

Dieter Maurer

Acoustics of the Vowel

Preliminaries

Peter Lang

Acoustics of the Vowel

Preliminaries

It seems as if the fundamentals of how we produce vowels and how they are acoustically represented have been clarified: we phonate and articulate. Using our vocal chords, we produce a vocal sound or noise which is then shaped into a specific vowel sound by the resonances of the pharyngeal, oral, and nasal cavities, that is, the vocal tract. Accordingly, the acoustic description of vowels relates to vowel-specific patterns of relative energy maxima in the sound spectra, known as patterns of formants.

The intellectual and empirical reasoning presented in this treatise, however, gives rise to scepticism with respect to this understanding of the sound of the vowel. The reflections and materials presented provide reason to argue that, up to now, a comprehensible theory of the acoustics of the voice and of voiced speech sounds is lacking, and consequently, no satisfying understanding of vowels as an achievement and particular formal accomplishment of the voice exists. Thus, the question of the acoustics of the vowel—and with it the question of the acoustics of the voice itself—proves to be an unresolved fundamental problem.

Dieter Maurer

Acoustics of the Vowel

Preliminaries



PETER LANG

Bern · Berlin · Bruxelles · Frankfurt am Main · New York · Oxford · Wien

Bibliographic information published by die Deutsche Nationalbibliothek

Die Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available on the Internet at <http://dnb.d-nb.de>.

British Library Cataloguing-in-Publication Data: A catalogue record for this book is available from The British Library, Great Britain

Library of Congress Control Number: 2015959255

Published with the support of the Swiss National Science Foundation within the scope of the pilot project OAPEN-CH.

This book is an open access book and available on www.oapen.org and www.peterlang.com.



It is distributed under the terms of the Creative Commons Attribution, Noncommercial, No Derivatives (CC-BY-NC-ND). License, which permits any non-commercial use, and distribution, provided no modifications are made and the original author(s) and source are credited.

Published as volume 12 of the series *subTexte*, edited by Anton Rey, Institute for the Performing Arts and Film, University of the Arts, Zurich. www.zhdk.ch/index.php?id=subtexte

Layout and cover design: Jacques Borel, Zurich

ISBN 978-3-0343-2031-3 pb.

ISBN 978-3-0351-0912-2 eBook

© Peter Lang AG, International Academic Publishers, Bern 2016
Hochfeldstrasse 32, CH-3012 Bern, Switzerland
info@peterlang.com, www.peterlang.com

All rights reserved.

All parts of this publication are protected by copyright.

Any utilisation outside the strict limits of the copyright law, without the permission of the publisher, is forbidden and liable to prosecution.

Contents

Acknowledgements

1 Introduction

Part I Prevailing Theory and Empirical References

14 1 Prevailing Theory

- 1.1 General Acoustic Characteristics of Vowel Sounds
- 1.2 Language-Specific Acoustic Characteristics of Vowel Sounds
- 1.3 Speaker Group-Specific Acoustic Characteristics of Vowel Sounds
- 1.4 Phonation Type-Specific Acoustic Characteristics of Vowel Sounds and Limitation to Voiced Oral Sounds
- 1.5 Limitation to Isolated Vowel Sounds
- 1.6 Limitation to Vowel Sounds as Monophthongs with Quasi-Constant Sound Characteristics
- 1.7 Speech Community-Specific Acoustic Characteristics of Vowel Sounds
- 1.8 The Prevailing Theory of Physical Vowel Representation
- 1.9 Formalising Prevailing Theory
- 1.10 Illustration

21 2 Prevailing Empirical References

- 2.1 General References
- 2.2 Empirical Reference for Standard German
- 2.3 Other Statistical References

Part II Reflections

32 3 Vowels and Number of Formants

- 3.1 Inconstant Number of Vowel-Specific Relative Spectral Energy Maxima in Sounds of Back Vowels and of /a–ɑ/
- 3.2 Inconstant Correspondence between Vowel-Specific Relative Spectral Energy Maxima and Calculated Vowel-Specific Formant Patterns
- 3.3 Inconstant Number of Vowel-Specific Relative Spectral Energy Maxima and of Calculated Vowel-Specific Formants
- 3.4 Addition: “Spurious” Formants

- 3.5 Addition: “Flat” Vowel Spectra
- 3.6 Addition: Inconstant Number of Vowel-Specific Formants in Synthesis
- 35 4 Vowels and Fundamental Frequency**
 - 4.1 Fundamental Frequency, First Formant and “Grade” of Vowels
 - 4.2 Fundamental Frequency, Spectral Envelope, Formant Pattern and “Grade” of Vowels
- 38 5 Formant Patterns and Speaker Groups**
 - 5.1 Fundamental Frequency, Spectral Envelope, Formant Pattern and “Grade” of Vowels Uttered by Children, Women and Men
 - 5.2 One Vowel, Different Formant Patterns
 - 5.3 Different Vowels, One Formant Pattern
 - 5.4 A Gap in the Reasoning
 - 5.5 Addition: Formant Patterns of Voiced and Whispered Vowel Sounds
- 45 6 Terms of Reference, Methods of Formant Estimation**
 - 6.1 Formant and Sound Spectrum
 - 6.2 Speaker Group and Vocal-Tract Size
 - 6.3 Formant Analysis and Objectivisation
 - 6.4 Formant Analysis, Fundamental Frequency and Speaker Group or Vocal-Tract Size
 - 6.5 Addition: Parameter Adjustments in Formant Analysis and Inconsistent References to Vocal-Tract Size
 - 6.6 Addition: Spectrum, Formant Pattern, Resynthesis
 - 6.7 Addition: Formant Analysis and Objectivity with Regard to Synthesised Vowel Sounds
 - 6.8 Addition: Formant Patterns and Resynthesis outside of the Framework of Prevailing Theory

Part III Experiences and Observations

- 56 7 Unsystematic Correspondence between Vowels, Patterns of Relative Spectral Energy Maxima and Formant Patterns**
- 7.1 Inconstant Number of Vowel-Specific Relative Spectral Energy Maxima and Incongruence of Vowel-Specific Formant Patterns
 - 7.2 Partial Lack of Manifestation of Vowel-Specific Relative Spectral Energy Maxima
 - 7.3 Addition: Resynthesis and Synthesis
- 59 8 Lack of Correspondence between Vowels and Patterns of Relative Spectral Energy Maxima or Formant Patterns**
- 8.1 Dependence of Vowel-Specific, Relative Spectral Energy Maxima and Lower Formants ≤ 1.5 kHz on Fundamental Frequency
 - 8.2 Vowel Perception at Fundamental Frequencies above Statistical Values of the First-Formant Frequency
 - 8.3 “Inversions” of Relative Spectral Energy Maxima and Minima and “Inverse” Formant Patterns in Sounds of Individual Vowels
 - 8.4 Addition: Whispered Vowel Sounds, Fundamental-Frequency Dependence of Vowel-Specific Spectral Characteristics and “Inversions”
 - 8.5 Addition: Resynthesis and Synthesis
- 64 9 Ambiguous Correspondence between Vowels and Patterns of Relative Spectral Energy Maxima or Formant Patterns or Complete Spectral Envelopes**
- 9.1 Ambiguous Patterns of Relative Spectral Energy Maxima and Ambiguous Formant Patterns
 - 9.2 Ambiguous Spectral Envelopes
 - 9.3 Ambiguity and Individual Vowels
 - 9.4 Addition: Resynthesis and Synthesis
- 66 10 Lack of Correspondence between Patterns of Relative Spectral Energy Maxima or Formant Patterns and Speaker Groups or Vocal-Tract Sizes**
- 10.1 Similar Patterns of Relative Spectral Maxima and Similar Formant Patterns ≤ 1.5 kHz for Different Speaker Groups or Different Vocal-Tract Sizes
 - 10.2 The Dichotomy of the Vowel Spectrum

- 10.3 Addition: Whispered Vowel Sounds and Speaker Groups or Vocal-Tract Sizes
- 10.4 Addition: Vowel Imitations by Birds
- 10.5 Addition: Resynthesis and Synthesis

70 11 Lack of Correlation between Methodological Limitations of Formant Determination and Limitations of Vowel Perception

- 11.1 Vowel Perception at Fundamental Frequencies > 350 Hz
- 11.2 Lack of Correspondence between Methodological Problems of Formant Pattern Estimation at Fundamental Frequencies ≤ 350 Hz and Impaired Vowel Perception
- 11.3 Addition: Lack of Methodological Basis of Determining Formant Patterns for Vowel Mimicry by Birds

Part IV Falsification

74 12 Empirical Falsification despite Methodological Limitations of Determining Patterns of Relative Spectral Envelope Maxima or Formant Patterns

- 12.1 Lack of Methodological Basis for Verifying Prevailing Theory
- 12.2 Systematic Divergence of Empirical Findings from Predictions of Prevailing Theory
- 12.3 Empirical Findings Directly Contradicting Prevailing Theory

Part V Commentary

78 13 Preliminaries

- 13.1 Impediments to Adjusting Prevailing Theory
- 13.2 Prevailing Theory as an Index
- 13.3 Excursus: Vowel Quality and Harmonic Spectrum
- 13.4 “Forefield”
- 13.5 Two Approaches
- 13.6 Phenomenology
- 13.7 Theory Building

87 Afterword

Materials

Materials Part I

- 98 M1 Prevailing Theory**
- 102 M2 Prevailing Empirical References**

Materials Part II

- 106 M3 Vowels and Number of Formants**
- 107 M4 Vowels and Fundamental Frequency**
- 112 M5 Formant Patterns and Speaker Groups**
- 118 M6 Terms of Reference, Methods of Formant Estimation**

Materials Part III

- 128 Note on the Method**
- 132 M7 Unsystematic Correspondence between Vowels, Patterns of Relative Spectral Energy Maxima and Formant Patterns**
 - M7.1 Inconstant Number of Vowel-Specific Relative Spectral Energy Maxima and Incongruence of Vowel-Specific Formant Patterns
 - M7.2 Partial Lack of Manifestation of Vowel-Specific Relative Spectral Energy Maxima
- 158 M8 Lack of Correspondence between Vowels and Patterns of Relative Spectral Energy Maxima or Formant Patterns**
 - M8.1 Dependence of Vowel-Specific, Relative Spectral Energy Maxima and Lower Formants ≤ 1.5 kHz on Fundamental Frequency
 - M8.2 Vowel Perception at Fundamental Frequencies above Statistical Values of the Respective First Formant Frequency
 - M8.3 “Inversions” of Relative Spectral Energy Maxima and Minima and “Inverse” Formant Patterns in Sounds of Individual Vowels

- 187 M9 Ambiguous Correspondence between Vowels and Patterns of Relative Spectral Energy Maxima or Formant Patterns or Complete Spectral Envelopes**
- M9.1 Ambiguous Patterns of Relative Spectral Energy Maxima and Ambiguous Formant Patterns
 - M9.2 Ambiguous Spectral Envelopes
 - M9.3 Ambiguity and Individual Vowels
- 217 M10 Lack of Correspondence between Patterns of Relative Spectral Energy Maxima or Formant Patterns and Age- and Gender-Related Speaker Groups or Vocal-Tract Sizes**
- M10.1 Similar Patterns of Relative Spectral Maxima and Similar Formant Patterns ≤ 1.5 kHz for Different Age and Gender-Related Speaker Groups or Vocal-Tract Sizes
 - M10.2 The Dichotomy of the Vowel Spectrum
 - M10.A Addition: Vowel Imitations by Birds
- 249 M11 Lack of Correlation between Methodological Limitations of Formant Determination and Limitations of Vowel Perception**
- M11.1 Vowel Perception at Fundamental Frequencies > 350 Hz
 - M11.2 Lack of Correspondence between Methodological Problems of Formant Pattern Estimation at Fundamental Frequencies ≤ 350 Hz and Impaired Vowel Perception

Experiments

- 252 E1 Number of Relative Spectral Energy Maxima and Number of Formants**
- E1.1 Sounds of Back Vowels Showing only One Lower Spectral Peak ≤ 1.5 kHz
 - E1.2 Sounds of Back Vowels Showing only One Pronounced Lower Formant ≤ 1.5 kHz
 - E1.3 Sounds of Single Front Vowels Showing Non-Corresponding F2 and F3
 - E1.4 Sounds of Back Vowels Showing No Pronounced Spectral Peak ≤ 1.5 kHz
 - E1.5 Sounds of Front Vowels Showing No Pronounced Spectral Peak > 2 kHz

- 254 E2 Patterns of Relative Spectral Energy Maxima, Formant Patterns and Fundamental Frequency**
- E2.1 Sounds of Single Vowels Produced at Different F0 Exhibiting Different Spectral Peaks and Different Calculated Formant Patterns: Part 1, Dependence of Formant Patterns on F0
 - E2.2 Sounds of Single Vowels Produced at Different F0 Exhibiting Different Spectral Peaks and Different Calculated Formant Patterns: Part 2, Vowel Intelligibility for Sounds at $F_0 > 500$ Hz
 - E2.3 Sounds of Single Vowels Produced at Different F0 Exhibiting Different Spectral Peaks and Different Calculated Formant Patterns: Part 3, Resynthesising a Formant Pattern at Different F0
 - E2.4 Sounds of Single Back Vowels Produced at Different F0 Exhibiting Inverse Spectral Peaks
 - E2.5 Special Note Concerning Inconstant Numerical Relationship between Calculated F0 and Formant Patterns
- 257 E3 Formant Pattern Ambiguity**
- E3.1 Formant Pattern Ambiguity in Natural Vocalisations
 - E3.2 Formant Pattern Ambiguity in Model Synthesis
- 258 E4 Patterns of Relative Spectral Energy Maxima, Formant Patterns and Age- and Gender-Related Vocal-Tract Sizes**
- E4.1 Comparison of Vowel-Specific Spectral Characteristics of Children, Women and Men Related to Different and Similar F0 of Vocalisations: Part 1, Natural Vocalisations
 - E4.2 Comparison of Vowel-Specific Spectral Characteristics of Children, Women and Men Related to Different and Similar F0 of Vocalisations: Part 2, Resynthesis
- 260 E5 Patterns of Relative Spectral Energy Maxima, Formant Patterns and Phonation Types**
- E5.1 Whispered Sounds Compared with Voiced Sounds at Different F0 in Utterances of a Single Speaker
 - E5.2 Whispered Sounds Compared with Voiced Sounds at Different F0 in Utterances of Speakers of Different Speaker Groups
 - E5.3 Sounds of Back Vowels Showing Three Spectral Peaks ≤ 1.5 kHz
 - E5.4 Sounds of Front Vowels Showing Two Spectral Peaks ≤ 1.5 kHz

262	E6	Patterns of Relative Spectral Energy Maxima, Formant Patterns and Vowel Imitation by Birds
	E6.1	Direct Comparisons of Selected Sounds of Humans and Birds
	E6.2	Resynthesis Relating to “Anomalous” Formant Patterns of Sounds of Birds
263	E7	Anomalous Vowel Spectra
	E7.1	Spectra with Increasing Number of Harmonics Equal in Amplitude (“Flat” Vowel Spectra)
	E7.2	Spectra with Increasing Number of Harmonic Pairs Showing Equal Amplitude Differences (“Ridged” Parts of Vowel Spectra)
264	E8	Aspects of Method
	E8.1	Formant Pattern Estimation Related to Non-Standard Parameters
	E8.2	Formant Pattern Estimation at $F_0 > 350$ Hz
	E8.3	Resynthesis of Sounds at Varying F_0 and Subsequent Formant Pattern Estimation
268		List of Figures
274		List of Tables
275		References