Phonetica 2011;68:198–199
DOI: 10.1159/000331902

Pavel Machač, Radek Skarnitzl

Principles of Phonetic Segmentation

Epocha Publishing House, Prague, Czech Republic 2009
146 pp.
ISBN 978–80–7425–032–3

Within the speech sciences, consistency in speech segmentation is an issue, for example in corpora development and for reproducibility across studies. The objective of *Principles of Phonetic Segmentation* is to provide well-grounded and consistent segmentation conventions. The book is targeted as a user-friendly reference for psycholinguists, phoneticians and speech technologists.

In the first chapter the authors refer to inter-labeller discrepancies in previous literature, thereby defending their motivation to help reduce such discrepancies. This introduction is followed by definitions of what constitutes speech sounds and how boundaries between them can be conceptualized. The authors are clear that segment boundaries are to some extent arbitrary, however, necessary. The first chapter also includes explanations concerning language material (English and Czech, read and spontaneous speech), terminology and key reference points in the spectrograms.

The segmentation conventions are described in chapters 2–12, which form the main part of the book. The descriptions are based on consonant boundaries and are organized according to (i) consonant types and (ii) whether the consonant is intervocalic or following/preceding other consonants. The main consonant types described are plosives (chapter 2), fricatives (chapter 3), nasals (chapter 4), trills (chapter 5), glides (chapter 6), lateral alveolar approximants (chapter 7), obstruents (chapters 8 and 9) and liquids (chapter 9). Vowels are treated as a homogeneous category for the purposes of this book. Chapters 2–7 focus on intervocalic (i.e. vowel-consonant-vowel) segment boundaries, whereas chapters 8–10 focus on consonant-consonant boundaries. Chapter 11 deals with the segmentation of glottal stops preceding vowels, and chapter 12 addresses segment boundaries at the beginning and end of utterances.

Each chapter is structured to first give an articulatory and acoustic introduction to the different sound groups, followed by a description of general segmentation rules, a set of less obvious examples, additional segmentation guidelines, and a summary. Throughout the text the descriptions are illustrated by spectrograms and waveforms, highlighting the cues and elements described. The final chapter 13 summarizes the book and offers evidence that the inter-labeller segmentation based on the conventions presented in this book gives a high level of consistency compared to previous work. The final chapter also gives information on which segment boundaries are drawn least consistently based on the same conventions.

The book is in most respects user-friendly for the target audience, offering rigorous and sensible segmentation conventions. The most user-friendly aspect of the book lies in the extensive use of visual illustrations for the overall clearly written descriptions. Different examples are used to support the individual types of segment boundaries (e.g., plosive-vowel). Moreover, both canonical and less canonical examples are presented, which is likely to be useful for a user who, naturally, has to deal with a large amount of variation in his/her material. The book is also useful in that it gives a range of relevant tips on how to identify certain acoustic features in a spectrogram. These features include anti-formants (in nasals and laterals) and typical formant values for different sound types as well as for transitions between them. As such, this book also works as a reference and guide for phonetic analyses, supplementing more extensive literature on this matter.

The least user-friendly aspect of the book lies in its organization: It is not entirely transparent how the sound categories are defined and divided, and they are not indexed, which together might affect the user's ability to locate descriptions of particular segment boundaries. At first glance, it appears that the chapters are organized according to the phonetic properties of the corresponding consonants, i.e. plosives, glides, etc. However at the same time the descriptions are organized according to some phonological association. As an example, taps/flaps are initially described under intervocalic plosives

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

0031–8388/11/0683–0198
$38.00/0
Accessible online at:
www.karger.com/pho

(chapter 2), because they occur as allophones of 't'-type sounds in some varieties of English. But sounds that are phonetically similar to taps are also described in chapter 5, i.e. intervocalic trills (referring mainly to the Czech material). Also, the chapter on lateral alveolar approximants (a phonetic description) includes a description of vocalized /l/s (in English), which is confusing with regard to the previously defined sound category. It appears that the descriptions are based on allophonic associations in Czech and English, which might seem natural as these are the languages from which the material is collected. However, the rationale for this organization is not clear with respect to other languages, and might not be that helpful to non-Czech/English users.

A few further points deserve brief commentary. First, the authors make it explicit on several occasions that segmentation may not be possible without careful listening, particularly when the visual information in the spectrogram and/or waveform for some reason is hard to interpret. This is a point worth stressing for the reader, however, the examples used to illustrate these difficulties are not always particularly helpful since there is no audio data available. Second, the illustrations could have been used even more carefully: since the authors do not always indicate what parts of the spectrograms and/or waveforms they are describing, there are some ambiguities in the descriptions. Third, the approach taken in this book is more oriented towards finding evidence for the existence of presupposed segment boundaries, rather than towards providing a comprehensive description of the phonetic characteristics of different sound types, from which boundaries can then be defined and derived in a second step. Thus, this book might be most suited for researchers who are working with pre-established strings of segments, for example, scripted speech of a known language.

Overall *Principles of Phonetic Segmentation* is a successful effort at providing rigorous conventions for speech segmentation, and I would generally recommend it to the target audience, which in my view should include both students and professionals, who have previous experience with spectrograms and waveforms and general knowledge of phonetics and phonology. For this audience *Principles of Phonetic Segmentation* works well as a starting point and companion for developing segmentation conventions: Compared to alternative publications it contains more and better illustrated examples aimed at locating segment boundaries. However, the book is not on its own suitable as a reference for phonetic analysis, and because it does not offer any comprehensive background or discussion on the purposes of segmentation, this book is mainly for researchers who already know their reasons for segmenting speech.

Rein Ove Sikveland, York, UK