1988 (Nye, 2006, p. 72), ". . . the philosophy of the whole enterprise was that you don't build devices where you can simulate the output and test the output. We were quite in agreement that *the acceptability of the signal to the subject was a critical point*" (italics added). They chose to develop acoustic outputs, undertaking a search for intelligible sounds; another group (Ra-

[3] In view of Caryl Haskins' leadership position at the NDRC, the selection of the laboratory he founded to serve this role has the appearance of a conflict of interest. Indeed, George Corner expressed to Bush some reservations along those lines (Hauger, 1995, p. 23). However, Haskins Laboratories did fit Bush's ideal closely. Moreover, the Laboratories assigned all patents to the government and did not take profits for the work (Hauger, 1995). According to Hauger (1995, p. 24), "there is no evidence that [Caryl] Haskins or Bush had any desire for personal benefit from the project."

[4] In fact, Caryl Haskins did not long involve himself directly in the research of the Haskins Laboratories after the war, although he continued to serve the larger science research community in several capacities (e.g., as president of Sigma Xi, the science academic honorary society), and he independently nourished and vigorously pursued biological research on insect evolution and ecology. In 1956, he began a long term as president and scientific director of the Carnegie Institution of Washington (where he succeeded Vannevar Bush).

[5] The early research on the reading machine was not published as it was ongoing. For information about the earliest period of research on the reading machine at Haskins Laboratories, we relied on a book sponsored by the Committee on Sensory Devices and the National Research Council summarizing progress of research on sensory aids as of 1947 and placing it in the context of other work in aid of the blind community (Zahl, 1950). We drew especially on the chapter on reading machine research prepared by Cooper. We also made use of later retrospective reviews by Cooper et al. 1984; Hauger, 1995; Liberman, 1996, and Nye & Bliss, 1970, and we drew upon extensive oral history tape recordings, giving recollections by the founders and other early researchers. These were made at Haskins Laboratories in 1988–1989 and transcribed by Patrick Nye (2006).

dio Inventions, Inc.) supported by the CSD tackled tactile outputs requiring braille-like reading skills.[6]

At the outset, Cooper and Liberman surmised that a direct translation device having discrete nonspeech sounds as output would provide acceptable signals for reading machine users as long as the sounds were discriminable. Such an output would be a kind of alphabet, a sound alphabet in analogy to an optical alphabet, which is a highly successful stand-in for speech. Moreover, they mistakenly believed initially that speech itself was composed of discrete alphabet-like acoustic segments for phonemes so that, plausibly, substitution of one acoustic alphabet for another should work well (e.g., remarks by Liberman in Nye, 2006, p. 71).

An existing reading machine, the Optophone, reflected these beliefs. It was invented in England in 1912 by Fournier d′Albe (1920), exploiting technology that was advanced for the time (Nye & Bliss, 1970). Improvements were made over the next 10 years, and the device went into commercial production for a time. A book was placed on a glass plate and scanned with five vertically positioned spots of light emitted from a narrow slit. Each spot was modulated at a different audio frequency (Figure 1). When a spot intercepted a letter part, a musical tone was sounded having a frequency associated with that light spot. Accordingly, a user heard combinations of chordal tones (harmonica-like) that changed rapidly as different letters passed beneath the slit.

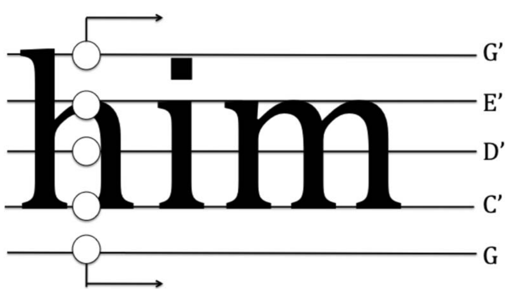

Figure 1. Adapted from Cooper (1950). The figure shows how the Optophone scanned a line of text. The scanning spots moved horizontally over a line of text. Where black print was encountered by a spot, a tone at the associated frequency was produced. The tones are labeled by their musical names on the right of the figure. Sometimes, as depicted just one tone was emitted. More commonly two or more tones generated a chord.

Contemporary accounts (e.g., Jameson, 1932) indicated that people could learn to use the Optophone to identify words of a text although at very slow rates for the most part (Cooper, Gaitenby, & Nye, 1984). Even after much practice, most users read fewer than five words per minute, and without full accuracy. However, there was a blind woman who used and demonstrated an Optophone for many years and who could read as many as 60 words per minute. By way of comparison, however, an average sighted adult reader of ordinary print can read in the neighborhood of 200 to 300 words per minute (e.g., Crowder, 1982).

Shortly after Haskins Laboratories took on the reading machine project, Cooper traveled to London to study the Optophone, by then a museum piece, and learn more about how it had performed for its users. On his return, he borrowed an Optophone from the American Foundation for the Blind and restored it to working order. This enabled Haskins researchers to compare users' performance with the Optophone to their performance with the expanded acoustic coding possibilities that were being devised under the CSD program. They found that Haskins subjects performed using the Optophone at about the same slow rate as had been reported for most users in England. The main difficulties were in identification of certain letters, especially a,e,o,c, for which the chord sequences, like the letters themselves, were much alike (Cooper et al., 1984). Even prolonged training did not enable users to overcome this problem.

At about the time Haskins researchers began to study the Optophone, they fashioned their own device for transforming visual patterns into auditory patterns (Cooper et al., 1984; Liberman, 1996). Like the Optophone, it had a basic optical-electric component in which photocells were placed behind a narrow slit. These were coupled

[6] Other research groups under contract with the CSD were RCA Laboratories, for developing an improved Optophone, an early reading machine (described later in this section), and for work toward a letter recognition device, and Radio Inventions, Inc., which redesigned the Visagraph, a device invented in the 1930s which could accept existing printed material, but transformed it into raised characters that the blind could perceive by touch. The Naval Medical Research Institute and the Institute of Human Adjustment of the University of Michigan conducted independent tests of some of the pilot devices from Haskins and the other laboratories (Cooper, 1950).

via appropriate circuitry to a sound source that produced tones modulated by moment-to-moment changes in the optical pattern as print material moved across the slit. Instead of building a mechanically operated scanning system with a fixed scanning rate such as that of the Optophone, Cooper fashioned a device that passed a film negative of the text across a fixed slit using a modified 16-mm movie projector. With this system, it was possible to vary the rate at which the text moved across the slit, enabling the experimenter to measure the effect of scanning rate on a subject's ability to identify the words conveyed by the acoustic code. Furthermore, the Haskins scanning device was equipped with circuits suitable for creating a variety of acoustic outputs. Researchers discovered that amplitude variation did not enable users to recognize the letters whereas frequency variation was more promising. (See Cooper's (1950) summary of these unpublished findings.) However, researchers were prepared to test whether more complex nonspeech signals that had several perceived dimensions of variation, such as pitch, loudness, timbre, duration, and rhythm would prove even more discriminable and learnable.

To determine which acoustic signals would be easiest to distinguish and learn by naïve listeners required that Haskins researchers develop appropriate listening tests. Psychologist Alvin Liberman devised a series of tests. The simplest was a list of eight four-letter words, meant to enable rapid elimination of outputs that were not viable. In addition to word lists, more varied and demanding listening tests involving sentence material were contrived. Testing involved presentation of words and sentences mechanically scanned at rates from 50 to 150 words per minute.

During the last half of the 1940s, Haskins researchers worked intensively to create auditory output signals that were compact, multidimensional, and complex. However, with the exception of one sighted participant who was found to be peeking beneath her opaque goggles, subjects never attained high accuracy levels on any of the varieties of acoustic alphabets that Haskins researchers tried, and they read at rates between 4 and 10 words per minute, nowhere near 100 words per minute that Cooper assumed would be required for a practically useful reading machine (Cooper, 1950; Cooper et al., 1984).

Aside from the crucial problem that letter-by-letter decoding was very slow, another difficulty arose when materials were presented to simulate faster scanning rates. It was that the letter sounds merged auditorily; they did not maintain their discrete identities (e.g., Cooper, 1950; see also Liberman, 1996, p. 5), explaining why performance at faster scanning rates declined precipitously. This merging also meant that what had been learned at a slow scanning rate would not transfer to faster rates. It occurred to the Haskins researchers (Cooper, 1950, pp. 523–524; Cooper et al., 1984 p. 55; Studdert-Kennedy & Liberman, 1963) that the nonspeech acoustic alphabets they had tried might be doomed to fail as a class because of their discrete nature. Presented at rapid rates, perhaps they were exceeding the limits of the temporal resolving capabilities of the ear. That is, inherent limits on auditory resolution would set an upper bound (of about 12 to 15 per second) on the rate at which discrete sounds could be heard as distinct, separate, and ordered (Miller & Taylor, 1948). These limits likely underlay the phenomenal merging of the discrete sounds at rapid scanning rates.

On this reasoning, maximal reading rates should be quite slow for any acoustic alphabet. With the Optophone there are about 3 chords per letter. A presentation rate of 5 or 6 letters per second (i.e., one average English word per second), or about 60 words per minute, would be an absolute upper limit on letter-by-letter reading. Use of Morse, the international telegraph code, falls within these limits. But for many projected uses of a reading machine this was too slow, being close to the rate at which Cooper (1950) judged that the machine would serve more as a toy or curiosity than a useful tool. Moreover, in fact, the actual performance rates of users were far lower (Cooper et al., 1984, p. 55).

Of course, this raised for Cooper and Liberman the challenging question of how speech itself evades the limits of the temporal resolving power of the ear (Cooper, 1950; Liberman, 1996). The answer that awaited Haskins researchers' later explorations is that speech is coarticulated; consonants and vowels overlap in long sequences of continuous, rather than temporally discrete, acoustic patterns. Although Cooper (1950) does not use the term "coarticulation" (nor do any early papers by Haskins researchers, e.g., Liberman, 1957; Liberman et al., 1952, 1954), he does recognize that acoustic signals for words provide many fewer "distinctive elements" than do words in Morse code. Later papers (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967) refer to "parallel transmission" and "restructuring."

A clue was available that there was something special about the way that speech packages successive consonants and vowels. Just one of a number of candidate acoustic outputs that Haskins researchers tested proved highly learnable by naïve listeners. This output was an enciphered version of speech. A talker spoke the enciphered forms of test words in which each vowel letter was replaced by another vowel and each consonant by another consonant. This system was called *Wuhzi*, after the enciphered form of *have*. Of course, the talker coarticulated or overlapped the substituted phonemes just as he would in producing standard English words. In a test comparing the learnability of the nonspeech acoustic alphabets and Wuhzi, Liberman (1996. p. 5) reports that Wuhzi was "in a class by itself." Hearing Wuhzi, subjects learned to read the test materials (eight four letter words) to perfection in about 15 trials, whereas with the other systems, performance trailed markedly even after many more trials.

In retrospect, the lesson here should have been that there is something distinctive about speech itself that sets it apart from the other codes attempted in this early Haskins work. However, Cooper and colleagues did not yet conclude that speech was distinctive in this respect. The lessons that *were* learned from the period of CSD support during the 1940s were more in the nature of what would not work than what would. In summarizing the understanding at that time, Cooper and his associates (1984, p. 60**)** quote from the final report to the CSD in mid 1947: "One of the principal conclusions to be drawn from the work done thus far is that a successful reading machine must present its information in word-like units, not letter-by-letter."

Later in 1947, the OSRD, which had created the CSD in 1944 and provided funds for its activities, was dismantled as the nation gradually resumed a peacetime footing. At about that time, the CSD itself was dissolved, bringing to an end funding for the reading machine project and other sensory aids projects. The reading machine, in particular, had come to be seen as a long-term goal that would require a prolonged period of basic research along several fronts, needs which could not be supplied by a temporary agency. Some of these needs were met later by the Department of Veterans Affairs, but there was a 10 year hiatus between the end of CSD support and the beginning of VA funding. This history is recounted by Cooper (1950); Cooper et al., (1984), and Hauger, (1995).

The Second Phase: New Tools and New Ideas

Cooper and Liberman did not conclude from their research under CSD support that reading machine outputs had to be speech. Drawing that conclusion would have doomed the enterprise for a long time, because speech was not yet a feasible reading machine output. At that time, there was no automatic way to transform print into intelligible speech. Instead, they considered the idea that speech is suitable for conveying language because it has a special *kind* of complex acoustic structure. Perhaps, by better understanding its physical makeup, other, feasible, acoustic substitutes for print could be devised.

To facilitate exploration of that idea, they turned their attention to a new tool for making speech visible and therefore more readily subject to study, the sound spectrograph. The spectrograph was invented during World War II at the Bell Telephone Laboratories. It was used initially for military purposes and classified until after the war, becoming available to civilian researchers only in the late 1940s. It displays the frequency distribution of acoustic energy along the time axis in some acoustic event.

Before development of the spectrograph, a common tool to make speech visible was the oscillograph, which displays variations in acoustic amplitude over time, but does not provide a revealing display of the distribution of frequencies in the signal. Figure 2 displays an oscillogram and a wide-band spectrogram of the sentence *Where were you a year ago* to show the marked differences in the information that each provides.[7]

Cooper saw in the spectrograph a potentially valuable tool for furthering the reading machine project. By showing the frequency composition of speech and its variation over time in a psychologically relevant way, the spectrograph of-

[7] One way to characterize the difference in oscillographic and spectrographic displays (Studdert-Kennedy, personal communication, December 31, 2013) is to note that the former represents the raw acoustic signal whereas the spectrographic display is analogous to its physiological transformation. It represents the output of the cochlea, which performs something like a Fourier analysis.

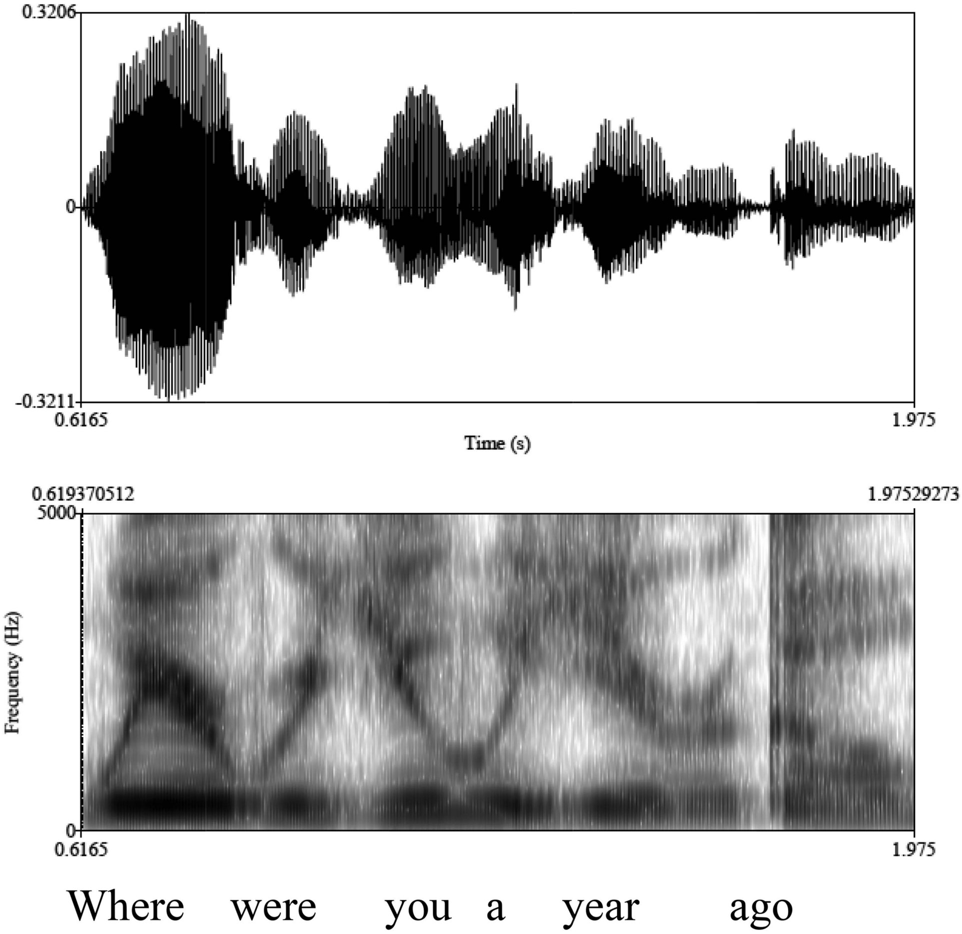

Figure 2. Oscillogram (top) and spectrogram (bottom) of the sentence *Where were you a year ago* illustrating the much greater visibility of acoustic structure in the latter. The oscillogram displays time on the *x*-axis and amplitude on the *y*-axis. It shows periodicity in the voiced parts of the signal as periodic up and down variation in amplitude, but it does not provide information about the vocal tract resonances (formants) that characterize vowels and vowel-like phonetic segments. The spectrogram displays time on the *x*-axis and frequency on the *y*-axis, with amplitude variation represented by variation in darkness. Formants are clearly visible as variation over time and frequency of dark bands of energy.

fered a way to explore the informative acoustic patterning of speech and other complex sounds (Potter, Kopp, & Green, 1947). Thereby it promised to facilitate development of a learnable nonspeech acoustic code for a reading machine. Accordingly, Cooper built spectrographs at Haskins Labs using specifications published by the Bell researchers (Koenig, Dunn, & Lacy, 1946) and modifying the design to meet the special requirements of the Haskins research.

Initially, Cooper and Liberman thought that the displays provided by the spectrograph could be a fruitful source of hypotheses about general principles of perceptual organization (Gestalt or the like) inherent in speech and other patterned material and common to acoustic and visual patterns. However, testing the hypotheses required developing a complementary machine, one that would convert visible spectrographic patterns back into sound and would also allow experimenters to change portions of the patterns. This became the Pattern Playback, invented and built by Franklin Cooper (Cooper, 1950).

As shown in Figure 3, the Pattern Playback consisted of a light source directed toward a rotating disk (a "tone wheel") representing, in concentric rings, the first 50 harmonics of a 120-Hz fundamental frequency. (The fundamental frequency is the lowest frequency in a periodic signal; its harmonics are integral multiples of the fundamental.) In one mode, light modulated at those frequencies was transmitted through a negative of a real spectrogram. In the other mode, light was reflected from a schematic spectrogram painted in white on a transparent acetate belt into a photocell. As the spectrogram was scanned along its time axis at a fixed rate, the frequency-modulated electrical output of the photocell was converted to sound.

Cooper and Liberman ascertained that spectrograms of natural speech when reconverted to sound by the Playback were intelligible, as they had anticipated. This proved that the spectrograph captures, and the Playback reproduces, enough of the phonetic content to enable the output to be understood. They also determined that hand-painted schematic copies of speech spectrograms are generally intelligible when made audible. Figure 4 shows a negative of a spectrogram of the sentence fragment *Never kill a snake* produced by a human speaker (top) and as schematically painted for the Playback (bottom). This fragment was from a set of 20 sentences devised earlier at Harvard to test telephone systems (Egan, 1948). Prominent in the figure are bands of energy in different frequency ranges that convey much linguistic information to listeners; these energy bands are produced by resonances of the vocal tract known as "formants."

Research with the spectrograph and Playback began along two fronts. One, described in *Proceedings of the National Academy of Sciences* (Cooper, Liberman, & Borst, 1951), was specifically meant to test the idea, alluded to above, that some principles of organization underlying perception of visual forms are shared by perception of acoustic forms. To test this idea directly, the researchers painted geometric forms (e.g., triangles, ellipses) onto acetate belts for input to the Pattern Playback. They asked whether varieties of distinct visual shapes retained their category memberships when transformed to the auditory modality. For example, would triangles of different sizes and types in different orientations constitute a category of sounds distinct from a category of sounds synthesized from various ellipses? In general, with some exceptions (e.g., that spatial rotation did not generally preserve visual category identity in the sound domain), they showed that category membership of visible geometric forms was retained in the acoustic domain. This supported their expectation that some principles of organization of visual and auditory forms were common to vision and hearing. An implication was that patterns visible in spectrographic displays of speech might be used to inspire design of nonspeech acoustic signals for a reading ma-

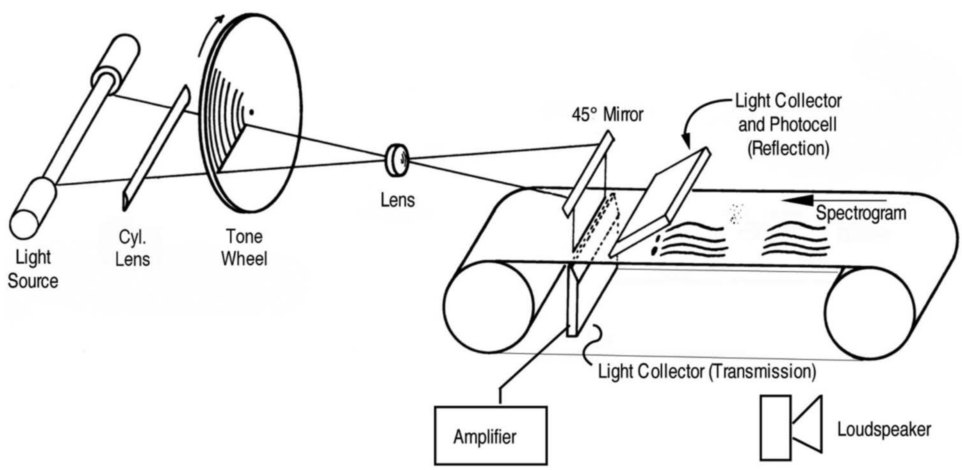

Figure 3. Schematic drawing of the Haskins Pattern Playback. See text for a description.

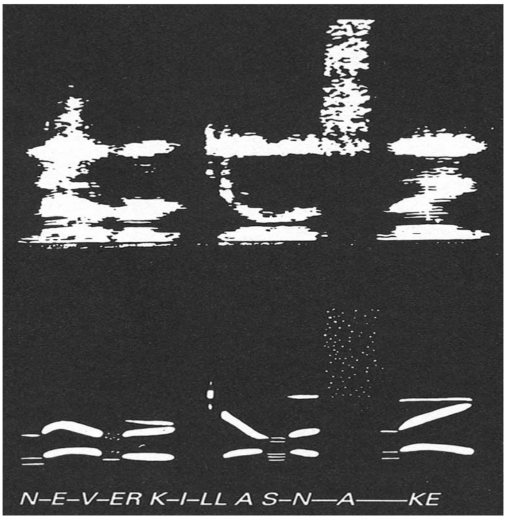

Figure 4. Negative of a spectrogram of the sentence fragment *Never kill a snake* (top) and a schematic copy for input to the Pattern Playback (bottom).

chine that would prove more discriminable and learnable than the failed acoustic alphabets. However, this effort soon proved to be another blind alley (Liberman, 1996).[8]

In contrast, the second line of research undertaken with the spectrograph and Playback was productive in promoting the reading machine goal and also proved critical for shaping the future of basic research at Haskins Laboratories. Haskins researchers began to explore speech for itself. Hand-painted spectrograms of individual syllables enabled them to test hypotheses about how acoustic patterns provided specific kinds of phonetic information by reproducing in schematic form just the bits of acoustic structure of interest. In turn, this enabled researchers to identify "acoustic cues" supporting perception of the consonants and vowels of speech.

The researchers learned that, as valuable as the development of the spectrograph was for speech research, spectrograms of words and phrases cannot usually be read (Liberman et al., 1967). Speech displayed spectrographically is not transparent to consonants and vowels. The opacity reflects two consequences of the fact that speech is coarticulated; that is, as we have noted, the articulatory gestures for successive consonants and vowels in an utterance overlap temporally.[9] Accordingly, neither the acoustic speech signal nor the spectrographic representation of it is composed of phoneme-sized segments (the *segmentation problem*), and the acoustic signal is everywhere highly context sensitive (the *invariance problem*). For both reasons, as Figure 2 reveals, it is very difficult to read off the succession of consonants and vowels of an utterance that the spectrograph represents. (In contrast, alphabetic print represents consonants and vowels much more transparently by distinct symbols.)

The Pattern Playback had inherent limitations of its own. Although playback speech is intelligible, it is not natural sounding. The speech is produced at an invariant fundamental frequency and so is monotone. Likewise, amplitude modulation (variations in darkness on a spectrogram) is not reproduced in the sound output. For both of these reasons, the prosody of the speech is unnatural. On top of that, the schematic spectrograms created by the Haskins researchers were deliberately simplified to test the importance of selected bits of acoustic structure.

Fortunately for this research enterprise, the unnaturalness of playback speech from painted spectrograms proved to be a virtue. As Liberman put it much later in 1988 (Nye, 2006, p. 91), "the success of the method depended on there being an almost complete orthogonality between what you might call naturalness on the one hand and phonetic information on the other." In this way, very spare acoustic patterns could be tested even with naïve listeners to find the minimal informational requirements for perception of particular consonants and vowels.

This line of research proved pivotal, serving indeed to shift the direction of Haskins research toward the investigation of speech as a special

[8] Later, however, it was taken up by others for different purposes; for example, Bregman, 1990; but see Remez, Rubin, Berns, Pardo, & Lang, 1994.

[9] That speech is coarticulated had been reported at least since some time in the 19th century (see Kühnert & Nolan, 1997, for a history). (The term itself is newer, however, dating, according to Kühnert & Nolan, from a publication by Menzerath and De Lacerda, 1933, *Koartikulation, Lautabgrenzung und Steuerung*). However, acoustic phonetics was in its infancy, and Liberman and Cooper were likely unfamiliar with early speculations about coarticulation. Indeed, as noted, they assumed that speech sounds are discrete and acoustically particulate like an alphabet. Understanding that this assumption was misleadingly wrong was essential for the further development of Haskins research on the reading machine and on speech itself.

kind of acoustic signal reflecting in critical ways how it is produced. Two findings led to a revolution particularly in Liberman's thinking (Liberman, 1996; Nye, 2006, p. 89). We describe one finding here and one in the next section. The first occurred as Liberman successively repositioned a schematic spectrogram of *Never kill a snake with your bare hands* (Figure 4) manually on the Playback slowly pulling it back and forth searching for the spectrographic patterning that conveyed the /l/ sound in *kill*. To his surprise, he could find no section on the spectrogram, which when converted to sound, sounded like /l/. Acoustic signals from selected spectrogram locations sounded vowel-like, not /l/-like. This finding sensitized the researchers to two important facts: (a) the relevance of dynamic change in the signal to perception of segmental phonetic information and (b) the pervasiveness of coarticulation in speech articulation and therefore of its acoustic effects.

The Third Phase: Speech Is Special–A Reading Machine Must Talk

The discovery by Liberman, Cooper and Delattre that, indeed, speech is nothing like an acoustic alphabet was reinforced by findings one of which Liberman (1996, p. 13) characterized as an "epiphany." In this research, Haskins researchers opted to work with isolated consonant-vowel (CV) syllables, instead of sentences, to permit closer, better controlled, examination of acoustic encoding of phonetic properties.

In the study producing the second finding alluded to above that led Liberman to look at speech and its perception in a new way (Liberman, Delattre, & Cooper, 1952), the researchers generated CV syllables, such as those shown in Figure 5. They created patterns in which two-formant vowels followed a brief burst of energy at a variety of center frequencies. Listeners identified these patterns as syllables having voiceless stop consonants (/p/, /t/, or /k/) at their onsets. The critical finding was obtained with syllables composed of a burst of energy centered at 1,440 Hz followed by a variety of two formant vowel patterns. Listeners identified the initial stop predominantly as /p/ before the vowels /i/ and /u/, but as /k/ before /a/. Liberman and colleagues recognized that this was not just any context effect. Rather, listeners responded as if they recognized that a burst of energy centered at 1,440 Hz had to be produced by a consonant closure at the lips (i.e., /p/) if the coarticulating vowels were /i/ and /u/, but by the tongue body (i.e., /k/) if the coarticulating vowel was /a/. That is, listeners' perceptions tracked the consonantal articulations, not the acoustic cues that signaled them. In short, when *different* consonantal gestures of the vocal tract gave rise to the *same* acoustic cue, owing to coarticulation with different vowel gestures, they were heard as *different* consonants. Schatz (1954) confirmed this finding with natural speech.

A subsequent finding was complementary (Liberman, Delattre, Cooper, & Gerstman, 1954). In this study, the researchers cued consonants, not by bursts of energy, but by transitions of the formants into the steady-state patterns for the vowels. They determined that the transitions of the second formant distinguished the consonants by place of articulation. Remarkably, as Figure 6 shows, the second formant transitions leading to more /d/ responses than to /b/ or /g/ responses were markedly different for the syllables with vowels /i/ and /u/. For listeners to report /d/ responses before /i/ required a high rise in the second formant transition frequency. But for them to give the same response before /u/ required a low fall. In isolation, these acoustic cues sounded like they looked (a high pitched "chirp" in the first case and a low pitched one in the second). Neither sounded like the consonant /d/ (or like *any* speech sound). In syllable context, they sounded alike, and like /d/. This can be understood if listeners are supposed to track articulation. Release of the same tongue-tip constriction gesture for /d/ into a vocal tract configuration for /i/, which has a high second formant, will lead to a rapidly increasing second formant transition; its release into a vocal tract configuration for /u/, having a low second formant, will lead to a low falling second formant transition. Listeners appear to hear the invariant consonant constriction gesture, not the variable second formant frequencies. In this case, the *same* consonantal gestures give rise, owing to coarticulation with different vowel gestures, to *different* acoustic cues, but are heard as the *same* consonant.

These findings made speech perception a more puzzling and challenging research domain than the researchers had anticipated. (And, as they appreciated later, it made the

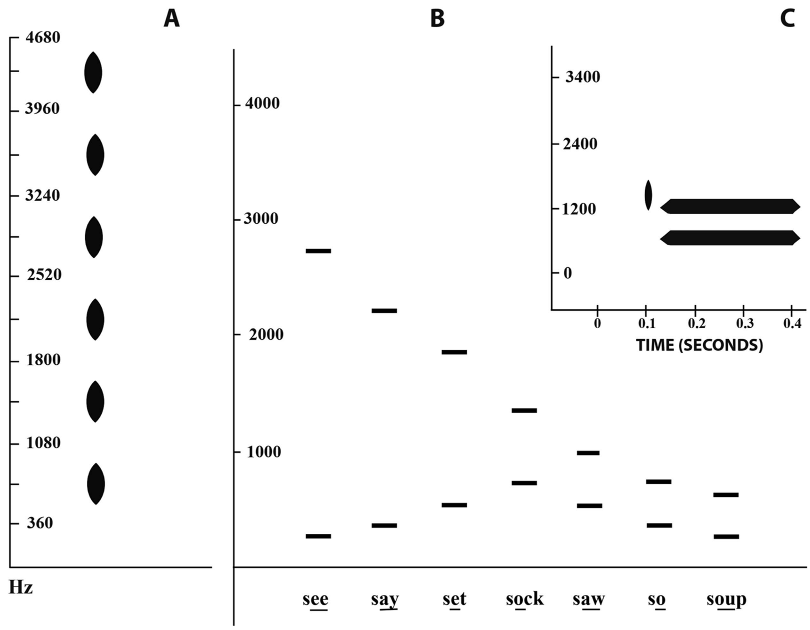

Figure 5. Adapted from Liberman et al. (1952). Schematic depiction of stop bursts (teardrop-shaped patterns on the left) and several two-formant vowel patterns identified as the vowels in key words along the *x*-axis. With stop bursts appended to vowel patterns as in the inset, acoustic signals are heard as consonant-vowel syllables with initial voiceless stop consonants, /p/, /t/, or /k/.

success of print as a means to convey language all the more astonishing.) Coarticulation not only gave rise to an acoustic signal that evaded limits of the temporal resolving power of the ear allowing rapid transmission of information, it also seemed to necessitate a special kind of perceptual process that recovers the discrete phonemes which are merged in the acoustic signal.

These findings and the interpretation of them had two major consequences for research at Haskins Laboratories in the middle

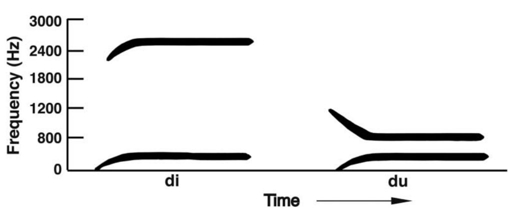

Figure 6. Adapted from Liberman et al. (1954). Formant patterns corresponding to syllables heard predominantly as /di/ and /du/ illustrating an effect of coarticulation.

1950s: First, researchers redirected their research on the reading machine (see especially, Cooper et al., 1984). The researchers now concluded that the output of the device would have to be speech itself. There could be no learnable nonspeech acoustic surrogate. Accordingly, they investigated ways to provide speech from a phonetic description of print.

A second consequence was that Liberman, Cooper, and colleagues chose to pursue basic research on speech for its own sake. The early findings just described were provocative. Discovering that the speech signal is nothing like an acoustic alphabet, but rather is an "encoded" signal (Liberman et al., 1967) as a result of coarticulatory overlap of gestures does not by itself explain why speech is such an effective signal for perceivers or why acoustic alphabets are not. Hints from the earliest investigations that listeners track articulation suggest an answer; however, it is an answer that appears to put speech perception in a category of its own, an idea that Liberman and Cooper came to embrace.

These innovative findings are among the legacies of the research at Haskins Laboratories on the reading machine. Haskins researchers never intended to produce a market-ready device as we have discussed, and they never did. Raymond Kurzweil accomplished that in the middle 1970s. In the following pages, we describe the Haskins researchers' subsequent work to produce satisfactory speech outputs for a reading machine. In a final section, we outline the multipronged development of their basic research program on speech. We preface these sections with a brief discussion of changes over the decades in funding sources at Haskins Laboratories that reflect changes in their research focus and changes in the way research was funded in the post-World War II era.

Growth of Federal Support for Haskins Research[10]

Following the end of CSD funding, the Laboratories' staff shrank considerably in size. But during the long funding hiatus, from the mid-1940s to the mid-1950s, Haskins researchers continued to pursue research relevant to the goal of a high-performance reading machine under a modest, long-term grant from the Carnegie Corporation. Moreover, the Laboratories took on other projects in applied acoustics supported by federal military science contractors, such as the Office of Naval Research, the National Security Agency, and Air Force Cambridge Research Center, which received their funding from the large defense budgets of the Cold War years (Solovey, 2013). These projects enabled this portion of the Laboratories to expand again gradually.[11]

After the 10-year funding hiatus, support for reading machine research was picked up by the Department of Veterans Affairs. Other support for the Laboratories came primarily from Caryl Haskins himself (who consistently donated between $20,000 and about $75,000 each year through at least 1970), the Carnegie Corporation, and a variety of other sources, primarily foundations. However, as the research focus of reading machine researchers shifted increasingly to include research on speech, the proportion of the Laboratories' funding for these projects from such federal sources as National Science Foundation, the military agencies previously mentioned, and especially the National Institutes of Health grew steadily from constituting about 30% of Laboratories income for reading machine related projects in the early 1950s to close to 90% by the end of the 1960s. This trend reflects the large increases in federal spending for research in psychology and the behavioral sciences, chiefly research based in university departments (Capshew, 1999). From the 1960s on, the Department of Health, Education, and Welfare (especially the National Institutes of Health) was a major patron promoting huge expansion of behavioral research in many domains (op cit). Today, research on speech, language, and reading at Haskins Labs is almost entirely funded by federal research grants, including a program project grant that has been continuously funded by the National Institute of Child Health and Human Development since 1966.

Prerecorded Compiled Speech Versus Speech Synthesis for Reading Machine Output

Beginning in 1957 and continuing for 21 years, the Prosthetics and Sensory Aids Service of the Department of Veterans Affairs (PSAS) supported reading machine development at Haskins Laboratories and three other sites (Cooper et al., 1984). The research pursued at Haskins was along two lines consistent with the Laboratories' original goal of creating a satisfactory reading machine output. However, now the outputs were exclusively speech. The first research line explored compiled speech as an output. This would consist of a database, an electronically stored dictionary, of recorded words that would be retrieved and sequenced as the text required and output acoustically to the user. Any words not in the database, rare words or words from specialized vocabularies, would be spelled out letter by letter. The second line of inquiry was more ambitious: the object was to synthesize speech from the ground up from acoustic specifications.

Of the two ways a reading machine could be made to speak, compiled speech from stored

[10] The information in this section was obtained by examining financial documents archived at Haskins Laboratories.

[11] In the meantime, biological research by other members of Haskins Laboratories was also expanding and was later to receive major federal support.

acoustic recordings of spoken words was for a time the more expedient option. However, the researchers quickly learned that compiled speech would present difficulties for the listener that would reduce its appeal as a long-term solution. In particular (e.g., Gaitenby, 1965), speakers produce the same word in different ways in different contexts and, moreover, fluent speech has a melody that concatenation of invariant recordings of words could not duplicate. Moreover, although the compiled speech was intelligible, listeners were bothered by the fact that words not in the dictionary database (of 7,200 words) had to be spelled out, letter by letter. The presence of even a few spelled-out words in a passage interrupts the flow of speech so much that listeners found it distracting (Cooper et al., 1984).

The First Synthesis by Rule

The term "synthesis" is usually reserved for an artificial system that generates and combines subword units to create speech out of its elementary units. Haskins researchers' past research with acoustic alphabets, and their investigations of speech itself using the Pattern Playback and spectrograph, led them to a view that synthesis of speech from its elements ultimately would be the more promising direction to take in this line of research. One reason invoked a principle of economy. A few dozen sets of synthesis instructions representing phoneme units would, in principle, constitute the raw materials for generating all the words of a language.[12] Accordingly, the second line of research sought audible outputs for a reading machine by creating speech starting from the letters of the text, which after being transcribed as instructions for phonemes according to rules of synthesis, were to be concatenated to form the words of the text.

Research seeking the acoustic cues for speech perception, necessary for development of synthesis by rule, was underway for several years before researchers attempted to systematize this growing body of knowledge. To undertake this systematization, the Laboratories hired Frances Ingemann, a recent PhD in Linguistics from Indiana University who had received phonetics training at the University of Copenhagen (Willhite, 2000). Her task was to cull generalizations from the existing body of research and formulate rules for painting schematic spectrograms to be rendered acoustically by the Pattern Playback. Following the interpretation of the speech cues adopted by Cooper, Liberman, and the Haskins pioneer speech researchers, she developed rules at the level of phoneme segments (not words, not syllables or demisyllables). Ingemann found that the rules could be formulated most economically by attending to the subphonemic features shared by classes of phonemes, members of which shared some acoustic properties (Cooper et al., 1984; see also Liberman, Ingemann, Lisker, Delattre, & Cooper, 1959). Regarding shared acoustic properties, for example, each vowel is characterized by a distinctive formant pattern, and stop consonants by an interval of silence followed by a stop burst and then formant transitions into the following vowel. As we saw in the example of stop consonant cues, the transitions are not merely links between consonants and vowels; they provide crucial consonantal (and vocalic e.g., Strange, 1987) information. Accordingly, rules could be written that applied to a class of phonemes rather than just to one. Cooper et al. (1984, p. 74) characterize the rules as generating speech that, produced by the Pattern Playback, was "for the most part, fully intelligible" albeit "wooden and machine-like." Informal testing (Liberman et al., 1959) suggested that sentence intelligibility ranged from 60–100%. But an adequate test of the rules was not possible at that time because there was no way to produce synthetic-speech versions of whole paragraphs of connected text, in sufficient quantity.

Automation of Speech Synthesis

Valuable as it was for the pioneering work on speech acoustics, speech perception, and synthesis, the Pattern Playback had limitations that ultimately caused it to be shelved. Most critical for the purposes at hand, synthesis with the Playback, implemented with rules such as Ingemann's, required a human agent to draw the spectrographic patterns and physically place them on the playback device. But to be useful for a reading machine, rules for synthesis of connected speech would have to be imple-

[12] In practice, things were more complicated. Several versions of each phoneme would be required to produce intelligible speech.

mented by a machine automatically without human intervention. The Playback also has limitations of fidelity, which result, as we noted, from the process used for converting spectrograms to sound. Both of these limitations made Playback speech unsuitable for simulating the output of a reading machine.

Synthesis tools, which were invented in the precomputer era, came into their own when matched with a computer and a synthesis program. The advent of laboratory computers in about 1960 greatly expanded the possibilities for machine control of synthesis and improvement of synthesis quality (as it also opened up a variety of other applications to linguistic studies [e.g., Crowther-Heyck, 1999]). The goal of machine control was further served by the availability of resonance synthesizers, a breed of synthesizers that differ in design from the Pattern Playback and is readily adaptable to the computer.[13]

A resonance synthesizer (Klatt, 1980) replaces a human operator with electrical pattern and function generators. Electrical circuitry is used to represent the frequency of the laryngeal tone and the first few resonances of the speech signal (formants). The synthesizer can be driven by control signals from a digital computer replacing a human agent with a computer program. The program's parameters correspond to dimensions of natural speech observable in spectrograms: in addition to formant frequencies and amplitudes and fundamental frequency, a hiss source corresponds to unvoiced noisy portions of utterances and a variable frequency buzz source corresponds to voiced portions. The resonance synthesizer is a more comprehensive tool than the Pattern Playback. By permitting the simulation of voice pitch through control of fundamental frequency, for example, it enables greater acoustic fidelity and naturalness of the speech output.

Regarding the role of computers in speech synthesis, Mattingly (1968, p. 39) notes that a computer can be used for three distinct tasks: implementing rules of synthesis by calculating parameter values, "on line" transmission of parameter values to a synthesizer, and simulation of the synthesizer itself. It is for the first of these tasks that a computer proves indispensable when there is a need, as there would be for a reading machine, to produce a quantity of synthetic speech in real time. Without a computer, the calculations could not be done except laboriously by hand. As noted, the synthesizer device itself can also be simulated by computer. An advantage of simulation is that the design of the synthesizer is more easily modified when it is realized in software than when it is a hardware device.

Automated speech synthesis became an accomplished fact in 1961 when Bell Labs researchers, physicist J. L. Kelly and psychologist L. J. Gerstman, who earlier had been a member of Haskins Labs and a pioneer contributor to research on the speech cues, demonstrated speech synthesis under computer control (Kelly & Gerstman, 1961). Kelly and Gerstman matched a resonance synthesizer (simulated by computer) with a synthesis program specifying rules for synthesis, based on the Haskins-Ingemann rules (Mattingly, 1968). Although the Kelly-Gerstman system was more a demonstration of possibilities than a working tool, the live exhibition of computer synthesis before the 1961 meeting of the Acoustical Society of America was a landmark in the development of machine simulation. One result of this feat was to bring a high performance reading machine closer to realization. It also opened the door to many other applications of synthesis as well.

At about that time, Ignatius Mattingly, then a linguistic researcher at the Defense Department, working with J. N. Holmes and J. N. Shearme of the Joint Speech Research Unit, Eastcote, England developed a program for the synthesis of British English (Holmes, Mattingly, & Shearme, 1964). The system produced speech that sounded remarkably good for that early date. Joining Haskins Laboratories in 1966, Mattingly worked intensively on synthesis, adapting and augmenting the existing Haskins rules for American English (Liberman et al., 1959).

Mattingly's rules for computer synthesis of American English and the routines for their implementation became the foundation for later developments in synthesis at Haskins Laboratories, as well as a major influence on developments elsewhere (Klatt, 1987). One outcome is

[13] The history of speech synthesis by rule and its implementation by synthesizers of various kinds is reviewed by Klatt (1987). For a review of the earlier period, see Mattingly (1968) and a collection of seminal articles edited by Flanagan and Rabiner (1973).

that, in Cooper's words, "pessimism about the prospects for using synthetic speech in a reading machine changed to optimism largely on the basis of Mattingly's successful undertaking" (Cooper et al., 1984, p. 76). It would be only another few years before a working model of a complete reading machine was produced by inventor Raymond Kurzweil. In the meantime, there were still significant obstacles to overcome. The rules were (and are still) incomplete and partly wrong, resulting in speech that is not quite human sounding. Nevertheless, they are perfectly explicit and therefore testable, and capable of improvement.

A functional text-to-speech application, such as a reading machine for the blind, must generate speech from print automatically. Accordingly, this requires an intermediary set of routines that converts printed letters (which have been identified by an optical character recognizer) into the corresponding speech units which in turn drive the synthetic speech output. A letter-to-phoneme converter became the third component of a fully realized plan for a reading machine (Cooper et al., 1984; Klatt, 1987).

The design of a letter-to-phoneme converter was pursued at MIT by Jonathan Allen and his students beginning in the 1960s. Allen maintained that a sufficiently powerful algorithm capable of accurately transcribing any written word in English requires that the program successfully simulate abilities of a skilled human reader who knows the language well (Allen, 1976). For the reader of English, the conversion from letters to phonemes is not a trivial matter. Experienced readers have learned that these correspondences are not fixed; they reflect a word's pedigree, its origin in Anglo-Saxon, Greek or Latin roots, or a borrowing from one or another modern language (Venezky, 1970) or spellings that did not keep abreast of sound changes in the language. One consequence is that letters often do not correspond one-to-one to phonemes. For example, the vowel letter < I > and the digraph <*ea*> each has many phonemic renderings (pronunciations) in English. For reasons such as this, English writing is complexly related to units of speech by a great variety of rules, some of very narrow scope.

Allen's print-to-phoneme algorithm (Allen, 1976; Allen, Hunnicutt, & Klatt, 1987) decomposes complex words into their roots and affixes, then finds their phonetic equivalents and assembles their phonetic spellings. A dictionary of root forms and a list of exceptions is incorporated into the system for this process. Roughly 20,000 items had to be stored in computer memory, but can be used to generate word specifications for a vocabulary many times as large.[14]

The 1980s saw the commercial development of synthesis and synthesizers. These developments, which are beyond the time frame of this article, owed much to decades of basic and applied research at Haskins Laboratories. When Kurzweil's company finally created a prototype reading machine in 1976 (and subsequently produced more compact units that were sold at a cost low enough to place them in the homes of some blind users), it had available an off-the-shelf commercial speech synthesizer with accompanying rules program. Within a year or two other better and cheaper alternatives became available and were incorporated into later models of the Kurzweil Reading Machine (Hauger, 1995). A description of the Kurzweil reading machine appears in the Proceedings of the AAAS (Kurzweil, 1976).

Evolution of Speech Research at Haskins Laboratories: An Overview

The reading machine project at Haskins Labs, conducted with the goal of a concrete application, was a catalyst for development of first one, then an integrated array of productive lines of basic research. Moreover, the reading machine project and other research at the Laboratories intended to yield basic scientific knowledge, were highly interactive. Basic discoveries about speech led reading machine researchers to choose synthetic speech as the appropriate output of the reading machine, and to discard all nonspeech acoustic codes. For its part, progress in basic research on speech perception depends critically on discovery of the acoustic cues for consonants and vowels, and discovery of the cues depended on use of synthetic speech. Knowledge of the cues was also indispensable for making a reading machine talk.

[14] This development of an orthographic link became incorporated, along with rules for synthesis created by Dennis Klatt (1980), into a complete computer-based text-to-speech system called MITalk (Allen, Hunnicutt, & Klatt, 1987; Klatt, 1987).

Research on the acoustic cues for speech perception has occupied researchers at Haskins Laboratories over several decades (e.g., Abramson, 2004; Delattre, 1958; Liberman et al., 1952). This research has led to a clearer understanding of the information supporting speech perception in a variety of languages.

The central question raised by the findings of Liberman and colleagues (1952, 1954) reviewed earlier is whether speech perception and production are to be regarded as separate domains or whether they are importantly linked. It was noted that, when the distinction could be made, listeners' perceptions were found to adhere more closely to articulatory speech actions than to the acoustic cues that inform about the actions. This led Liberman and colleagues to develop a "motor theory" of speech perception in which they claimed that listeners use highly context-dependent acoustic speech cues to recover speech motor invariants that they proposed mapped more directly to phonetic segments than did the acoustic cues.

Early formulations of the motor theory regarded the links between perception and production as having been forged during language acquisition by associating various acoustic cues and various articulatory responses, reflecting Liberman's training as a behaviorist (e.g., Liberman, 1957). However, the theory underwent a biological reinterpretation when Donald Shankweiler introduced Laboratories' researchers to a behavioral tool, dichotic listening, for studying cerebral hemispheric specialization for different varieties of auditory stimuli (e.g., Kimura, 1961, 1967; Shankweiler, 1966). Using dichotic listening, Shankweiler and Studdert-Kennedy (1967) found that meaningless consonant-vowel syllables showed a right ear/left hemisphere advantage, linking speech phonetics to the left-hemisphere specialization for language itself. Later formulations of the motor theory (e.g., Liberman et al., 1967; Liberman & Mattingly, 1985) invoked a specialization of the brain, eventually identified as a brain module (Fodor, 1983), responsible for both production and perception of coarticulated speech.

Early research designed to test the motor theory used electromyography to seek invariant muscle commands in speech production. That project failed to find invariants (see MacNeilage, 1970, for a review). However, it led Katherine Harris and colleagues to recognize speech production and coarticulation as important domains for experimental research (e.g., Bell-Berti & Harris, 1981; Lisker & Abramson, 1971). Harris, Bell-Berti and other Haskins researchers made valuable research contributions to the study of speech physiology, and especially coarticulation (e.g., Bell-Berti & Harris, 1982; Gelfer, Bell-Berti, & Harris, 1989).

A significant development of this research line was achieved by embedding the study of speech within the context of a general theory of action (e.g., Turvey, 1977a). Turvey, also well-known for his espousal of Gibson's ecological theory of perception (Gibson, 1966, 1979; Hayes-Roth, 1977), underscored that coarse-grained actions are produced by *synergies* (e.g., for speech, lip closing achieved by coordinated actions of the jaw, upper lip and lower lip), rather than movements of individual articulators or individual muscles. Turvey and colleagues proposed that synergies constitute a psychologically relevant level of description. In speech, the smallest units of action achieved by synergies were proposed to be *phonetic* gestures of the vocal tract, the linguistically relevant constituents of word forms (e.g., Fowler, Rubin, Remez, & Turvey, 1980).

Seen not only as units of speech production, but also as the articulatory events perceived by listeners, the idea of phonetic gestures became a growth point for several further theoretical developments and lines of research at the Laboratories. One was development of an articulatory phonology (e.g., Browman & Goldstein, 1986, 1992), a linguistic phonology in which primitive units of language form were claimed to be the very actions of the vocal tract produced by talkers and perceived by listeners. A second related development was an understanding of the importance of "parity" to human linguistic communication (e.g., Liberman & Mattingly, 1989: Liberman & Whalen, 2000). This was the idea that elements of speech action, of speech perception, and of the language itself *must* be the same in part because individuals know and use the same language in both roles, as talkers and as listeners. Furthermore, for language to serve as a vehicle for communication between people, listeners have to perceive what (other) talkers say; that is, they must, successfully for the most part, identify the language forms produced by speakers in order to discern reliably what a speaker is attempting to communicate.

A third development reflected the interest among the Laboratories' researchers in the biological specialization for language, specifically here, in its evolution. Following Abler (1989); Studdert-Kennedy (1998, 2000) underscored a "particulate principle," that operates in nature whenever a hierarchy of diverse forms arises from a limited set of building blocks. In these systems, elementary units (in physics, particles; in chemistry, elements; in biology, genes; in language, phonetic gestures for consonants and vowels) combine without blending to make "infinite use of finite means" (von Humboldt, 1836/1972). In language, combinations of a few phonetic gestures form an indefinitely large number of words, which themselves form an indefinitely large number of sentences.

A productive family of research lines at the Laboratories developed from the researchers' understanding that speech is part of the human specialization for language. If, as motor theorists proposed, speech perception depends on an evolved specialization of the brain for producing speech and perceiving it by ear, Mattingly asked, (e.g., Mattingly, 1991), how is reading at all possible when it is a process that depends on accessing elements of the spoken language *by eye* (e.g., Frost, 1998)? Addressing this paradox led to research on skilled reading comparing different writing systems (e.g., Lukatela & Turvey, 1998), and research on reading development (e.g., I. Y. Liberman, Shankweiler, & A. Liberman, 1989) and reading disability (Shankweiler, & I. Y. Liberman, 1989) pioneered by Isabelle Liberman and Donald Shankweiler. Cross linguistic work on skilled reading led to an understanding of the several variable factors in language and orthography that mediate the link between print and phonology (e.g., Frost & Katz, 1992; Mattingly, 1984). In reading development, I. Y. Liberman and colleagues (I. Y. Liberman, 1973; I. Y. Liberman, Shankweiler, Fischer, & Carter, 1974) proposed that phonological awareness, a metalinguistic skill, should underlie successful learning to decode words, an idea that has been repeatedly confirmed by research (Brady, Braze, & Fowler, 2011; Byrne, 1998) and has led to applications of basic research findings to reading instruction for typically developing and disabled readers (e.g., Shankweiler & A. E. Fowler, 2004). As for reading disability, Shankweiler and colleagues proposed a phonological deficit hypothesis (e.g., Shankweiler, I. Y., Liberman, Mark, Fowler, & Fischer, 1979) that disabled readers develop deficient phonological representations of spoken words that impede word decoding in reading. Beginning in the early 1990s, Haskins researchers and their collaborators have explored the neuropsychology of reading, successfully tracing brain activity patterns associated with reading skill differences in several languages using neuroimaging techniques (e.g., Shaywitz et al., 1998; Pugh et al., 2000).

The Catalyzing Effect of Haskins Laboratories Research in the Research Community

The legacy of research on the reading machine was not what the pioneers or their sponsors had anticipated. Decades of work toward this goal led to a number of dead ends as researchers were constrained by their mistaken notions about speech as a sound alphabet, and limitations of the technologies available (especially, no optical character recognition and no synthetic speech). Even though these barriers were eventually lifted in later years, the early work was not wasted. Haskins research established definitively that to achieve rapid rates of reading, blind users would need to hear speech itself, restricting further reading machine developments to achievement of that goal. Moreover, their work, especially on speech synthesis, had an impact (not always acknowledged) on the transformative commercial products that were eventually developed by others.

At Haskins Laboratories and elsewhere there have been lasting effects. One impact is methodological. Among their many contributions, Haskins researchers pioneered the study of "categorical perception," a finding that listeners respond to some speech segments almost as if they perceive only their category membership (e.g., bilabial, voiced), but not their within-category variability. The original test for categorical perception developed by Liberman and colleagues (Liberman, Harris, Hoffman, & Griffith, 1957) is still frequently cited and discussed. More generally, the technique used to detect categorical perception, by varying speech stimuli incrementally along some acoustic dimension and observing the identification and discrimination response patterns, is now standard procedure in much speech research, including research on speech development in infancy (as pioneered by Eimas, Siqueland, Jusczyk, & Vigorito, 1971). A second impact of Haskins research is theoretical: The motor theory of speech perception of Liberman and colleagues, Fowler and Best's direct perception alternative (e.g., Fowler, 1986; Best, 1995), Turvey's theoretical perspective on action and as applied to the speech domain by Saltzman, Kelso and Munhall (e.g., Saltzman & Kelso, 1987; Saltzman & Munhall, 1989), Browman and Goldstein's articulatory phonology—all have had an impact on speech research. In reading, the formative ideas include the parasitic relation of speech and reading as modulated by variations among orthographies and scripts (I. Y. Liberman, Liberman, Mattingly, & Shankweiler, 1980) and the importance of phonological awareness for beginning readers (I.Y. Liberman, Shankweiler, & Liberman, 1989). These theoretical approaches have had an impact, not because they are universally accepted by the scientific community, but, by exciting controversy, they have stimulated critical discussion and research. Today, 70 years after the faltering beginnings of the reading machine project, Haskins Laboratories is best known for its pioneering research on speech and reading, but it also deserves to be known for the pioneering work on the reading machine that stimulated these developments.

References

Abler, W. (1989). On the particulate principle of self-diversifying systems. *Journal of Social and Biological Structures, 12,* 1–13. http://dx.doi.org/10.1016/0140-1750(89)90015-8

Abramson, A. S. (2004). The plausibility of phonetic explanations of tonogenesis. In G. Fant, H. Fujisaki, C. Jianfen, & Y. Xu (Eds.), *From traditional phonology to modern speech processing: Festschrift for Professor Wu Zongji's 95th birthday* (pp. 17–29). Beijing, China: Foreign Language Teaching and Research Press.

Allen, J. (1976). Synthesis of speech from unrestricted text. *Proceedings of the IEEE, 64,* 433–442. http://dx.doi.org/10.1109/PROC.1976.10152

Allen, J., Hunnicutt, M. S., & Klatt, D. (1987). *From text to speech: The MITalk system*. New York, NY: Cambridge University Press.

Bell-Berti, F., & Harris, K. S. (1981). A temporal model of speech production. *Phonetica, 38,* 9–20. http://dx.doi.org/10.1159/000260011

Bell-Berti, F., & Harris, K. S. (1982). Temporal patterns of coarticulation: Lip rounding. *The Journal of the Acoustical Society of America, 71,* 449–454. http://dx.doi.org/10.1121/1.387466

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 171–204). Baltimore, MD: York Press.

Brady, S. A., Braze, F. D., & Fowler, C. A. (Eds.). (2011). *Explaining individual differences in reading: Theory and evidence*. New York, NY: Taylor and Francis.

Bregman, A. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: The MIT Press.

Browman, C., & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook, 3,* 219–252. http://dx.doi.org/10.1017/S0952675700000658

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica, 49,* 155–180. http://dx.doi.org/10.1159/000261913

Byrne, B. (1998). *The foundation of literacy: The child's acquisition of the alphabetic principle*. Hove, United Kingdom: The Psychology Press.

Capshew, J. H. (1999). *Psychologists on the march: Science, practice, and professional identity in America, 1929–1969*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511572944

Constant, E. W. (1980). *The origins of the turbojet revolution*. Baltimore, MD: Johns Hopkins University Press.

Cooper, F. S. (1950). Research on reading machines for the blind. In P. Zahl (Ed.), *Blindness: Modern approaches to the unseen environment* (pp. 512–543). Princeton, NJ: Princeton University Press.

Cooper, F. S., Gaitenby, J. H., & Nye, P. W. (1984). Evolution of reading machines for the blind: Haskins Laboratories' research as a case history. *Journal of Rehabilitation Research and Development, 21,* 51–87.

Cooper, F. S., Liberman, A. M., & Borst, J. M. (1951). The interconversion of audible and visible patterns as a basis for research in the perception of speech. *Proceedings of the National Academy of Sciences of the United States of America, 37,* 318–325. http://dx.doi.org/10.1073/pnas.37.5.318

Crowder, R. G. (1982). *The psychology of reading: An introduction*. New York, NY: Oxford University Press.

Crowther-Heyck, H. (1999). George A. Miller, language, and the computer metaphor of mind. *History of Psychology, 2,* 37–64. http://dx.doi.org/10.1037/1093-4510.2.1.37

Delattre, P. C. (1958). Les indices acoustiques de la parole: Premier rapport. *Phonetica, 2,* 108–118. http://dx.doi.org/10.1159/000257850

Egan, J. P. (1948). Articulation testing methods. *Laryngoscope, 58,* 955–991. http://dx.doi.org/10.1288/00005537-194809000-00002

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science, 171,* 303–306. http://dx.doi.org/10.1126/science.171.3968.303

Flanagan, J. L., & Rabiner, L. R. (Eds.). (1973). *Benchmark papers in acoustics: Speech synthesis*. Stroudsburg, PA: Dowden, Hutchinson, & Ross.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.

Fournier d'Albe, E. E. (1920). The Optophone: An instrument for reading by ear. *Nature, 105,* 295–296. http://dx.doi.org/10.1038/105295a0

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics, 14,* 3–28.

Fowler, C. A., Rubin, P., Remez, R. E., & Turvey, M. T. (1980). Implications for speech production of a general theory of action. In B. Butterworth (Ed.), *Language production* (pp. 373–420). New York, NY: Academic Press.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123,* 71–99. http://dx.doi.org/10.1037/0033-2909.123.1.71

Frost, R., & Katz, L. (Eds.; 1992). *Orthography, phonology, morphology, and meaning*. Amsterdam, the Netherlands: North-Holland.

Gaitenby, J. H. (1965). The elastic word. *Haskins Laboratories Status Reports on Speech Research, 2,* 3.1.1271–3.

Gelfer, C. E., Bell-Berti, F., & Harris, K. S. (1989). Determining the extent of coarticulation: Effects of experimental design. *The Journal of the Acoustical Society of America, 86,* 2443–2445. http://dx.doi.org/10.1121/1.398452

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston, MA: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Haskins Laboratories. (1953). History. Retrieved from http://www.haskins.yale.edu/history.html

Hauger, J. S. (1995). *Reading machines for the blind: A study of federally supported technology development and innovation*. PhD dissertation, Virginia Polytechnic Institute and State University, Blacksburg.

Hayes-Roth, F. (1977). Critique of Turvey's "Contrasting orientations to the theory of visual information processing." *Psychological Review, 84,* 531–535. http://dx.doi.org/10.1037/0033-295X.84.6.531

Holmes, J. N., Mattingly, I. G., & Shearme, J. N. (1964). Speech synthesis by rule. *Language and Speech, 7,* 127–143.

Jameson, M. (1932). The Optophone, or how the blind may read ordinary print by ear. *And There was Light, 1,* 18–22.

Kelly, J. L., Jr., & Gerstman, L. J. (1961). An artificial talker driven from a phonetic input. [A]. *The Journal of the Acoustical Society of America, 33,* 835. http://dx.doi.org/10.1121/1.1936801

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology, 15,* 166–171. http://dx.doi.org/10.1037/h0083219

Kimura, D. (1967). Functional asymmetry of the brain in dichotic listening. *Cortex, 3,* 163–178. http://dx.doi.org/10.1016/S0010-9452(67)80010-8

Klatt, D. (1980). Software for a cascade/parallel formant synthesizer. *The Journal of the Acoustical Society of America, 67,* 971–995. http://dx.doi.org/10.1121/1.383940

Klatt, D. H. (1987). Review of text-to-speech conversion for English. *The Journal of the Acoustical Society of America, 82,* 737–793. http://dx.doi.org/10.1121/1.395275

Koenig, W., Dunn, H. K., & Lacy, L. Y. (1946). The sound spectrograph. *The Journal of the Acoustical Society of America, 18,* 19–49. http://dx.doi.org/10.1121/1.1916342

Kühnert, B., & Nolan, F. (1997). The origin of coarticulation. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation de Universität München, 35,* 61–75. Retrieved from http://ed268.univparis3.fr/lpp/pages/EQUIPE/kuehnert/bk_pub/kuehnert_nolan99.pdf

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.

Kurzweil, R. (1976). The Kurzweil reading machine: A technical overview. In M. R. Redden & W. Schwandt (Eds.), *Science, technology, and the handicapped. American Association for the Advancement of Science Meeting, Boston* (pp. 3–7). Washington, DC: American Association for the Advancement of Science Publication 76-R-11.

Layton, E. T. (1971). Mirror-image twins: The communities of science and technology in 19th-century America. *Technology and Culture, 12,* 562–580. http://dx.doi.org/10.2307/3102571

Layton, E. T. (1974). Technology as knowledge. *Technology and Culture, 15,* 31–41. http://dx.doi.org/10.2307/3102759

Liberman, A. M. (1957). Some results of research on speech perception. *Journal of the Acoustical Society of America, 29,* 117–123. http://dx.doi.org/10.1121/1.1908635

Liberman, A. M. (1996). *Speech: A special code*. Cambridge, MA: Bradford Books.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74,* 431–461. http://dx.doi.org/10.1037/h0020279

Liberman, A. M., Delattre, P., & Cooper, F. S. (1952). The role of selected stimulus-variables in the perception of the unvoiced stop consonants. *The American Journal of Psychology, 65,* 497–516. http://dx.doi.org/10.2307/1418032

Liberman, A. M., Delattre, P., Cooper, F. S., & Gerstman, L. (1954). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs: General and Applied, 68,* 1–13. http://dx.doi.org/10.1037/h0093673

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54,* 358–368. http://dx.doi.org/10.1037/h0044417

Liberman, A. M., Ingemann, F., Lisker, L., Delattre, P., & Cooper, F. S. (1959). Minimal rules for synthesizing speech. *Journal of the Acoustical Society of America, 31,* 1490–1499. http://dx.doi.org/10.1121/1.1936079

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory revised. *Cognition, 21,* 1–36. http://dx.doi.org/10.1016/0010-0277(85)90021-6

Liberman, A. M., & Mattingly, I. G. (1989). A specialization for speech perception. *Science, 243,* 489–494. http://dx.doi.org/10.1126/science.2643163

Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences, 4,* 187–196. http://dx.doi.org/10.1016/S1364-6613(00)01471-6

Liberman, I. Y. (1973). Segmentation of the spoken word and reading acquisition. *Bulletin of the Orton Society, XXIII, 6,* 5–77.

Liberman, I. Y., Liberman, A. M., Mattingly, I. G., & Shankweiler, D. (1980). Orthography and the beginning reader. In J. F. Kavanagh & R. L. Venezky (Eds.), *Orthography, reading, and dyslexia* (pp. 137–153). Baltimore. MD: University Park Press.

Liberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology, 18,* 201–212. http://dx.doi.org/10.1016/0022-0965(74)90101-5

Liberman, I. Y., Shankweiler, D., & Liberman, A. M. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1–33). Ann Arbor, MI: The University of Michigan Press.

Lisker, L., & Abramson, A. S. (1971). Distinctive features and laryngeal control. *Language, 47,* 767–785. http://dx.doi.org/10.2307/412155

Lukatela, G., & Turvey, M. T. (1998). Reading in two alphabets. *American Psychologist, 53,* 1057–1072. http://dx.doi.org/10.1037/0003-066X.53.9.1057

MacNeilage, P. F. (1970). Motor control of serial ordering of speech. *Psychological Review, 77,* 182–196. http://dx.doi.org/10.1037/h0029070

Mattingly, I. G. (1968). *Synthesis-by-rule of General American English*. Supplement to the Haskins Laboratories Status Report on Speech Research. New York, NY: Haskins Laboratories.

Mattingly, I. G. (1984). Reading, linguistic awareness, and language acquisition. In J. Downing & R. Valtin (Eds.), *Language awareness and learning to read* (pp. 9–25). New York, NY: Springer-Verlag. http://dx.doi.org/10.1007/978-1-4613-8248-5_2

Mattingly, I. G. (1991). Reading and the biological function of linguistic representations. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception* (pp. 339–346). Hillsdale, NJ: Erlbaum.

Menzerath, P., & De Lacerda, A. (1933). *Koartikulation, Lautabgrenzung und Steuerung* [Coarticulation, Phonetic Segment Boundaries, and Control]. Berlin, German: Fred. Dümmlers.

Miller, G. A., & Taylor, W. G. (1948). The perception of repeated bursts of noise. *Journal of the Acoustical Society of America, 20,* 171–182. http://dx.doi.org/10.1121/1.1906360

Nye, P. W. (2006). *An oral history of Haskins Laboratories*. Retrieved from http://www.haskins.yale.edu/history/OH/HL_Oral_History.pdf

Nye, P. W., & Bliss, J. C. (1970). Sensory aids for the blind: A challenging problems with lessons for the future. *Proceedings of the IEEE, 58,* 1878–1898. http://dx.doi.org/10.1109/PROC.1970.8061

Petrosky, H. (2006). *Success through failure: The paradox of design*. Princeton, NJ: Princeton University Press.

Potter, R. K., Kopp, G. A., & Green, H. L. (1947). *Visible speech*. New York: Van Nostrand.

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., . . . Shaywitz, B. A. (2000). Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Mental Retardation and Developmental Disabilities Research Reviews, 6,* 207–213. http://dx.doi.org/10.1002/1098-2779(2000)6:3<207::AID-MRDD8>3.0.CO;2-P

Remez, R. E., Rubin, P. E., Berns, S. M., Pardo, J. S., & Lang, J. M. (1994). On the perceptual organization of speech. *Psychological Review, 101,* 129–156. http://dx.doi.org/10.1037/0033-295X.101.1.129

Saltzman, E., & Kelso, J. A. S. (1987). Skilled actions: A task-dynamic approach. *Psychological Review, 94,* 84–106. http://dx.doi.org/10.1037/0033-295X.94.1.84

Saltzman, E., & Munhall, K. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology, 1,* 333–382. http://dx.doi.org/10.1207/s15326969eco0104_2

Schatz, C. D. (1954). The role of context in the perception of stops. *Language, 30,* 47–56. http://dx.doi.org/10.2307/410219

Shankweiler, D. (1966). Effects of temporal-lobe damage of perception of dichotically presented melodies. *Journal of Comparative and Physiological Psychology, 62,* 115–119. http://dx.doi.org/10.1037/h0023470

Shankweiler, D., & Fowler, A. E. (2004). Questions people ask about the role of phonological processes in learning to read. *Reading and Writing, 17,* 483–515. http://dx.doi.org/10.1023/B:READ.0000044598.81628.e6

Shankweiler, D., & Liberman, I. Y. (1989). *Phonology and reading disability: Solving the reading puzzle*. Ann Arbor, MI: The University of Michigan Press.

Shankweiler, D., Liberman, I. Y., Mark, L. S., Fowler, C. A., & Fischer, F. W. (1979). The speech code and learning to read. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 531–545. http://dx.doi.org/10.1037/0278-7393.5.6.531

Shankweiler, D., & Studdert-Kennedy, M. (1967). Identification of consonants and vowels presented to left and right ears. *The Quarterly Journal of Experimental Psychology, 19,* 59–63. http://dx.doi.org/10.1080/14640746708400069

Shaywitz, S. E., Shaywitz, B. A., Pugh, K. R., Fulbright, R. K., Constable, R. T., Mencl, W. E., . . . Gore, J. C. (1998). Functional disruption in the organization of the brain for reading in dyslexia. *Proceedings of the National Academy of Sciences of the United States of America, 95,* 2636–2641. http://dx.doi.org/10.1073/pnas.95.5.2636

Solovey, M. (2013). *Shaky foundations: The politics-patronage-social science nexus in Cold War America*. New Brunswick, NJ: Rutgers University Press.

Strange, W. (1987). Information for vowels in formant transitions. *Journal of Memory and Language, 26,* 550–557. http://dx.doi.org/10.1016/0749-596X(87)90141-0

Studdert-Kennedy, M. (1998). The particulate orgins of language generativity: From syllable to gesture. In J. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language* (pp. 202–221). New York, NY: Cambridge University Press.

Studdert-Kennedy, M. (2000). Evolutionary implications of the particulate principle: Imitation and the dissociation of phonetic form from semantic function. In C. Knight, M. Studdert-Kennedy, & J. R. Hurford (Eds.), *The evolutionary emergence of language* (pp. 161–176). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511606441.011

Studdert-Kennedy, M., & Liberman, A. M. (1963). Psychological considerations in design of auditory displays for reading machines. *Proceedings of the International Congress on Technology and Blindness, 1,* 289–304.

Turvey, M. T. (1977). Preliminaries to a theory of action with reference to vision. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing* (pp. 211–265). Hillsdale, NJ: Erlbaum.

Venezky, R. (1970). *The structure of English orthography*. The Hague, Switzerland: Mouton. http://dx.doi.org/10.1515/9783110804478

Vincenti, W. G. (1990). *What engineers know and how they know it*. Baltimore, MD: Johns Hopkins University Press.

Von Humboldt, W. (1972). *Linguistic variability and intellectual development* (G. C. Buck & F. A. Raven, Trans.). Philadelphia, PA: University of Pennsylvania Press. (Original work published 1836)

Willhite, J. (2000). *An interview with Frances Ingemann*. Oral History Project, University of Kansas. Retrieved from http://groups.ku.edu/~endacottsociety/History/OralHistoryTranscripts ingemann.wpd.pdf

Zahl, P. (1950). *Blindness: Modern approaches to the unseen environment*. Princeton, NJ: Princeton University Press.

Received May 9, 2014
Revision received August 14, 2014
Accepted September 15, 2014 ■