ALVIN M. LIBERMAN and FRANKLIN S. COOPER

IN SEARCH OF THE ACOUSTIC CUES[1]

It was our good fortune to have worked closely with Pierre Delattre in pursuit of a common goal: to find the acoustic cues that underlie the perception of speech. We had come to this research from different backgrounds and by different routes. At the time we started to work together, Pierre Delattre was a phonetician who knew about speech that which was known and worth knowing; we two writers of this paper were psychologist and physicist who could claim, in speech, no more than an active curiosity. He had long understood why it might be important to find the acoustic cues; we had only very recently come to appreciate why that might be a worthwhile thing to do. On this occasion it is appropriate, we hope, to say how we were led to wonder about speech, and to describe some of our experiences in the early stages of the research. This will show, though only in part, why we had to rely so heavily on Pierre Delattre.

Our interest in speech, and our appreciation of it, developed out of an attempt to solve a practical problem. In the mid-forties we were given an opportunity to work with various other people in the design and evaluation of prosthetic devices for the blind. One of these devices was, or was to have been, a reading machine (Cooper, 1950; Studdert-Kennedy and Cooper, 1966; Studdert-Kennedy and Liberman, 1963). The point of this device was to convert print to sound and thus, with appropriate arrangements for scanning, make it possible for the blind to read any book, magazine, or newspaper. This seemed to us an eminently reasonable thing to do. At that time there were no print recognizers — devices that identify letters of the alphabet — as there are now, but it was technologically possible, by scanning the print with a photosensitive element, to use the contours of the print to control the parameters of an acoustic signal.

We assumed when we began this work that the problem was straightforward and rather easily soluble: the device had only to produce, for the 26 letters of the alphabet, 26 corresponding and discriminably different sounds; the user had only to learn to associate each sound with the appropriate letter, and to come, by practice, to do this

---

[1] The authors are deeply indebted to the Carnegie Corporation of New York for its support of these research studies during their uncertain beginnings.

with accuracy and speed. In fact, it was not difficult to convert the letters into sounds, even into sounds that were discriminably different. This had been done not too badly as early as 1914 (D'Albe, 1924). The difficulty was, rather, that our human subjects could not, even after considerable practice, perceive these sounds rapidly or well. Top performance was under twenty words per minute — a tenth or less of normal rates for speech. Perhaps we should not have been surprised, for the same lack of success had been the outcome of attempts to build reading machines before we entered the arena, and it has been the experience of those who have followed us (Freiberger, 1966). In any case, the difficulties we encountered in trying to substitute arbitrary sounds for speech taught us to respect the sounds of speech as very special and uniquely efficient vehicles for the transmission of phonemic information. Accordingly, we developed a keen interest in finding out why these sounds are perceived so well.

For one who wishes to study the perception of speech, the first task is obvious enough and not different from that which confronts the scientist who sets out to investigate the perception of anything else: it is to find the cues — the physical stimuli — that control the perception. In the case of speech this is, for several reasons, somewhat difficult. One must first do something about the fact that speech is invisible and transitory. It was our good luck that workers at Bell Telephone Laboratories had recently developed the sound spectrograph. As all students of speech know, this device converts the speech signal into a visible and permanent display that makes reasonably good sense to the eye — better sense, surely, than the oscillogram, the only display previously available. But good as the spectrogram is, one cannot, simply by looking at it, find the information-bearing features of speech. Part of the difficulty is that one sees several features for each phonemic difference, and so cannot discover by inspection which of the several features is doing the work.

But there is another and far more interesting difficulty, one that goes to the very heart of speech. It arises from the fact that the relation between sound (hence spectrogram) and phoneme is complex. To appreciate the nature of this complexity, and the difficulty it creates, we should imagine how easy the task might have been if the relation between sound and phoneme were simple. In that case we might suppose that each phoneme would be represented in the sound stream by a unit signal. Hence there would be a reasonably obvious segmentation of the physical signal corresponding to the segments of the phonemic message — that is, there would be, in any utterance, as many acoustic segments as phonemes. In this simplest of all possible speech worlds it would also be true that the sound representing a particular phoneme would be essentially the same regardless of context. And, inasmuch as speech is quite easily perceived, we should suppose, too, that the important information-bearing elements would be among the most prominent parts of the acoustic signal, standing in strong contrast to a background of noise and linguistically irrelevant features. These are, of course, exactly the simple relations between signal and graphemic message that one finds on the printed page. If speech were alphabetic in the same way, then one could easily find and describe the acoustic cues simply by examining any reasonably appropriate representation of the acoustic signal. Indeed, the problem would in principle be not essentially different from that of describing the letter shapes of some exotic alphabet. But, as we have already indicated, the relation of speech to phoneme is not simple in this way. The segmentation of the acoustic signal does not correspond to the phoneme segments; the acoustic cues for particular phonemes are not, in general, the same in different contexts; and the most important cues are sometimes among the least prominent parts of the acoustic signal.

We hasten to say that this appreciation of the relation between sound and phoneme was a consequence of the experimental work we did jointly with Pierre Delattre; these things were not fully apparent to us when first we looked at spectrograms. What we saw then suggested only that the speech signal was complex in ways that would make it difficult to find the acoustic cues simply by examining spectrograms. It seemed reasonably clear that it would be most important — even necessary — to experiment with the speech signal — that is, to make changes in the presumably important parts of the acoustic pattern, and then determine the effects of these changes by listening to the modified sounds. We decided, then, to try to take advantage of the spectrographic display, not just as a means of examining the speech signal, but also, and even primarily, as a basis for modifying and reconstituting it in audible form. To that end, we developed the Pattern Playback and the techniques associated with its use. The device has been described elsewhere (Cooper, 1953), so we will say about it here only that it converts spectrograms — either photographic copies of real spectrograms or hand-painted versions — into sound.

Our first step with the Playback was to see whether or not we could produce reasonably intelligible phrases and sentences from simple, hand-painted spectrograms. For that purpose, we first copied from spectrograms of 'real' speech those parts of the pattern that seemed reasonably apparent to the eye and, also, those that impressed us, for other reasons, as likely to contain linguistic information. We were pleased, and perhaps a little surprised, to discover that, with a manageable amount of trial-and-error editing, these patterns produced fairly intelligible speech (Cooper, Liberman and Borst, 1951). But these early successes in producing synthetic sentences taught us very little about speech; they only demonstrated that it might be feasible to use simple, hand-painted patterns as a basis for experimenting with it. So far as the cues for speech perception were concerned, we could conclude only that they were somehow included in our simplified hand-painted patterns — that is, that they were to be found in patterns that contain only two or three formants, plus occasional click-like acoustic segments and patches of noise. The essential complexity of the relation between phoneme and sound — that is, the lack of correspondence in segmentation at those two levels and the context-conditioned variation of the acoustic signal — was not significantly reduced in these early synthetic spectrograms. As a consequence, we could not see the cues, even in our simplified patterns.

It was at this point that the actual research on speech began. Fortunately, it was also at this point that Pierre Delattre joined us. The collaboration that followed was active and close, so much so that it is hardly possible to say who did what. Still, it must be clear how much we two writers of this paper, almost totally naive about phonetics and related subjects, had to depend on Pierre Delattre's knowledge and ideas. It is, perhaps, less obvious, but certainly no less true, that such success as we had is also owing largely to his great skill as an experimenter, to his illuminating insights, to his tireless devotion to the work, and to his patience with problems or colleagues when they proved refractory.

Pierre Delattre and we undertook, then, to identify the information-bearing elements in speech. At the outset we proceeded partly by educated guesses and partly by sheer trial and error, using our hand-painted sentences as a base. When we thought we had succeeded in defining a cue, we regularized it, varied it systematically along what we guessed to be the appropriate dimensions, and then discovered what kind of result we got when we listened to the sounds and identified them as phonemes. If the results of that listening test confirmed our earlier impressions, we usually gave the test to a group of phonetically naive listeners. In this way we conducted many informal and formal experiments designed to find those aspects of the acoustic pattern that were (or were not) sufficient cues for the perception of the phonemic segments or subphonemic dimensions.

We will not describe the results of these experiments — they have appeared elsewhere[2] — but will, rather, remark on several aspects of our experience, especially with method, that may be of some general interest. First, we should emphasize how essential it was in the early stages that the techniques for synthesis were flexible and convenient. We have several times said that we proceeded largely by trial and error, by making and testing our guesses. We should say now that we have been talking about many thousands of trials and an only slightly smaller number of errors. If these many trial experiments could not have been done quickly, they probably could not have been done at all. It was, therefore, essential that the patterns could be prepared, converted to sound, and then modified with reasonable speed and facility. To do that, we had to sacrifice a great deal of naturalness and realism but, as later events proved, the loss was small and the gain large. All this may be obvious enough, at least in retrospect. It is, perhaps, a little less obvious that the patterns had also to be conceptually agreeable. That is, experimental convenience in this case required not only that one be able to paint, erase, and listen, but also that he deal with stimuli that could be kept in mind and thought about as reasonably distinctive visual patterns. That requirement is rather well met by the spectrogram. To have controlled the parameters of speech from other displays that make less sense as patterns would have made

---

[2] Cf., Cooper, Delattre, Liberman, Borst and Gerstman, 1952; Delattre, 1954, 1958, 1962; Delattre, Cooper and Liberman, 1952; Delattre, Cooper, Liberman and Gerstman, 1956; Delattre, Liberman and Cooper, 1951, 1955, 1964; Delattre, Liberman, Cooper and Gerstman, 1952; Harris, Hoffman, Liberman, Delattre and Cooper, 1958; Liberman, Delattre and Cooper, 1952, 1958; Liberman, Delattre, Cooper and Gerstman, 1954; Liberman, Delattre, Gerstman and Cooper, 1956; Liberman, Ingemann, Lisker, Delattre and Cooper, 1959; O'Connor, Gerstman, Liberman, Delattre, and Cooper, 1957.

our exploratory trial-and-error work vastly more difficult — as, indeed, we found it to be when confronted with the hundred or so knobs and dials of a later synthesizer.

Another aspect of our experience with method that may be of some general significance arose from the fact that we could not study all the phonemes at once. Given that we had, as a practical matter, to limit our attention to a few phonemes at a time, it was important that these few constitute a perceptually meaningful class. Consider, for example, what might have happened if we had chosen to find the cues on the basis of which a listener could distinguish among /s/, /a/, and /w/. In an experimental test, listeners might have sorted stimuli into these three classes with great consistency, but on the basis of superficial or even irrelevant distinctions: they might, perhaps, have identified any signal that contained some kind of noise as /s/, any signal in steady state as /a/, and one that contained a relatively slow change as /w/. The inadequacy of such results would, of course, have been revealed just as soon as we began to group the stimuli differently and to require, for example, that our subjects distinguish /s/ from /ʃ/. But such corrections as would inevitably have followed could have been made only slowly and painfully. In fact, we investigated the cues in terms of the standard linguistic dimensions and classes, one minimal class or dimension at a time. This usually drew our attention to a single acoustic attribute of the spectrographic pattern that was common to all members of the class, and that, when varied systematically, enabled the listener to distinguish among them. Thus, for example, we experimented with the cues that distinguish among the voiced stops (i.e., according to place of production). When we had isolated what we thought to be a sufficient cue — in the first instance, the transition of the second formant (Liberman, Delattre, Cooper and Gerstman, 1954) — we varied that parameter alone in patterns readily heard as voiced stops and asked phonetically naive listeners to identify each resulting sound as /b/, /d/, or /g/. The outcomes of such isolated experiments gave information about this cue for place of production that held up quite well when, in related studies, we investigated the cues that distinguish the voiced stops from voiceless stops (Liberman, Delattre and Cooper, 1958) or from nasal consonants (Delattre, Liberman and Cooper, 1955; Liberman, Delattre and Cooper, 1958), i.e., the same second-formant transitions served as cues for the same places of production regardless of manner or condition of voicing. This testifies to what is clearly an important fact about speech — namely, that the major linguistic dimensions of place, manner, and voicing (and the various classes of phonemes that can be formed from them) are largely independent in perception. There is very little interaction across dimensions or classes. If this were not so, we should have encountered very great practical difficulties, no matter how we might have chosen to group the phonemes.

It is, perhaps, equally fortunate for our research effort — and interesting as a fact about speech — that the several acoustic cues within a single phonemic dimension or class tend also to be independent in perception. In our first formal experiment on consonants, for example, we found that the frequency position of a burst of noise is a cue for perception of the voiceless stops according to place of production (Liberman, Delattre and Cooper, 1952). The patterns of that experiment consisted of highly schematized bursts followed by absolutely steady-state formants. We found, in a following experiment, as noted above, that transitions (rapid frequency changes) of the second formant are also cues for perception of stops according to place of production. These experiments were carried out with patterns that had no bursts. Subsequently we found that the results we had obtained separately with the bursts and the transitions held up reasonably well when both kinds of cues were combined in the same pattern (Delattre, 1958). Still later we found that transitions of the third formant were also cues for place of production (Harris, Hoffman, Liberman, Delattre and Cooper, 1958). The effects of all three cues (burst of noise and transitions of both second and third formants) proved to be largely independent of each other (Delattre, 1958; Hoffman 1958), at least in the qualitative sense that best values for each cue did not change when the other cues were varied; quantitatively, though, the effectiveness of multiple cues suggested some kind of vectorial addition of single-cue effects. Our research would have been enormously more complicated than it was if these three cues (and other sets of cues for other distinctions) had interacted to any considerable degree. More important, of course, we should have been confronted with a very different fact about speech. The true, and we think interesting, conclusion is that the different acoustic cues for particular phonemic distinctions are independent to an astonishing degree.

We remarked earlier that the synthetic speech we produced on the Playback was unnatural and unrealistic, especially in the early days. To some extent this was the price we had to pay for experimental convenience, but it was also true that we would not have known very well how to make the speech realistic even if it had been critically important for us to do so. It was, therefore, particularly fortunate for our research, and it may be of some general interest, that perception of the linguistic message is almost wholly independent of the general quality of the carrier, except in an all-or-nothing way. It is as if there were, in this regard, two independent sets of cues in the speech signal, one that tells the listener he is hearing speech (and also gives him other paralinguistic information, including data about voice quality) and another that indicates what the string of phonemes is. If the first set is so far off the mark that the listener does not hear the sounds as speech, then little or nothing can be learned about the cues for the phonemes. But if the experimenter manages to produce signals that are heard as speech, then, no matter how poor the quality, the information he obtains about the phonemic cues will have great generality; that is, the results will not be much changed when the same variations are made in patterns that sound like closer approximations to human speech. The primary hazard in using synthetic speech that lacks realism, or that has been stripped of multiple cues, is that the experimenter will fail to get data from some of his crew of listeners. Some of the subjects, especially those who are phonetically naive, will not be able to hear the patterns as speech; hence, their responses to the stimuli will be random and so add noise to the results. When the same experiment is done with stimuli that contain more of the 'realism' or 'naturalism' cues — or, when experimentation is possible with real speech — the pattern of results may be clearer, but it will be essentially the same pattern.

It is difficult to imagine synthetic speech more unnatural sounding than the stimuli used in our first experiment on the cues for stop consonants. As we pointed out earlier, these were highly schematized bursts of sound followed by two-formant steady-state vowels. We were able, nevertheless, to obtain results that proved to have some reasonable generality, both in other experiments with synthetic speech and in a parallel experiment on real speech (Schatz, 1954). In this connection we should remark on the fact that the Playback works on the basis of the first 50 harmonics of a 120-cycle fundamental; there is no provision for pitch inflection or for noise. The monotone nature of the synthetic speech proved to be no handicap in studying the segmental phonemes, and was a small price to pay for the freedom from instrumental complexity inherent in the variable-pitch synthesizers that we, and others, developed later (Cooper, 1962). The lack of adequate noise sources was a more important limitation, yet it was possible even so to do serious research on the fricatives (Delattre, 1958; Delattre, Liberman and Cooper, 1964). By putting dots of paint (in place of lines for formants) on the synthetic spectrogram, we produced a quasi-random effect that sounded more like a twittering of birds than a proper fricative noise. Nevertheless, in the context of certain patterns, listeners can hear fricatives, and it is possible to determine for example, that a salient cue for the distinction between /s/ and /ʃ/ is the frequency position of the lower edge of the band-limited twittering. Essentially that same result is obtained when the experimenter uses a synthesizer that produces true and proper noise (Harris, 1956; Heinz and Stevens, 1961).

As a result of research by many people working in various laboratories, a great deal is now known about the acoustic basis for speech.[3] Moreover, several investigators have systematized this knowledge by incorporating it into explicit rules by which reasonably intelligible speech can be synthesized.[4] This is to be contrasted with our first synthesized sentences, referred to earlier, which were produced by copying from spectrograms and then making revisions by trial and error. To synthesize speech by explicit rule — or 'free-hand' from thorough familiarity with the cues, as Pierre Delattre did so well — is surely an appropriate goal of the search for acoustic cues. It is also an example of a highly scientific kind of linguistic description, suggesting hypotheses that can be confirmed or denied (Lisker, Cooper and Liberman, 1962). Even though rules for synthesis, as they now exist, produce less than perfect speech, they are nonetheless perfectly explicit and perfectly testable.

---

[3] In this account of research shared with Pierre Delattre, the authors have not tried to reflect related efforts by other investigators, or even all of the research done in their own laboratories. One recent review to which the reader might turn for such information is Stevens and House, 1971. For a more nearly complete review of the Haskins work cf. Liberman, Cooper, Shankweiler and Studdert-Kennedy, 1967.

[4] Speech synthesis by rule is under active investigation in a number of laboratories. For a statement of the problem and an account of early work cf. Liberman, Ingemann, Lisker, Delattre and Cooper, 1959; for a review of other work, as well as a current contribution, cf. Mattingly, 1968.

We are still a long way from fully answering the question that arose from the work with reading machines and motivated our research: why are the sounds of speech perceived so well? But, as a result of the work we did with Pierre Delattre, we have come a fair distance, or so we think. We know now that the sounds of speech differ from our early reading machine signals in ways clearly calculated to make them more efficient. It is plain, for example, that, unlike the signals produced by our reading machines, the sounds of speech are not an alphabet on the phonemes, but a very complex code (Liberman, Cooper, Shankweiler and Studdert-Kennedy, 1967). One important feature of this code is that information about successive phonemic segments is transmitted in parallel. That is, information about two or more successive phonemic segments is carried simultaneously by the same acoustic cue. The second-formant transition, for example, is at every instant conveying information about the place of production of the consonant and also the identity of the following vowel. This reduces by a significant factor the number of discrete acoustic segments that must be perceived per unit time, and thus enables the listener to evade the severe limitations on rate set by the temporal resolving power of the ear. Given the kind of code that exists in the relation between sound and phoneme, the rate at which discrete acoustic segments would merge into an unanalyzable buzz is set by the number of syllables per second, not by the number of phonemes. But this advantage is gained at the cost of a complex relation between sound and phoneme: a corollary of the very efficient parallel transmission we find in speech is, as we pointed out earlier, that there is little correspondence between acoustic and phonemic segmentation and that the acoustic cues for many phonemes differ markedly with differences in context. For one who wants only to perceive speech, this great complexity appears to pose no problem at all, presumably because human beings have a readily available device that processes the incoming sounds and recovers the phoneme strings (Liberman, Cooper, Studdert-Kennedy, Harris and Shankweiler, 1968). Indeed, it is probably for this reason that students of language have for so long tended to think of speech as a simple alphabet. But, as we have tried to say, the hidden though very real complexities of speech do create difficulties for the investigator who would seek the acoustic cues that underlie its perception. In overcoming these difficulties, and others as well, we had the constant help of Pierre Delattre. Had we not been guided by him, we might still be floundering in the early exploratory phases of our research.

HASKINS LABORATORIES
NEW HAVEN, CONNECTICUT

REFERENCES

Cooper, F.S.,

1950 "Research on Reading Machines for the Blind", in Zahl, P.A. (ed.), *Blindness. Modern Approaches to the Unseen Environment* (Princeton, Princeton University Press): 512–43.

1953 "Some Instrumental Aids to Research on Speech", *Report on the Fourth Annual Round Table Meeting on Linguistics and Language Teaching* (Washington, Georgetown University Press): 46–53.

1962 "Speech synthesizers", *Proc. IVth Int. Congress of Phonetic Sciences* (Helsinki) (The Hague. Mouton and Co.) 3–13.

Cooper, F. S., P. C. Delattre, A. M. Liberman, J. M. Borst and L. J. Gerstman,
1952 "Some Experiments on the Perception of Synthetic Speech Sounds", *JASA* 24: 597–606.

Cooper, F. S., A. M. Liberman and J. M. Borst,
1951 "The Interconversion of Audible and Visible Patterns as a Basis for Research in the Perception of Speech", *Proc. National Academy of Sciences* 37: 318–25.

D'Albe, F.E.E.,
1924 *The Moon Element* (London, T. Fisher Unwin).

Delattre, P.C.,
1954 "Les attributs acoustiques de la nasalité vocalique et consonantique", *Studia Linguistica* 8: 103–09.
1958 "Les indices acoustiques de la parole: Premier rapport", *Phonetica* 2: 108–18, 226–51.
1962 "Le jeu des transitions des formants et la perception des consonnes", *Proc. IVth Int. Congress of Phonetic Sciences* (Helsinki, 1961) (The Hague, Mouton): 407–17.

Delattre, P.C., F.S. Cooper and A.M. Liberman,
1952 "Some Suggestions for Language Teaching Methods Arising from Research on the Acoustic Analysis and Synthesis of Speech. *Report on the Third Annual Round Table Meeting on Linguistics and Language Teaching* 2: 31–47.

Delattre, P. C., F. S. Cooper, A. M. Liberman and L. J. Gerstman,
1956 Speech synthesis as a research technique. *Proc. VIIIth Int. Congress of Linguistics* (London, 1952). (London, 1956, pp. 543–61 plus 9 figures).

Delattre, P.C., A.M. Liberman and F.S. Cooper,
1951 "Voyelles synthétiques à deux formants et voyelles cardinales", *Le Maître Phonétique* 96: 30–7.
1955 "Acoustic Loci and Transitional Cues for Consonants", *JASA* 27:769-73
1964 "Formant Transitions and Loci as Acoustic Correlates of Place of Articulation in American Fricatives", *Studia Linguistica* 18: 104–21.

Delattre, P. C., A. M. Liberman, F. S. Cooper and L. J. Gerstman,
1952 "An Experimental Study of the Acoustic Determinants of Vowel Color: Observations on One- and Two-Formant Vowels Synthesized from Spectrographic Patterns", *Word* 8: 195–210.

Freiberger, H.,
1966 "Research in Sensory Aids for the Blind", *Proc. Int. Conference on Sensory Devices for the Blind* (London, St. Dunstans): 427–34.

Harris, K.S.,
1956 "Some Acoustic Cues for the Fricative Consonants", *JASA* 28: 160(A).

Harris, K.S., H.S. Hoffman, A.M. Liberman, P.C. Delattre and F.S. Cooper,
1958 "Effect of Third-Formant Transitions on the Perception of the Voiced Stop Consonants", *JASA* 30: 122–26.

Heinz, J.M., and K.N. Stevens,
1961 "On the Properties of Voiceless Fricative Consonants", *JASA* 33: 589–96.

Hoffman, H.S.,
1958 "Study of Some Cues in the Perception of the Voiced Stop Consonants", *JASA* 30: 1035–41.

Liberman, A.M., F.S. Cooper, D.P. Shankweiler and M. Studdert-Kennedy,
1967 "Perception of the Speech Code", *Psychological Review* 74: 431–61.

Liberman, A.M., F.S. Cooper, M. Studdert-Kennedy, K.S. Harris and D.P. Shankweiler,
1968 "Some Observations on the Efficiency of Speech Sounds" (Paper presented at the XVIII Int. Congress of Psychology, Moscow, August, 1966) *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikations forschung* 21: 21–32.

Liberman, A.M., P.C. Delattre and F.S. Cooper,
1952 "The Role of Selected Stimulus Variables in the Perception of the Unvoiced Stop Consonants", *Am. J. of Psychology* 65: 497–516.
1958 "Some Cues for the Distinction Between Voiced and Voiceless Stops in Initial Position", *Language and Speech* 1: 153–67.

Liberman, A.M., P.C. Delattre, F.S. Cooper and L.J. Gerstman,
1954 "The Role of Consonant-Vowel Transitions in the Perception of the Stop and Nasal Consonants", *Psychological Monographs* 68: 8 (Whole No. 379).

Liberman, A.M., P.C. Delattre, L.J. Gerstman and F.S. Cooper,
1956 "Tempo of Frequency Change as a Cue for Distinguishing Classes of Speech Sounds", *J. of Experimental Psychology* 52: 127–37.

Liberman, A.M., F. Ingemann, L. Lisker, P.C. Delattre and F.S. Cooper,
1959 "Minimal rules for synthesizing speech", *JASA* 31: 1490–99.

Lisker, L., F.S. Cooper and A.M. Liberman,
1962 "The Uses of Experiment in Language Description", *Word* 18: 82–106.

Mattingly, I.G.,
1968 "Synthesis by Rule of General American English", Ph.D. Thesis (Yale University).

O'Connor, J.D., L.J. Gerstman, A.M. Liberman, P.C. Delattre and F.S. Cooper,
1957 "Acoustic Cues for the Perception of Initial /w,j,r,l/ in English", *Word* 13: 25–43.

Schatz, C.,
1954 "The Role of Context in the Perception of Stops", *Language* 30: 47–56.

Stevens, K.N., and A.S. House,
1971 "Speech Perception", in J. Tobias (ed.) *Foundations of Modern Auditory Theory* (New York, Academic Press).

Studdert-Kennedy, M., and F.S. Cooper,
1966 "High-Performance Reading Machines for the Blind: Psychological Problems, Technological Problems, and Status", *Proc. Int. Conference on Sensory Devices for the Blind* (London, St. Dunstans): 317–42.

Studdert-Kennedy, M., and A.M. Liberman,
1962 "Psychological Considerations in the Design of Auditory Displays for Reading Machines", *Proc. Int. Congress on Technology and Blindness* I, (New York): 289–304.