The assumption of serial operation of the three stages should be supported with evidence or disclaimed.

Massaro's discussion *of bottom-up versus to-down processing* is oversimplified. The statement that lipreading provides bottom-up information seems fair within the task of identifying articulatory motor movements, but lipreading would seem to provide top-down or contextual information for the task of deciding what was heard. This line of reasoning suggests that a visual speech display should be adequate to take the place of lexical context within the phoneme restoration effect. That is, if a brief segment of the sound track of a videotape were replaced by noise, it should be perceptually restored by the uninterrupted visual display even if the missing segment were embedded in a nonsense string.

The debate about *categorical perception* seems correct but incomplete. It ignores the fact that there generally are peaks in the discrimination function at the category boundaries, and that the origin of these peaks is controversial. Theoretically, they could be evidence for an even weaker, but more viable, form of categorical perception. In addition, the developmental discussion should interpret the finding that discrimination of sounds within a phonetic category in the infants native language declines markedly across the first year of life (see Werker 1989 for a review).

The debate about *modularity in speech perception* also seems incomplete. The support for modularity (e.g., Liberman & Mattingly 1989) comes not only from trading relations among cues, which Massaro discussed in detail, but also from duplex perception (speech and nonspeech percepts arising concurrently from the same stimuli). It can probably be accounted for by learned perceptual modes rather than innate modules, but it needs to be interpreted within FLMP.

Finally, although this is not truly a criticism, the locus of integration relative to the conventional, generic descriptions of information processing remains unspecified. These are interesting types of intermodal effects that should occur only if integration occurs sufficiently late in processing. For example, what if two complementary stimuli were presented within the McGurk paradigm (e.g., both visual /ga/ matched with auditor}' /ba/, and the converse), either by a male and a female speaking concurrently or by one person pronouncing two syllables in rapid succession? In which situations would the auditor)' and visual cues be recombined according to their phonetic identities? If recombinations occurred with successive presentation, would this suggest that perceptual syntheses of visual and auditory cues can be revised retroactively? Massaro's book is most useful in provoking such research questions.

Categorical perception of speech: A largely dead horse, surpassingly well kicked

Robert G. Crowder

Department of Psychology, Yale University, New Haven, CT 06520
Electronic mail: *rgcrowfgyalevm.bitnet*

Twenty years ago I enjoyed startling friends and frightening undergraduates with the strong form of the claim that speech perception was categorical. Soon, studies by Barclay (1972) and Pisoni (1973) among several others, convinced me that people, after all, had access to continuous information, even for continua of stop-consonant syllables. Although quibbling has continued about what Massaro (1987, p. 118) calls the weak form of the hypothesis, I assumed that many informed bystanders, like me, really didn't *need* convincing of the general point. Thus, to find so much of the present book devoted to a relentless assault on categorical perception struck me as anticlimactic. But I leave to others the debate whether this issue was already nearly closed.

If categorical perception, in its original sense, was too solidly discredited to have been taken seriously in the late 1980s, almost the opposite can be said of the Fuzzy Logical Model of Perception (FLMP), which is just as monotonously upheld by the research reported here. Not only is it victorious in the speech data, but in such diverse applications as person impressions, category judgments, and sentence interpretation as well. The FLMP simply works too well, everywhere it has been applied. Ironically, the case for this model would actually be enhanced by the discovery of some domain in which *it failed* to beat all contenders in sight. This would be the exception that proved the rule, a concept I used to think silly, but now appreciate. The FLMP attracts too much attention in this book for my taste.

Which is a pity to the extent that the FLMP and the argument about categorical perception get in the way of other commanding virtues in Massaro's approach. I count three of these, and *they* form the paradigm promised by the title of the book:

First, the book is data-intensive, without once losing sight of important ideas. Good ideas are, after all, cheap, and what distinguishes us as experimental psychologists from others in the cognitive sciences is that we demand public, empirical tests of our ideas, whereas others use different criteria for justifying their belief. Massaro's book affirms the centrality of carefully arranged and abundant evidence to our science. In an area where some main ideas have derived from linguistics, this emphasis is all the more welcome.

Second, and more specifically about the testability of ideas, Massaro makes it clear that a hypothesis is not properly confirmed or falsified qualitatively, in a vacuum, but must rather be tested against a viable alternative hypothesis, both of them preferably worked out in enough detail to make quantitative predictions. I don't claim, by the way, that my own hands are clean by this criterion, but I wonder if some of the false starts and arguing-past-one-another that have occurred in my specialty would have been avoided if we had adopted the model-comparison principle years ago.

Finally, although Massaro's book is closely focused on audiovisual speech perception, it is also seriously attentive to a wide range of other applications in psychology, including the development of perception, decision-making in general, and social impression-formation, as well as the other applications mentioned above. Of course, this Catholicism should not be interpreted as a purely transcendental virtue, since one of Massaro's strong contentions is that speech perception shares information-processing characteristics with other human activities, outside of language. Thus he is advancing this case when he reports parallel findings in other areas. More research programs, perhaps similar to this one in method, will be necessary to decide whether speech is "special" because (a) the stimuli themselves are special, because (b) the brain locus is demonstrably different from other kinds of processing, or because (c) the kind of information processing applied to speech signals is different from other modalities. Even scholars who might disagree vigorously on these three alternatives might nevertheless agree that Massaro's paradigm provides a good way of finding out.

Straw modules

Anne Cutler

MRC Applied Psychology Unit, 15 Chaucer Rd., Cambridge CB2 2EF, United Kingdom
Electronic mail: *anne@mrc-apu.cam.ac.uk*

Massaro's enterprise is impressive, and there are many ways he could have brought it to a resounding conclusion. But the final paragraph he chose begins "If the present generalizations of the

current research are reasonable, they strike a blow to the concept of modularity" (p. 281); his conclusion thus returns to the theme of his introduction (pp. 2-4), in which the concept of modularity was described, and its falsification in the following chapters prophesied. Massaro apparently sees modularity as the principal objective of his theoretical attack. Note that Massaro is not loath to tilt at sacred cows. Categorical perception is another prominent target of his lance. But modularity is the target he chose to set his sights upon in opening and in closing his book.

This is all very curious, because nothing in Massaro's work is crucially damaging to modularity; rather, his results offer the concept of modularity significant support. Consider the case for modularity as put by Fodor (1983; see also multiple book review in *BBS* 8(1) 1985). Fodor proposed nine properties of input systems by which their modularity was defined (Massaro lists these properties on pp. 2-3). In fact, these nine properties have not all been debated to the same extent. Most of the action in psycholinguistic argument has centred on principle five, the informational encapsulation of modular systems; this was perhaps only to be expected from a field which has disputed autonomy versus interaction above all else for the better part of two decades. To a lesser extent the principle of mandatoriness of operation has also proved contentious.

Now how does Massaro's research address the issues with which Psycholinguistics is obsessed? In one major way, at least, Massaro allies himself firmly with the autonomous camp. A crucial aspect of Massaro's Fuzzy Logical Model of Perception is the independence of separate sources of information in perception and recognition. Independence of information derived from separate sources is precisely what encapsulation requires; if the evaluation of one source of information is affected by what is coming in from another source of information, then the capsule has been penetrated and the principle is violated. Fodor argues that such dependent processing does not occur in modular input systems. Massaro is equally adamant that it does not occur in any of the domains of his research. Insofar as Massaro has made this case, he has made about the most important case for modularity.

That Massaro is perfectly aware of this is clear from his brief allusions to connectionist models. Interactive-activation models allow activation from higher levels of processing to feed back to lower levels of processing, thereby altering the operation of the lower-level processes and hence obliterating their independence. As Massaro points out, such models would make predictions about the listening-plus-lipreading situation which are not borne out by his data. For instance, it would be reasonable to expect an unambiguous lipreading cue (such as closed lips for a bilabial) to affect the relative activation of auditory features for various phonemes, rendering the features of bilabial phonemes more active and of nonbilabial phonemes less active. Neither the facilitative nor the inhibitory prediction finds support in Massaro's results; he argues that this outcome "contradicts a fundamental assumption of TRACE" (p. 171) and "provides important constraints on the assumptions that are viable for connectionist models" (p. 281). Thus he shows no reluctance to draw out the implications of his findings with respect to autonomy versus interaction of processing levels.

The autonomy-interaction debate has become far more sophisticated in recent years. It is no longer assumed that the mere presence of context effects poses problems for an autonomous view; context effects can occur when multiple outputs of lower-level processing - each even equiprobable given a degraded input - are evaluated at a higher level (see e.g. Norris 1986; Marslen-Wilson 1987). Again, this is a position of which Massaro approves: "the bottom-up information and the top-down information . . . function as independent sources . . . [which] illustrates once again that a positive effect of context does not necessarily imply that low-level processes . . . are dependent on higher-level processes" (p. 269). Clearly, he is an autonomist through and through.

Why, then, does Massaro consider himself to be arguing against modularity at all? The answer appears to be that he takes issue with another characteristic ascribed to modular systems, namely, domain-specificity. This is the principle that modular systems operate only upon their particular class of inputs, and are closely tuned to the specific properties of those inputs.

Domain-specificity is not one of the central pillars of the modularity temple. As Fodor points out, in one sense it is entirely trivial to say that the mechanisms which process language are domain-specific because language is what they process. But there are certain nontrivial aspects of domain-specificity. One is that processing by a modular system should be initiated only by input of its particular kind. Another is that the operations of the modular system should be sensitive to the particular characteristics of what is being processed.

What domain-specificity does *not* entail is that the operations of a modular system need be unlike the operations of other systems in every respect. Indeed, it would be astonishing if this were so. General constraints on processing by systems realised in neurological material will of course apply. Although a modular system will not draw upon resources shared with other systems, there is no reason why it should not carry out operations on its own type of input which are similar in kind to the operations carried out by other systems on their particular inputs. This is where Massaro appears to have misled himself: "by definition, a specialized process should not follow general principles" (p. 4). Massaro seems to have got hold of the wrong definition.

To clarify' this distinction, imagine, a coin-sorting device which consigns coins of different denominations to different processing operations on the basis of weight. Since each operation is initiated by the weight characteristics specific to a particular denomination of coin, the restrictive initiation principle of domain-specificity is satisfied. We can also build in lots of denomination-specific characteristics of the processing that goes on - pennies get painted black and shot out of a barrel, dimes get smashed to smithereens and mixed with sulphuric acid, quarters get baked in a cake, and so on. And we can have all these specific processes go on in entirely separate and noncommunicating compartments. But absolutely none of this alters the fact that the physical laws which govern the assessment of each coin's weight, and which further govern the way the coins tumble into the compartments, are the same for all coins; nor does the applicability of the laws of physics compromise the domain-specificity of the coin-processing. In fact, domain-specificity would not be compromised even if there were considerable similarity between the operations - if some part of each denomination's processing included a painting operation, for example.

Just the same sort of distinction is true in perceptual processing. It is never going to be an argument against modularity to point to commonalities between different input systems either in general constraints or in operating characteristics. Yet that seems to by why Massaro thinks he is against modularity: "principles . . . of bimodal speech perception prove to be relevant to a variety of behaviors" (p. 4). But the principles in question are extremely general - feature evaluation, integration, and pattern classification; "any model of recognition would have to assume stages functionally equivalent to these" (p. 152). Quite so. Thus Massaro's chief case against modularity is null. In his final paragraph he makes, for the first time, a slightly different claim, namely, that processing of a general cognitive kind has been shown to engage supposedly encapsulated modules; but this only widens the definition of general cognitive processing, quite apart from the fact that the violation of encapsulation requires engaging general cognitive resources by an encapsulated module rather than vice versa. He follows this, however, by reiterating the principal point of his research: "the same fundamental processes contribute to an impressive variety of perceptual and cognitive actions" (p. 281). Since there is

nothing in this general claim which is incompatible with modularity, it is a pleasure to welcome Massaro to the modular fold.

Models in the mind, modules on the lips

Beatrice de Gelder and Jean Vroomen

Departments of Psychology and Philosophy, Tilburg University, 5000 LE Tilburg, The Netherlands
Electronic mail: *beadegel(cvhtikub5.bitnet*

As a consequence of studies of both sign language and lipreading it is now widely agreed that a comprehensive theory of speech perception (and production) must capture the common properties of the different modalities in which speech can be perceived. Generally speaking, evidence from sign language and from lipreading has been welcomed by modularist speech theorists. In contrast, Massaro's contribution builds on more than a decade of his research on speech perception in the auditory and visual modality and arrives at a refutation of modularism. Modularists welcome lipreading evidence because they regard it as support for the view that speech perception draws upon a representational repertoire which is both language specific and amodal or abstract. Massaro believes that the bimodalitv of speech perception offers a spectacular insight into the general way the organism processes information, for example, by integrating multiple sources of information. The existence of this general process is the core of his refutation of modularism.

We believe that the issue of information integration is orthogonal to the matter of modularity of speech perception. What matters is the domain specificity of the representations over which integration processes operate. Next, we consider the relevance of language-specific deficits in the visual and auditory modality observed in developmental dyslexia.

1. Modularity is about representations. Massaro claims that information integration is the basic principle of information processing in all domains of cognitive functioning. What aspects of the claims of modularism (as epitomized in Fodor 1983) does this refute? The core of modularism is the view that all cognitive processes are inferential but that some of them are subserved by a special purpose computational system having a restricted database. So both modularism and Massaro agree that information processing is hypothesis testing and decision making. Going beyond this, modularism proposes a cognitive geography of decisions. Some decision processes are local because they are based on limited knowledge (e.g., they are domain specific and encapsulated). Massaro claims that a general model of decision making is enough - that one does not need to go through the trouble of designing a (modular) model of the internal knowledge environment of the organism to understand how decisions on inputs are reached. The crux of the disagreement is categorical versus continuous perception. In Massaro's view, categorical identification results from an integrative decision about a previous continuous evaluation process taking place separately in various dimensions. Modularism claims that evaluation itself is a decision process located in the language module and resulting in phonological representations. (This view leaves room for postmodular decisions about modular phonological representations).

How does evidence for modality specific coding square with either view? The answer is complicated by the fact that the notion of dimension and that of modality are orthogonal. Massaro uses the multimodality of speech input as a privileged example to show that information processing consists of integrating the input coming from various dimensions. This way he indicates that he treats modalities as dimensions and implies that phonological category decisions require integrating bimodal information. Obviously, in normal circumstances both the visual and the auditory information are present. Still, when auditory information is not deteriorated, it is perfectly possible to identify speech in the absence of information from the visual modality. But perception in the auditory modality still results from the integration of multiple features. Perception of sign language in congenitally deaf subjects offers a similar example of multiple dimensions of a stimulus within one modality. The real issue is categorical perception within one and the same modality.

2. *Developmental dyslexia.* Detailed studies of acquired dyslexia support the view that we are dealing with a specific impairment in the domain of written language skills (Shallice 1988) which leaves intact decision processes in other domains of linguistic and nonlinguistic information processing. Such disorders support modularism because as noted above, modularism is a thesis about the nature of representational resources required in decision processes. There is a similar consensus about the language specificity of the deficits observed in developmental reading disorders. We have explored the idea that reading acquisition disorders might be related to spoken language deficits. By looking also at speech perception in the visual modality one might get a fuller picture of the speech processing abilities of young dyslexics. Alternatively, this approach could reveal the existence of bimodal integration difficulties.

One could use a stimulus tape made by Massaro (Massaro & Cohen 1983a) to examine this issue. Our results (de Gelder & Vroomen 1988) show that young dyslexics have less robust auditory speech categories when their performance is compared with that of control groups (both reading age and chronological age). This confirms the findings of Werker and Tees (1986). Our subjects also lag in processing speech information in the visual modality. These modality-based deficits show a significant correlation which could suggest a deficit in representational resources shared by visual and auditory information processing. Does the McGurk illusion also suggest a common resource explanation? does it show late categorization within each modality? or does it illustrate late integration of both modalities, as Massaro believes it does?

When presented with auditory /ba/ and a visual /ga/, subject's report hearing a phoneme that is a fusion of both, a /da/. On presentation of an auditory /da/ and a visual /ba/, the subject reports a blend, /bda/. Massaro's model (which does not give the full details of the representations activated at the various stages in processing) explains the former case but might have difficulty with the latter. He treats the perception of/bda/ as the identification of a single integrated speech event which is no exception to the general integration formula. The /b/ in /bda/ is indeed an integrated percept because subjects never report /mda/ or /pda/. There is, however, no visual influence in the /da/ part of/bda/. In other words, there is conflict plus integration.

The extent to which there can be conflict between modalities would appear to be closely related to the robustness of the representations in each modality. Indirect evidence about robustness of modality-specific coding has been obtained by using a serial recall paradigm (Massaro 1987, p. 50). Adapting the paradigm of the auditory recency effect in serial recall, Campbell and Dodd (1980) found that lip-read lists show recency just as auditory presented lists do. This suggests that there is a common language source for both modalities.

In a recent study using a large population of subjects we find that a visual suffix has no influence on the recency effect of stimuli presented auditorily. More surprisingly, an auditory suffix with no visual articulation does not affect recency in a visually presented list (de Gelder *&c* Vroomen 1989). This suggests that besides shared resources for both modalities (modularism) or late integration of visual and auditory information (Massaro), there is still room for modality specific codes. One would need to know more about the influence of linguistic experience on speech perception to have strong views in these matters (Bertelson & de Gelder 1989).