REVIEW

Reviewed by **David Deterding**
University of Brunei Darussalam
dhdeter@gmail.com

This book covers a range of techniques for analysing speech spectra, reporting on the advantages and disadvantages of each method and describing the problems that may be encountered. After a brief introduction, Chapter 2 offers an overview of phonetics and signal processing, Chapter 3 provides a synopsis of the history of speech spectral analysis, Chapter 4 discusses the characteristics of the Fourier Power Spectrum, and Chapter 5 delves into alternative time-frequency representations of speech, including the Wigner-Ville Distribution and the Zhao-Atlas-Marks Distribution. In Chapter 6, the reassigned spectrogram is introduced, aiming to combine the good time resolution of the traditional wide-band spectrogram with the better frequency resolution of a narrow-band spectrogram, in order to allow the analyst to examine the formant frequencies of vowels without losing information about the fluctuating spectral patterns during each glottal cycle. Finally, Chapter 7 gives an overview of Linear Prediction techniques, especially the selection of parameter settings such as the order of the analysis and the choice of whether to use pre-emphasis or not.

The background of the author is in speech processing, and this is reflected in the authoritative summary of the various algorithms, the discussion of the implications of using each one, and the advice about how best they can be applied to the analysis of speech. As a result, the book provides a valuable overview of the history and current status of spectral analysis, offering some excellent material that many researchers will find exceptionally useful.

However, in contrast, the phonetics is a bit dodgy in places. For example, it is not clear why the vowel in *bought* is shown as [ɔə] (p. 9). On page 10, ten sounds are listed as the stops of American English: [pʰ b p tʰ t ɾ d kʰ k g]; but if allophones are to be included in this inventory of sounds, why is dark [ɫ] as well as clear [l] not listed (p. 11)? And then the definition of an approximant as 'a consonant whose articulatory constriction is slightly more open than that of a fricative' (p. 11) might seem fine until we note that the only consonant classified as an approximant is [l] (as [j w ɹ] are described instead as semivowels), and unfortunately this definition does not actually work too well for a lateral. Nevertheless, this is not a book about phonetics, so in reality a few flaws in the phonetic introduction do not have too much impact on its overall value in describing and analysing the various tools and techniques of spectral analysis.

One issue about the presentation of material is whether potential readers with a linguistics background can handle some of the mathematics that is included. The author does his best to put the most complicated technical details inside grey boxes 'for the benefit of those readers with sufficient background' (p. vii), yet still there remains quite a lot of dense material in the general-purpose text outside of these grey boxes. For example, one of the first equations in the introductory section on signal processing in Chapter 2 (p. 18) gives the Fourier series for the values c_k for all integers k over one period T as:

$$c_k = f_0 \int_T s(t) \exp[-2\pi i k f_0 t] \, dt$$

Now, this might seem really basic mathematics to engineers, but one wonders how many readers with a background in linguistics instead of engineering can cope with it, and how many might instead be put off by such equations throughout the book.

This is a pity, as there is plenty of excellent material which should be of value to a wide range of readers, especially those without a strong technical background. In the modern world in which we all have access to excellent, easy-to-use software such as Praat (Boersma & Weenink 2012), many researchers have lost sight of some of the underlying issues, such as the choice of window that should be adopted for each kind of analysis, or what order for linear prediction should be selected, and we are often not aware of the impact these different choices have on the results. Of course, in reality there are still many limitations on our understanding of speech spectral analysis, and these issues tend to get hidden away a bit too effectively in software such Praat, with the result that users may make easy use of the output without a proper understanding of the issues involved. And it is hoped that this book can offer an enhanced understanding about some of these issues for students and researchers involved in speech analysis.

One issue that might be raised is the extensive use of the author's own voice in the analysis. While this certainly works well in allowing a detailed comparison of various techniques of spectral analysis, in particular in demonstrating how the reassigned spectrum offers a superior insight into the fluctuating spectral characteristics of vowels, we are left with a nagging doubt about whether the same results would emerge for a greater range of other speakers. Nevertheless, it is of course not possible to cover all aspects of spectral analysis in a slim volume such as this, and the in-depth comparison of a range of spectral analysis techniques does offer a valuable insight into the benefits and drawbacks of a range of analytical tools.

In summary, this book is not a primer on speech analysis, and it does not deal with many practical issues, such as how to analyse formant transitions at the start and end of obstruents, or the details of how we should go about measuring voice onset time, and furthermore it almost entirely skips over the issues involved in tracking pitch. Instead, it focuses on methods of representing the speech spectrum, and it does this rather well. Indeed, it offers valuable information and advice about the spectral analysis of speech, and many speech researchers should benefit from the wealth of material that is presented, even if some might find the technical contents a bit overwhelming in places.

References

Boersma, Paul & David Weenink. 2012. Praat: Doing phonetics by computer. http://www.fon.hum.uva.nl/praat/ (accessed 15 October 2012).