

Paradigm lost

700

Michael Studdert-Kennedy

Haskins Laboratories, New Haven, CT 06510*

Electronic mail: haskins@yalevm.bitnet

The theme of *Paradigm* is that perception, in general, and speech perception, in particular, entails the integration of information from many sources. The theme is scarcely new, having been a commonplace of speech research almost since its inception (e.g. Jones 1948; Miller et al. 1951; Pollack & Pickett 1963; Scholes 1968). In fact, very little in *Paradigm* is new, beyond the impressive mathematical apparatus marshalled to support the argument. Yet it is precisely the mathematics that robs the discussion of its interest – at least for a reader concerned with the biological foundations of language, for it leads Massaro to treat all sources of information as equivalent, and so to miss the true import of his own and others' work on audiovisual speech perception.

Paradigm is concerned solely with process, with *how* we perceive, and is indifferent to *what* we perceive (cf. Repp 1988, p. 253). Accordingly, the author grants exactly the same status to information stemming from the perceiver's knowledge of phonology, syntax, semantics, and pragmatics as to information carried by the physical properties of the perceived event itself. There can be no question that all these forms of knowledge contribute to speech perception, and may even be essential to later stages in the "interactive" process of language acquisition that the author favors (*Paradigm*, p. 240). Yet there can equally be no question that for the infant learning to speak, it is the physical (acoustic, optic) properties of an utterance that must first be apprehended, for these are properties that specify the articulatory organization of one language rather than another.

The central weakness of *Paradigm's* approach to speech perception, then, is its failure to acknowledge that the process of integrating auditory and visual information is qualitatively different from the process of integrating "top-down" linguistic and

extralinguistic knowledge. While the latter process may be as general as *Paradigm* proposes, the former evidently rests on the specific fact that acoustic and optic information are structurally equivalent because they have a common origin in the speaker's articulation. Curiously, *Paradigm* cites (p. 51) the important experiment of Summerfield (1979) supporting the hypothesis that the integration of acoustic and optic-properties in speech engages a "common metric" of articulatory dynamics, but elects not to address the issue. *Paradigm* also conspicuously fails to acknowledge that it was Summerfield (1979; see also Summerfield 1983; 1987), not Massaro, who first demonstrated that, contrary to the original hypothesis of McGurk and MacDonald (1976), acoustic and optic information are analog rather than categorically discrete at their conflux.

Unhappily, *Paradigm's* selective and self-serving treatment of Summerfield's work is typical of its treatment of other work. For example, roughly one quarter of the text is devoted to demonstrating that categorical phonetic percepts derive from a decision process based on continuous auditory information. No reader who was unfamiliar with the literature on speech perception would suspect that this conclusion has been a commonplace of that literature for nearly 20 years, and was indeed well established before Massaro even entered the field.

In fact, the hypothesis of Studdert-Kennedy et al. (1970) that stop consonants (though not vowels) could be discriminated only if they were differently identified had already been undermined before it was published. Stevens (1968) and Sachs (1969) both attributed the poor within-category discrimination of stop consonants to the transient nature of their acoustic cues. Fujisaki and Kawashima (1970) proposed a model explaining the consonant-vowel difference as due to differential decay of auditory memory during a discrimination trial. Pisoni (1973) and his colleagues (Pisoni & Lazarus 1974; Pisoni & Tash 1974) supported Fujisaki's "dual process" model in a series of elegant and decisive experiments to which *Paradigm* gives no credit. Studdert-Kennedy and Shankweiler (1970) also attributed the consonant-vowel differences in dichotic ear advantages to differential loss of auditory information during transcassal transfer.

Moreover, all these studies and many others have been repeatedly acknowledged in review articles (e.g., Repp 1984; Studdert-Kennedy 1975; 1976; 1980; 1981). For example: "In short, consonants and vowels are distinguished in the experiments we have been considering, not by their phonetic class or the processes of assignment to that class, but by their acoustic characteristics and by the duration of their auditory stores" (Studdert-Kennedy 1976, p. 263). Compare this statement with the following from *Paradigm*: "If the sensory information is lost very quickly, continuous information could participate in the perceptual process but might not be readily accessible for introspective reports" (p. 120; p. 14 of target article). Massaro apparently does not recognize the correspondence between these two statements. But then he does not acknowledge many of the papers cited in this and the preceding paragraph, nor the many other relevant papers cited in those papers. Perhaps he has not read them. If so, one can only observe that those who do not study the literature are condemned to repeat it.

Finally, let me note that *Paradigm* grossly overrates the importance of categorical perception as support for the hypothesis that speech perception engages specialized cerebral processes. The principal supports for this hypothesis are: (1) the failure, despite attempts over many years, to devise an arbitrary acoustic substitute for speech in reading machines for the blind that would be any more efficient than Morse code, for which perceptual rates are scarcely one tenth those of speech (Liberman et al. 1967); (2) the repeated demonstration, in both brain-lesioned and normal subjects, of a double dissociation between left and right cerebral hemispheres in the perception of speech and nonspeech sounds (e.g., Blumstein 1981; Molfese & Betz 1988; Tartter 1988). That such findings have no place in the

Commentary/Massaro: Speech perception

grand design of a *Paradigm for Psychological Inquiry* invites the conclusion that this is a paradigm we would do well to lose.

ACKNOWLEDGMENT

Preparation of this comment was supported in part by NIHCD Grant 01994 to Haskins Laboratories.

NOTE

*Author's additional affiliations are University of Connecticut and Yale University.