

***Phonetische und linguistische Prinzipien des
forensischen Stimmenvergleichs***

Michael Jessen (2012)

LINCOM Studies in Phonetics 247 pp

Reviewed by Sara Neuhauser

Finally, after more than twenty years there is a new German textbook that provides an overview of the most common practices and techniques in forensic speaker identification but at the same time summarises recent and ongoing research in this area. Michael Jessen's *Phonetische und linguistische Prinzipien des forensischen Stimmenvergleichs* is an introductory textbook that is very welcome in the German-speaking forensic phonetic and academic phonetic community. Jessen aims at three different audiences: first, at students interested in phonetics and forensic speaker recognition; secondly, at phoneticians and linguists who give lectures or seminars on this topic or are forensic phonetic experts themselves and interested in the ongoing discussion about specific data and research questions or want an overview of the recent literature; following the work of Künzel (1987), his third audience are those working in the police force or in legal professions, such as lawyers and judges who will find information about topics and analyses relevant to expert reports on forensic speaker comparison.

Jessen, who is an expert on speaker identification and does research and forensic casework at the federal forensics laboratory of the German Bundes-

Affiliation

University of Jena, Germany

email: sara.neuhauser@uni-jena.de

kriminalamt, identifies two thematic foci: first, an introduction into the practices and techniques of forensic voice comparison, i.e. illustration of casework and description of methods; and secondly, the development of a theoretical basis for forensic voice comparison.

The book is divided into six chapters and an appendix, the latter providing a summarising overview of the German sound system and basic articulatory principles. Chapter 1, *Forensische Sprechererkennung und Tonträgeranalyse*, outlines the tasks of forensic speaker identification and forensic acoustics (audio analyses). It introduces the relevant terminology and discusses major practical problems for forensic casework, such as restrictions on the quality and quantity of speech material, e.g. frequency limitations of telephone-transmitted speech signals, or mismatch problems, such as session mismatch, missing text identity or dealing with different languages. The first chapter closes by describing and explaining a complex of features that may be used for speaker classification. These features are not speaker-specific but characterise a group of speakers by determining a speaker's gender and age or, among other things, whether they belong to a group of speakers with a particular dialect, sociolect or foreign accent. The reader is taken through a wealth of examples from the author's and his colleagues' casework experience, convincingly illustrating the practical relevance of theoretical considerations.

Chapter 2, *Forensischer Stimmenvergleich im Überblick*, provides an overview of the basic concepts and methods of forensic voice comparison and explains the differences between interindividual and intraindividual variation. It also lists and categorises relevant speech features characterising a speaker that are commonly used for voice comparison. The concepts of similarity and typicality in respect to speaker discrimination tasks are elaborately discussed, emphasising that the questioned identity of two speakers (unknown speaker and suspect) is more difficult to prove if their voices show a strong degree of similarity, but high typicality, i.e. the analysed pattern is relatively common in the entire population. In the following two sections the concept of multidimensionality of the voice and the need to analyse speech features independently of each other are explained, and the differences between systematic and holistic methods in voice comparison are described. A further section gives a brief survey of voice quality and summarises voice qualities that are the most commonly analysed in forensic voice comparison. Finally, the last section is devoted to listing additional speech features characterising a speaker and complementing the list of features previously mentioned, e.g. breathing patterns, pauses and disfluencies, intonation, nonverbal vocalisations, speech rhythm, long-term average spectra and acoustic characteristics of consonants.

Chapters 3 to 5, *Grundfrequenz* (Chapter 3), *Formantfrequenzen* (Chapter 4),

Sprechtempo und gefüllte Pausen (Chapter 5), present the acoustic and theoretical concepts of fundamental frequency, formant frequencies, speech rate and filled pauses, since these are the acoustic parameters identified to be the most crucial for phonetic analyses in forensic casework (cf. Cambier-Langeveld 2007; Gold and French 2011). However, it is slightly irritating that the author does not explain his reasons for choosing to focus on these parameters before Chapter 4. Nevertheless, these chapters are all transparently and logically structured. Each chapter begins by explaining the basic acoustic and theoretical concepts and methods of analyses. They continue by referring to the methods' application in forensic contexts, including a discussion of consequences for forensic analyses, results of recent research as well as perspectives for future research. Particular emphasis is placed on the analyses of fundamental frequency and its variability (Chapter 3), formant and long-term formant analyses (Chapter 4) and, in Chapter 5, on the analyses of articulation rate and fundamental frequency and rate of filled pauses. Jessen returns repeatedly to the discussion about the importance of typicality and provides necessary population statistics, e.g. for the distribution of fundamental frequency and its variability (Section 3.4), for long-term formant distribution (Section 4.3) and for the distribution of articulation rate (Section 5.3). These statistics are mainly based on the *Pool 2010* corpus compiled by the German Bundeskriminalamt (Jessen, Köster and Gfroerer 2005) but also on other studies.

Chapter 3 starts with step-by-step explanations of the acoustic concepts of fundamental frequency and how to determine a speaker's mean f_0 , with and without the help of programs for speech analyses, such as *praat* (Boersma and Weenink 2013). There is a clear explanation of editing pitch contours in *praat* to correct errors by the pitch extraction algorithm. Jessen also discusses which measure of central tendency (mean, mode or median) might be more appropriate to describe the average fundamental frequency of a speaker. Section 3.3 describes the *Pool 2010* speech data corpus. The following sections focus on conclusions from forensic casework and on research results on the distribution of f_0 and its variability. Finally, Jessen briefly summarises studies on the integration of fundamental frequency analyses in automatic speaker identification.

Chapter 4 approaches the subject of formant analysis. Jessen starts with a comprehensive explanation of the basic concepts of Fourier synthesis and Fourier analysis and continues with the source-filter theory and the modelling of the resonances of a neutral vocal tract. He then moves on to a description of how to analyse formant frequencies with the help of *WaveSurfer* (Sjölander and Beskow 2005) and examples showing how to display the distribution of

formant frequencies. As in Chapter 3, coloured figures with screenshots of *praat*'s or *WaveSurfer*'s windows or formant plots help the reader to follow the explanations. Emphasis here is placed on the application of long-term formant analysis in a forensic context. The method of analysis using automatic formant tracking is explained, together with the necessary requirements (such as adjusting the formant track) and possible restrictions, e.g. poor quality or quantity of speech material. The final two sections are devoted to examples from forensic casework, recent research results and further methodological perspectives, e.g. from speech technology but also from sociophonetics.

The first part of Chapter 5 begins by describing and discussing the differences between speech rate and articulation rate, introduces possible methods to measure articulation rate and provides an overview of relevant research results, e.g. statistics on the distribution of articulation rate in spontaneous and read speech. As in the former two chapters, the first part of Chapter 5 closes with special reference to further perspectives for research in speech technology but also for fundamental research in phonetics and phonology. The chapter then moves on to its second part focussing on filled pauses. The results of a recent research project on idiosyncratic patterns of filled pauses are summarised, providing a classification scheme as well as data on the rate and fundamental frequency of filled pauses in different recording conditions.

The final chapter, *Die Rolle der Linguistik*, provides a theoretical introduction, necessary for readers without any background in linguistics. In the former chapters, the author has assumed basic linguistic knowledge or simplified particular descriptions and has referred the reader to the final chapter. For example, in Section 2.3 speech features are classified with the categories *organic*, *idiolectal* and *habitual* although the term 'idiolectal' is not explained in detail before the final chapter. Chapter 6 starts by summarily introducing the linguistic subfields or levels (phonology, morphology, syntax etc.), interdisciplinary fields such as psycholinguistics or sociolinguistics, and also the linguistic framework, e.g. functionalism, generative linguistics or linguistic typology and language universals. However, the contribution of these (sub-)disciplines to forensic speaker identification is not discussed further. The second section of the chapter is devoted to explaining the terms 'phonetics' and 'phonology' and to separate them from each other. This separation is based on two criteria: first, using empirical methods vs theoretical methods and, secondly, to study gradient or categorical aspects of speech. At the beginning of his book Jessen sometimes mixes the levels of phonetic description, phonological abstraction and graphemic representation, e.g. on p. 58 he writes, 'wenn ein Plosiv als Frikativ oder Gleitlaut ... ausgesprochen wird (z.B. ein [b] als ein [w] wie in Englisch "would") oder wenn ein Vokal wie [i] oder [u] als sog.

Schwalaut ausgesprochen wird ...' (*a plosive is pronounced as a fricative or approximant, e.g. [b] pronounced as [w] as in English 'would' or a vowel such as [i] or [u] are pronounced as schwa*). This confusion might be caused by an intentional oversimplification for the non-linguistic reader, but it is clearly problematic for a textbook that is directed at students. However, in this final chapter Jessen attempts to make a clearer separation of phonetics and phonology. The third section of Chapter 6 is focussed on the concept of idiolect. In this context Jessen discusses individual combinations of language varieties (e.g. speakers showing patterns of two different dialects) or of specific features within a variety. He further outlines phonetic-linguistic features which should be independent of a variety. It seems a little strange therefore that he chooses voice onset time as an example of such a feature, since VOT of fortis plosives has been shown not only to differ between languages, but also between (German) dialects. The final section of this chapter discusses how systematic linguistic description can be applied in forensic contexts and which restrictions commonly occur in forensic casework, e.g. speech material in foreign languages. Finally, three casework examples illustrate the process of collecting, describing, interpreting and analysing linguistic features for the speaker-discrimination task within a forensic voice comparison. Here the book shows again its most valuable contribution – providing informative insights into the practices of forensic casework, in this case with special emphasis on expert reports. Unfortunately, the last paragraphs are devoted to formal notation of phonological rules that do not seem necessary here.

This book impresses most with its detailed explanation and description of technical terms and methods, as well as the explanation of theoretical concepts and the provision of a wealth of practical examples. Even complex issues are, in most cases at least, explained and described very carefully, but still in simple terms. In many cases descriptions of methods are supplemented with illustrative figures and screenshots, e.g. showing pitch analyses and editing of pitch contours in *praat*. Perhaps even more importantly, the reader profits from the author's extensive practical experience and his competence to use his experience to demonstrate advantages but also practical restrictions of the forensic application of the methods described. All those features help the reader to follow the text easily and will be very useful in phonetics classes not only restricted to forensic phonetics, but also in other fields of applied experimental phonetics. The book being written in German is clearly an unfortunate disadvantage for the international phonetic community. On the other hand, both German-speaking students and practitioners will probably welcome and perceive it rather as a reader-friendly advantage while exploring an exciting field of applied phonetics.

The biggest challenge for Jessen was probably in aiming at different audiences. And this might also be the biggest challenge for some of his readers. On the one hand, a forensic phonetic expert may find parts of this book too general and some descriptions or explanations too elementary or oversimplified. On the other hand, it seems unquestionable that most linguistic students or even the most interested lay person with only basic knowledge in statistics and speech technology will not be able to cope with concepts and methods such as Equal Error Rate (p. 92) or Gaussian Mixture Models (p. 95), which are only briefly mentioned. However, most of the tough parts can be found in the final sections of Chapters 3 to 5, where the advanced and interested reader is addressed and referred to further research topics and studies. Though Jessen sometimes seems to struggle with the balancing act of meeting the demand of his various addressees, my overall conclusion is that this is an extremely useful book not only to the German-speaking forensic phonetic community, but also for students and practitioners who come into contact with forensic phonetics.

References

- Boersma, P. and Weenink, D. (2013) Praat: doing phonetics by computer. Computer program, Version 5.3.51. <http://www.praat.org/>
- Cambier-Langeveld, T. (2007) Current methods in forensic speaker identification: results of a collaborative exercise. *International Journal of Speech, Language and the Law* 14(2): 223–243.
- Gold, E. and French, P. (2011) International practices in forensic speaker comparison. *International Journal of Speech, Language and the Law* 18(2): 293–307.
- Jessen, M., Köster, O. and Gfroerer, S. (2005) Influence of vocal effort on average and variability of fundamental frequency. *International Journal of Speech, Language and the Law* 12(2): 174–213.
- Künzel, H.J. (1987) *Sprechererkennung: Grundzüge forensischer Sprachverarbeitung*. Heidelberg: Kriminalistik Verlag.
- Sjölander, K. and Beskow, J. (2005). Wavesurfer. Computer program, Version 1.8.5. <http://www.speech.kth.se/wavesurfer/>