DOI: 10.1159/000078663

Keith Johnson Acoustic and Auditory Phonetics;

2nd ed. Blackwell, Oxford 2003 192 pp.; £ 17.99 ISBN 1-405-10123-7

This volume is a revised and expanded (second) edition of a textbook first published in 1997. Already the successful first edition received high praise and it seems to be a somewhat unrewarding task to (re-)review a book that has been labelled as '(...) well-thoughtout textbook (...) invaluable for students who want to know more than the basic elements of acoustic phonetics' (P. Ladefoged), 'the best phonetics textbook I've ever used' (J. Pierrehumbert), and 'an excellent phonetics text for students of linguistic phonetics and speech science' (C.A. Fowler) [all citations taken from the back cover of the bookl. However, this second edition contains a new chapter on speech perception and additional sections on digital filtering and cross-linguistic vowel and consonant perception. Thus, the present review sets out to give a brief sketch of the book's contents and its didactical techniques with special attention to the newly added topics.

Written for students in linguistic phonetics, in speech and hearing science, as well as in branches of electrical engineering and cognitive psychology dealing with speech, the book offers a nontechnical introduction to the acoustic theory of speech production, to acoustic properties of major classes of speech sounds, and to the auditory representation and perception of speech. These topics are conveyed in a lucid, sometimes almost colloquial style, which reflects the author's teaching experience and his awareness of the students' difficulties in grasping abstract concepts of acoustic and perceptual phonetics. The great didactic value of the textbook under review is enhanced by a lot of grey-shadowed boxes dealing with 'semi-related stuff' (as the author puts it). These semi-related stuff boxes informally address digressive questions of students, ranging from underwater speech to the aerodynamics of freeways, connecting acoustic terms and perceptual phenomena to everyday experience. One can highly agree with the author 'that some of the most interesting stuff in the book is in the boxes' (p. 2).

Before going to the description and discussion of the various chapters there are two general positive remarks to be made which apply to the textbook as a whole. First, the author does not only deal with acoustic and auditory phonetics in different chapters, he also successfully combines these two aspects of speech sounds throughout the book, thus demonstrating that the linguistic significance of speech acoustics greatly depends on the auditory systems of human beings. Second, despite the highly technical foundation of acoustics the author refrains from introducing too many formulas although 'the really interesting part of acoustic phonetics starts when you get out a calculator' (p. 2). Even in discussing several analysis techniques of digital sound processing (p. 28ff.) he succeeds in reducing mathematics to practical instructions without loss of (phonetically relevant) information.

The volume under review is subdivided into nine chapters, followed by references (pp. 168-173), answers to selected shortanswer questions (pp. 174-178) and an index (pp. 179–182). Each chapter is completed by exercises, consisting of a paragraph of 'sufficient jargon' (i.e. a list of technical terms dealt with in the preceding text) and of a series of short-answer questions. Unfortunately, not all of these questions are answered in the appendix mentioned above. Otherwise the exercises could have met the requirements of a textbook for an autodidactic study of the field. By the way, there is a typographical error in the answer section concerning the formula for antiformant frequency (p. 178); the numerator of the formula should read ' $(2 \times 1 - 1)c$ '

instead of ' $(2 \times 1 + 1)$ c', which in fact can be detected by reading carefully the relevant passages of the textbook, but nevertheless may puzzle a reader who is not too familiar with mathematics.

Chapter 1 (pp. 3–18) introduces basic acoustic knowledge such as sensation and propagation of sound as well as types of sound, and describes properties and usage of acoustic filters.

Chapter 2 (pp. 19-45) takes into account the fact that nowadays instrumentation used in acoustic phonetics has turned from oscilloscopes and spectrographs to hard- and software of computers, consequently deals with digital signal processing and gives an intuitive grasp of how computers perform acoustic analysis. Starting from the distinction of continuous versus discrete signals, problems of sampling and quantization in analog-to-digital conversion are presented, followed by a discussion of several signal analysis methods. It is particularly in this section, dealing among others with fast Fourier transform, digital filtering and linear predictive coding, that the author gets the basic concepts across without forcing the reader into mathematical details.

Chapter 3 (pp. 46-58) on basic audition briefly describes the anatomy and function of the peripheral auditory system and then discusses the differences between the acoustic and the auditory representations of sound, i.e. the mismatch of frequency and sound pressure/loudness scales of, for instance, a spectrograph versus the auditory system. This is finally demonstrated by a computer simulation of auditory response using a publicly available auditory model (ftp://ftp.ling. ohio-state.edu/pub/phonetics/ear). The explanations given in this chapter seem somewhat scarce and the drawings of the auditory system (p. 47) and the inner ear (p. 52) do not really convey much information. However, with the 'semi-related stuff box' on decibels (p. 50) and the figures comparing acoustic power spectra and auditory spectra (p. 53ff.) the essential topics appear to be emphasized.

Chapter 4 (pp. 59–78), new in the present second edition, deals with speech perception and tries to demonstrate that perceptual representations of speech are not solely based

on the auditory input but rather make use of additional information (e.g. the visual image of the talker) and the background of a lifetime of experience with language. Initially, the reader is guided through the necessary steps of a speech perception experiment on English fricatives which result in a confusion matrix. After discussing how to compute perceptual distances from a confusion matrix the perceptual maps of fricatives and of the feature [place] are presented. The latter is based on multimodal (auditory and visual) perception data and is included to demonstrate the McGurk effect. However, in the similarity data for 'b', 'd', and 'g' given on p. 71 there is something odd, as applying the explicitly worked out formula (p. 68) two of the three values simply cannot be derived from the response proportions in the corresponding table 4.3, on which they are claimed to be based. A final section impressively points out the limits of perceptual universality by the results of a discrimination test, which compared similarity judgments of American English and Mandarin listeners on Mandarin Chinese tones. Whereas Chinese listeners distinguish basically level tones from contour tones, English listeners rely on the pitch at the end of a syllable, thus distinguishing tones that end low from those that end high.

Chapter 5 (pp. 79–101) develops the acoustic theory of speech perception based on the source-filter model, mentioning also the quantal theory of speech and discussing LPC spectral analysis to capture information about the vocal tract filter function.

The remaining four chapters review the acoustic properties of major classes of speech sounds and relate them to their auditory characteristics and perceptual attributes.

Chapter 6 (pp. 102–119) deals with vowels and gives a detailed account of tube models of vowel production as well as of perturbation theory. The auditory and perceptual sections of this chapter clearly demonstrate the striking differences concerning the display of F₁ and F₂ between acoustic and auditory spectra, and compare perceptual vowel spaces for English, German, Thai and Turkish listeners. Based on exactly the same acoustic signals the divergence in the geometry of the perceptual

spaces are due to the different linguistic backgrounds of the listeners.

Chapter 7 (pp. 120–134) on fricatives follows the same sucessful structure: after the description of the (turbulent noise) source and the vocal tract filtering, quantal aspects in fricative production are surveyed, acoustic and auditory properties are presented, and finally dimensions of fricative perception are discussed by comparing the results of an identification test on English voiceless fricatives judged by American English and Japanese listeners. This section once more shows how specific linguistic dimensions are reflected in the perception of identical acoustic stimuli.

Chapter 8 (pp. 135-148), dealing with stops and affricates, differs from the preceding chapters in that it has, obviously enough, to take into account the dynamic aspects of the sound classes involved. As a prerequisite the acoustics of different phonation types is discussed (although relevant as well for fricatives and vowels), before sound sources and vocal tract filtering in stops and affricates are exemplified. Again there is a section on auditory properties of stops (illustrated by a figure comparing spectrograms and cochleagrams of [ba], [da], [ga] syllables) and finally the section on stop perception convincingly argues that a phenomenon like palatalization, changing a velar stop to a palatal affricate before /i/, cannot be explained as a purely articulatory process (that ought to result in a palatal stop, i.e. [k] > [c] / i), but is due to perception as well; as the perceptual distance between velar stops and palatal affricates is much smaller when followed by /i/ (as compared to that of the [-a] environment), the articulatorily driven change from [k] to [tʃ] is more likely to remain perceptually 'undetected'.

The concluding chapter 9 (pp. 149-168) on nasals and laterals mainly explains the more complicated vocal tract filtering characteristics due to specific side branches (mouth cavity in nasals, supralingual pocket of air in laterals) and the resulting changes of the spectral shape caused by anti-formants. For nasalized vowels the complex interplay between nasal formants, oral formants, and nasal anti-formants is lucidly described and the relatively small number of distinctive nasal vowels in languages seems to reflect their acoustic and perceptual peculiarities. A concluding section on nasal consonant perception displays perceptual maps of nasals for speakers of Malayalam, English, and Marathi, and convincingly speculates on language-specific attention to either the VC formant transitions or the CV release transitions of nasals in intervocalic position.

To sum up, the book under review covers the whole range of basic and some more advanced knowledge on acoustic and auditory phonetics by clear descriptions and easy-to-follow explanations. Notwithstanding the few critical remarks, the whole structure and approach make the book satisfying to read and rewarding to use.

H. Grassegger, Graz