

The Auditory Processing of Speech: From Sound to Words. Edited by
M. E. H. Schouten. Berlin: Mouton de Gruyter, 1992. 380 pp. DM 188.00

Reviewed by
BRUNO H. REPP*

This book contains the proceedings of the second conference on the Psychophysics of Speech Perception, held in Utrecht in July of 1991. The first conference on the same topic had resulted in a similar volume (Schouten, 1987). The present book contains 35 papers grouped into six chapters, compared to 46 papers in seven chapters in the earlier volume. The coverage of topics is quite similar, though this conference includes a few papers on word recognition, at the cost of fewer theoretical papers and fewer papers on perception of phonetic speech contrasts. Longer "keynote" papers alternate with shorter contributions. Nearly all authors are well-known experts in psychoacoustics or speech perception, and the quality of the contributions is generally high. The editor's introduction could have been reproduced here in lieu of a review, as it provides good summaries of the contents of all papers and includes some personal comments as well.

Chapter 1, "The Auditory System in Relation to Speech Perception", contains eight papers (by Duifhuis, Horst *et al.*, Delgutte and Cariani, Sachs *et al.*, Oomens *et al.*, Patterson *et al.*, Kohlrausch *et al.*, and Ehret) and is mainly for those interested in models of peripheral processing. The research reported here is concerned with explaining how the auditory system manages to transmit the information needed for speech perception, rather than with accounting for what is ultimately done with that information. The extremely lucid keynote paper by Patterson *et al.* on auditory models as preprocessors for automatic speech recognizers deserves special mention, as does Ehret's intriguing report on categorical perception and left-hemisphere advantage for species-specific acoustic communication signals in mice. Ehret makes the important point that, rather than comparing perception of speech stimuli by humans and animals, it may be more enlightening to compare human perception of speech with animals' perception of signals that are meaningful to them.

The second chapter, "Separation of Simultaneous Signals", presents progress reports from within a rather narrow but extremely active area of research that has attracted some of the best minds in psychoacoustics and (formerly) speech perception. All papers have a strong psychoacoustic orientation but are relevant in various degrees to the important problem of separating simultaneous speech streams from different sources. Palmer discusses auditory nerve responses to simultaneous vowels. Cooke briefly presents a computer model of peripheral auditory grouping. Darwin summarizes a series of ingenious experiments using speech and complex tones. For example, he has demonstrated that changing the fundamental frequency of the second formant of a synthetic syllable independently of the other formants makes that formant audible as a separate auditory object but subtracts it from the phonetic percept only when the harmonics in the

* Address correspondence to Bruno H. Repp, Haskins Laboratories, 270 Crown Street, New Haven, CT 06511-6695.

region of the formant can be resolved by the auditory system; when the harmonics are unresolved (at a low fundamental frequency), phonetic integration persists despite auditory segregation — a striking instance of “duplex perception” (cf. Liberman, 1982). In the following paper, Carlyon investigates the role of pitch-pulse asynchrony in the detection of fundamental frequency differences. Then, Summerfield presents experiments showing that coherent frequency modulation does not serve as a grouping cue in the separation of simultaneous vowels. Moore’s keynote paper provides a lucid and authoritative review of two “hot” phenomena in psychoacoustics, comodulation masking release and modulation discrimination interference, with some appropriately cautious speculations on their possible relevance to speech perception. In the final three papers, Fantini and Moore, van den Brink *et al.*, and Festen present empirical contributions relevant to these topics.

The third chapter, “Perception of Spectral Change and Timbre”, is a mixed bag. Houtsma samples a few interesting recent findings on pitch and timbre perception but (despite a promise in his title) has relatively little to say about the pertinence of these results to speech perception. Similarly, Versfeld’s work on the discriminability of changes in the spectral slope of noises and two-tone complexes is elegant but of uncertain relevance to speech perception, especially since the spectral change investigated is between rather than within stimuli. Lacerda’s paper on infants’ discrimination of synthetic speech-like stimuli is impressive only in terms of the number of babies tested; the results seem confusing and statistically unreliable. In the following paper, Schwartz *et al.* kindly make reference to my suggestion at the first conference of an “articulatory psychophysics” (Repp, 1987), but their discussion does not resemble what I had in mind, being concerned mainly with auditory modelling. Finally, Espinoza-Varas presents studies of burst level discrimination and gap detection, but seemingly loses track of his original motivation, which was to explore the correlation between tests of auditory acuity and accuracy in speech perception tasks.

Chapter 4, “Loss of Spectral and Temporal Resolution”, contains only three papers. The first one, by Shannon *et al.*, discusses the important topic of amplitude-envelope information in speech perception, particularly in patients with auditory brainstem or cochlear implants. These authors show that a nonlinear external amplitude transformation, which simulates the normal compressive cochlear function, benefits perception in these patients. The following paper by van Son *et al.* deals with perception of complex steady-state sounds by hearing-impaired listeners, at some remove from speech perception. Finally, ter Keurs *et al.* briefly summarize their elegant research on the effects of artificial spectral smearing (simulating limited spectral resolution in the hearing-impaired) on phoneme perception.

Chapter 5, “Phoneme Perception”, reflects the decline of what was once a thriving area of research. The first paper, by Uchanski *et al.*, deals with effects of token variability on vowel identification scores, a topic of some methodological importance in intelligibility testing. Its relevance to natural speech perception may be limited, however, by the fact that effects of token variability seem to arise primarily from listeners’ ability to recognize individual tokens when there are few of them — that is, an artifact of testing method. The following contribution, by Li and Pastore, is quite problematic. It presents labeling, reaction time, and similarity judgment data for synthetic consonant-vowel stimuli varying in the onset frequencies of the second and third formants (F2 and

F3). The goal was to evaluate a prototype model of consonant identification (without reference to Massaro's extensive work in this area; see, e.g., Massaro and Oden, 1980). Contrary to the authors' statement, the /ba/ stimulus with the shortest reaction time was not at all "near the center" of that category but immediately adjacent to stimuli labeled as /da/. Moreover, the set of stimuli to be identified as /ba/ or /da/ included tokens that were identified as /ga/ in an initial experiment. (In that experiment, F3 started at a lower frequency than F2 in some stimuli!) The two dimensions of a multidimensional scaling solution are interpreted as "category membership" and "category goodness", even though their respective correlation with F2 and F3 frequency is noted. A comment about phonetic trading relations seems misguided because F2 and F3 in this instance do not cue the same phonetic contrast. This flawed paper ends with the reassuring statement that the conclusions "do not necessarily reflect the views of the National Science Foundation". Schouten and van Hessen, in the following paper, eschew an introduction and immediately present results from categorical-perception experiments with stop consonants and vowels. They propose that stop consonants are discriminated on the basis of distances from a category prototype, rather than by a direct comparison of their auditory representations. This model makes their results for stop consonants appear even more categorical, but those for vowels even less so. A belief that this is the way it should be seems to be the only criterion for the validity of the model. Finally, Sawusch comments on Uchanski *et al.* and briefly reports data from a priming study suggesting that phonemes have position-specific perceptual representations.

The final chapter, "Word Perception and Beyond" opens with an excellent keynote paper by Frauenfelder on the interface between phonetic and lexical processing. After an instructive summary of findings from monitoring experiments, he reports some data of his own which suggest that phonetic factors may account, at least in part, for effects attributed to syllable structure in earlier studies. Frauenfelder stresses the importance and relative neglect of phonetic stimulus properties in studies of lexical processing. Unfortunately, that lesson is lost on Nusbaum and Henly, who describe speeded classification experiments showing that there is more integrality of phonemes within words than across word boundaries. While they kept the phonemic environment constant, it seems that they did not control phonetic word boundary cues. These are precisely the focus of Quené's following paper. Using ambiguous Dutch utterances such as *maag oud* vs. *ma goud*, he shows that phonetic cues are indeed important, though not entirely sufficient for word boundary decisions. Interestingly, artificially lengthened vowel duration had only an effect in utterances originally pronounced as *maag oud*, whereas lengthened consonant closure duration had only an effect when *ma goud* was intended. According to Quené, this contingency supports a "direct-mapping" model of lexical access. The penultimate paper, by Bosman *et al.*, might have been included in Chapter 4, as it deals with speech intelligibility in the hearing-impaired. The constraints imposed on phoneme perception by meaningful as compared to meaningless syllables (it is not clear whether they were presented in separate tests or intermixed) are shown to be effective in both normal-hearing and hearing-impaired listeners, perhaps somewhat more strongly in the latter. Finally, Terhardt's keynote paper presents a sweeping and somewhat idiosyncratic view of auditory information processing, his main point being that discrete recoding and hierarchical processing start right at the periphery. Although this paper presumably is the "beyond" in the chapter's title, it could just as well have been placed at the beginning of this volume.

Like all conference proceedings, and particularly those that adhere closely to the original format of the papers, the present volume has the drawback that most contributions summarize research that has meanwhile been published elsewhere in much greater detail (primarily in *The Journal of the Acoustical Society of America*). Therefore, the book has relatively little to offer to the specialist who is likely to be already familiar with most of the work. However, it would be a useful source of information for those in related fields who would like to inform themselves about what is going on in a very active subarea of psychoacoustic research. The high quality of most contributions and the generally careful editing (though there are some problems with phonetic symbols) make this book a worthwhile acquisition for libraries and individuals who can afford the prices of European publishers.

(Received: April 25, 1994)

REFERENCES

- LIBERMAN, A. M. (1982). On finding that speech is special. *American Psychologist*, *37*, 148 – 167.
- MASSARO, D. W., and ODEN, G. C. (1980). Speech perception: A framework for research and theory. In N. J. Lass (Ed.), *Speech and Language: Advances in Research and Practice* (pp. 129 – 165). New York: Academic Press.
- REPP, B. H. (1987). The role of psychophysics in understanding speech perception. In M. E. H. Schouten (Ed.), *The Psychophysics of Speech Perception* (pp. 3 – 27). Dordrecht: Martinus Nijhoff.
- SCHOUTEN, M. E. H. (Ed.) (1987). *The Psychophysics of Speech Perception*. Dordrecht: Martinus Nijhoff.