Book Announcement: Relevant Acoustic Phonetics of L2 English Focus on Intelligibility

Ettien Koffi
St. Cloud State University, enkoffi@stcloudstate.edu

Follow this and additional works at: https://repository.stcloudstate.edu/stcloud_ling

Part of the Applied Linguistics Commons

Recommended Citation
Available at: https://repository.stcloudstate.edu/stcloud_ling/vol10/iss1/7

This Article is brought to you for free and open access by theRepository at St. Cloud State. It has been accepted for inclusion in Linguistic Portfolios by an authorized editor of theRepository at St. Cloud State. For more information, please contact tdsteman@stcloudstate.edu.
Relevant Acoustic Phonetics of L2 English Focus on Intelligibility

Koffi: Book Announcement: Relevant Acoustic Phonetics of L2 English Focus

Published by theRepository at St. Cloud State, 2021
Intelligibility is the ultimate goal of human communication. However, measuring it objectively remained elusive until the 1940s when physicist Harvey Fletcher pioneered a psychoacoustic methodology for doing so. Another physicist, von Bekesy, demonstrated clinically that Fletcher’s theory of Critical Bands was anchored in anatomical and auditory reality. Fletcher’s and Bekesy’s approach to intelligibility has revolutionized contemporary understanding of the processes involved in encoding and decoding speech signals. Their insights are applied in this book to account for the intelligibility of the pronunciation of 67 non-native speakers from the following language backgrounds—10 Arabic, 10 Japanese, 10 Korean, 10 Mandarin, 11 Serbian and Croatian “the Slavic Group,” 6 Somali, and 10 Spanish speakers who read the Speech Accent Archive elicitation paragraph. Their pronunciation is analyzed instrumentally and compared and contrasted with that of 10 native speakers of General American English (GAE) who read the same paragraph. The data-driven intelligibility analyses proposed in this book help answer the following questions:

1. Can L2 speakers of English whose native language lacks a segment/segments or a suprasegment/suprasegments manage to produce it/them intelligibly?

2. If they cannot, what segments or suprasegments do they use to substitute for it/them?

3. Do the compensatory strategies used interfere with intelligibility?

The findings reported in this book are based on nearly 12,000 measured speech tokens produced by all the participants. This includes some 2,000 vowels, more than 500 stops, over 3,000 fricatives, nearly 1,200 nasals, about 1,500 approximants, over 1,200 syllables onsets, as many as 800 syllable codas, more than 1,600 measurement of F0/pitch, intensity, and duration measurements of no fewer than 539 disyllabic words. These measurements are in keeping with Baken and Orlikoff’s (2000:3) six principles of what constitutes a relevant acoustic phonetic analysis and also in accordance with widely accepted Just Noticeable Difference (JNDs) thresholds, and relative functional load (RFL) calculations provided by Catford (1987).

This book fulfills the long-awaited wishes of many luminaries of acoustic phonetics who have been calling for the application of acoustic phonetic findings to pronunciation teaching. They have often lamented the fact that pronunciation teaching still relies primarily on auditory impressions instead of being grounded in acoustic phonetic research and findings. I have endeavored long and hard to make acoustic phonetics understandable and relevant to the practical pursuit of pronunciation...
teaching. The instrumental approach used in this book makes the assessment of the intelligibility of L2-accented English more scientific. Science is understood here in a Ladefogean sense:

One of the objectives of any science is to be able to measure the things that are being described so that they can be expressed in terms of valid, reliable, and significant numbers that other people can check (Ladefoged 2001:166).

The primary goal of this book is for assessing intelligibility scientifically. However, users will also garner insights into how speech signals are extracted for speech-enabled artificial systems, i.e., automatic speech recognition (ASR) or Text-to-Speech (TTS), and for various voice biometric applications.
Acknowledgments

This book has been 15 years in the making. For this reason, the acknowledgments section should be long. However, since space is a rare commodity in a book, many individuals who have helped will not be mentioned by name. Instead, they will fall into the broader category of nameless people that I thank collectively. Even so, some people deserve to be named for the unique support and insights that they contributed to this book. Among them are my family members for their patience and longsuffering in being my “experimental” subjects and sounding boards on various aspects of the research that led to this book. My wife, Kim, has been a good sport in letting me stop her in mid-stream in a conversation to have her say a word again and again just to satisfy my insatiable appetite for the verification of allophonic realizations of such and such segments. The number of times my kids have rolled their eyes at me when I would ask their friends who come to our house to pronounce a word or two is incalculable. My son, Jeremy, threatened not to bring his friends over again because of my constant requests that they repeat words such as “ice” and “bike” when I was verifying the evidence of [ai] raising to [ʌɪ] in Central Minnesota. When my daughter, Corinne, was younger, she would quip and tell her friends that I measured vowels for a living. Later, when she was in sixth grade, she did her science project by comparing her pronunciation of vowels, with those of her mom and her grandmother. Her poster was acclaimed by the panelists in the “Most Innovative Award” category. Furthermore, her success was the singular act that convinced me that anybody could use Praat to do acoustic phonetic analyses because she did most of the spectrographic visualization and measurements by herself.

In the 15 years or so that it has taken me to put this book together, I have benefited from a wide variety of inputs from my students on what works, what doesn’t, and what is confusing. These insights have helped me hone my pedagogical strategies to spectrographic visualization techniques and measurements. Many graduates from the program are teaching English pronunciation confidently at universities, high schools, and elementary schools both in the US and around the globe. Some are applying their acquired skills in their practice as speech language pathologists, while others are working in the speech technology field or in speech-enabled artificial intelligent systems. I have also benefited from insights of colleagues who have attended my presentations at the Acoustical Society of America, the Annual Conference on African Linguistics, and the Pronunciation Second Language Teaching and Learning meetings. Anonymous peer reviewers of papers that I have sent for publication have been “annoying” with various requests for clarification, fearing that linguists or
applied linguists at large may not understand the highly specialized acoustic phonetic jargon and formulas that I use. Even though these requests have been aggravating at times, with hindsight, I see that they have helped me make acoustic phonetics accessible to many.

Last but not least, I want to express my heartfelt gratitude to my former colleague and friend, Emeritus Professor Rhoda Fagerland, for her extraordinary editing abilities. She has carefully edited this book, pointing out typos, inconsistencies, and other undesirable malapropisms. I’m extremely grateful to her for finding the time to edit in a similar manner two previous books and many papers that I have published. I’m extremely grateful to the Acquisitions Editor, Vijay Primlani, and the very competent staff at CRC Press, for their diligent work on this book. In spite of their keen editing eyes, I cannot guarantee that this book is completely error-free, but I can say with confidence that it would have contained many more errors without their expertise! Finally, I would like to thank all the giants of acoustic phonetics on whose shoulders I have stood to write this book.
### Contents

**Preface** iii  
**Acknowledgments** v  
**List of Abbreviations** xi  

1. **Review and Preview of Essential Concepts** 1  
   1.0 Introduction 1  
   1.1 A Quick Overview of Acoustic Phonetics 1  
   1.2 Hierarchical Organization of Speech 3  
   1.3 Syllables 14  
   1.4 Phonetic Transcription 15  
   1.5 Summary 18  

2. **Pictorial Definitions of Key Acoustic Phonetic Concepts** 21  
   2.0 Introduction 21  
   2.1 A Short History of Instrumental Phonetic Analysis 21  
   2.2 The Praat Revolution 23  
   2.3 An Overview of Spectrograms and Waveforms 24  
   2.4 Quantitative Analysis of Key Acoustic Realities 27  
   2.5 Textgrid and Annotations 36  
   2.6 Correlate Hierarchy 37  
   2.7 Summary 40  

3. **Preliminary Remarks on the Instrumental Assessment of Intelligibility** 43  
   3.0 Introduction 43  
   3.1 Grappling with the Concept of Intelligibility in Sociolinguistics 43  
   3.2 Speech Intelligibility Research on the Perception of Segments 48  
   3.3 Overview of the Auditory-Perceptual Theory 55  
   3.4 Intelligibility and World Englishes 61  
   3.5 Facilitating and Impeding Factors in L2 Intelligibility 66  
   3.6 Summary 70  

4. **The Acoustic Vowel Spaces of L2 English** 72  
   4.0 Introduction 72  
   4.1 A New Perspective on Vowel Intelligibility 73
# Relevant Acoustic Phonetics of L2 English: Focus on Intelligibility

## 4. Methodological Background
- 4.2 Methodological Background: 74
- 4.3 Methodological Rationale for /hvd/ Words: 78
- 4.4 L2-Accented English Vowels: 79
- 4.5 The Acoustic Vowel Space of Arabic English: 80
- 4.6 The Acoustic Vowel Space of Japanese English: 83
- 4.7 The Acoustic Vowel Space of Korean English: 86
- 4.8 The Acoustic Vowel Space of Mandarin English: 88
- 4.9 The Acoustic Vowel Space of Slavic English: 90
- 4.10 The Acoustic Vowel Space of Somali English: 93
- 4.11 The Acoustic Vowel Space of Spanish English: 95
- 4.12 Summary: 99

## 5. The Acoustic Phonetic Correlates of Stops in L2 English: 102
- 5.0 Introduction: 102
- 5.1 Basic Articulatory Overview: 103
- 5.2 The Perception of English Stop Consonants: 104
- 5.3 VOT in GAE and World Languages: 106
- 5.4 Relevance of VOT in Intelligibility Studies: 107
- 5.5 Data Analysis and Procedures: 110
- 5.6 Inter-speaker Variability in VOT: 111
- 5.7 The VOT of Arabic-Accented English Stops: 114
- 5.8 The VOT of Japanese-Accented English Stops: 116
- 5.9 The VOT of Korean-Accented English Stops: 117
- 5.10 The VOT of Mandarin-Accented English Stops: 119
- 5.11 The VOT of Slavic-Accented English Stops: 121
- 5.12 The VOT of Somali-Accented English Stops: 123
- 5.13 The VOT of Spanish-Accented English Stops: 126
- 5.14 Summary: 128

## 6. The Acoustic Phonetic Correlates of Fricatives in L2 English: 130
- 6.0 Introduction: 130
- 6.1 Definitions and Overview of English Fricatives: 131
- 6.2 An Overview of the Acoustics of English Fricatives: 131
- 6.3 Fricatives in GAE: 133
- 6.4 Fricatives in Arabic-Accented English: 137
- 6.5 Fricatives in Japanese-Accented English: 141
- 6.6 Fricatives in Korean-Accented English: 146
- 6.7 Fricatives in Mandarin-Accented English: 151
- 6.8 Fricatives in Slavic-Accented English: 155
- 6.9 Fricatives in Somali-Accented English: 159
- 6.10 Fricatives in Spanish-Accented English: 162
- 6.11 An Excursus into Affricates: 166
- 6.12 An Excursus into Duration: 167
- 6.13 Summary: 168
7. The Acoustic Phonetic Correlates of Nasals in L2 English 171
   7.0 Introduction 171
   7.1 The Articulatory Characteristics of Nasals 172
   7.2 An Overview of the Acoustics of Nasals 173
   7.3 Intensity and Duration of Nasals in GAE 179
   7.4 Nasals in Arabic-Accented English 181
   7.5 Nasals in Japanese-Accented English 183
   7.6 Nasals in Korean-Accented English 185
   7.7 Nasals in Mandarin-Accented English 186
   7.8 Nasals in Slavic-Accented English 188
   7.9 Nasals in Somali-Accented English 190
   7.10 Nasals in Spanish-Accented English 192
   7.11 Excursuses about Nasality and Nasalization 194
   7.12 Summary 198

8. The Acoustic Phonetic Correlates of Approximants in L2 English 200
   8.0 Introduction 200
   8.1 An Overview of the Articulation of Laterals 200
   8.2 Overview of Approximants in GAE 203
   8.3 Approximants in Arabic-Accented English 209
   8.4 Approximants in Japanese-Accented English 211
   8.5 Approximants in Korean-Accented English 215
   8.6 Approximants in Mandarin-Accented English 218
   8.7 Approximants in Slavic-Accented English 222
   8.8 Approximants in Somali-Accented English 224
   8.9 Approximants in Spanish-Accented English 225
   8.10 Excursus 227
   8.11 Summary 231

9. The Acoustic Phonetic of Complex Onset and Coda Clusters in L2 English 234
   9.0 Introduction 234
   9.1 The Importance of the Syllable: A Short History 234
   9.2 Overview of English Syllable Structures 244
   9.3 Syllable Structures in the SAA Text 246
   9.4 Overview of Arabic Syllabic Structures 248
   9.5 Overview of Japanese Syllabic Structures 250
   9.6 Overview of Korean Syllabic Structures 252
   9.7 Overview of Mandarin Syllabic Structures 255
   9.8 Overview of Slavic Syllabic Structures 257
   9.9 Overview of Somali Syllabic Structures 258
   9.10 Overview of Spanish Syllabic Structures 260
   9.11 General Observations and Constraint Ranking 263
   9.12 Excursus 271
   9.13 Summary 273
10. The Acoustic Phonetic Correlates of Lexical Stress in L2 English

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0</td>
<td>Introduction</td>
<td>275</td>
</tr>
<tr>
<td>10.1</td>
<td>A Quick Review of English Lexical Stress</td>
<td>276</td>
</tr>
<tr>
<td>10.2</td>
<td>A Phonological Definition of Lexical Stress</td>
<td>277</td>
</tr>
<tr>
<td>10.3</td>
<td>Linguistic and Paralinguistic Factors that Affect Lexical Stress</td>
<td>277</td>
</tr>
<tr>
<td>10.4</td>
<td>Measuring and Ranking of the Acoustic Correlates of Lexical Stress</td>
<td>280</td>
</tr>
<tr>
<td>10.5</td>
<td>Lexical Stress in GAE</td>
<td>287</td>
</tr>
<tr>
<td>10.6</td>
<td>Lexical Stress in Arabic</td>
<td>289</td>
</tr>
<tr>
<td>10.7</td>
<td>Lexical Stress in Japanese: A Prototypical Pitch-Accent Language</td>
<td>292</td>
</tr>
<tr>
<td>10.8</td>
<td>Lexical Stress in Korean</td>
<td>295</td>
</tr>
<tr>
<td>10.9</td>
<td>Lexical Stress in Mandarin</td>
<td>296</td>
</tr>
<tr>
<td>10.10</td>
<td>Lexical Stress in Serbian and Croatian</td>
<td>299</td>
</tr>
<tr>
<td>10.11</td>
<td>Lexical Stress in Somali</td>
<td>302</td>
</tr>
<tr>
<td>10.12</td>
<td>Lexical Stress in Spanish</td>
<td>304</td>
</tr>
<tr>
<td>10.13</td>
<td>The Role of Suprasegmentals in the Intelligibility of L2-Accented English</td>
<td>308</td>
</tr>
<tr>
<td>10.14</td>
<td>Summary</td>
<td>310</td>
</tr>
</tbody>
</table>

References 312

Index 325

About the Author 331
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>Articulation Index or Artificial Intelligence</td>
</tr>
<tr>
<td>AMI</td>
<td>Acoustic Masking and Intelligibility</td>
</tr>
<tr>
<td>ANSI</td>
<td>American National Standards Institute</td>
</tr>
<tr>
<td>AOA</td>
<td>Age of Arrival</td>
</tr>
<tr>
<td>ASR</td>
<td>Automatic Speech Recognition</td>
</tr>
<tr>
<td>CANS</td>
<td>Central Auditory Nervous System</td>
</tr>
<tr>
<td>CAPT</td>
<td>Computer Assisted Pronunciation Teaching</td>
</tr>
<tr>
<td>CBT</td>
<td>Critical Band Theory</td>
</tr>
<tr>
<td>COG</td>
<td>Center of Gravity</td>
</tr>
<tr>
<td>FIC</td>
<td>Fricative Intelligibility Criterion</td>
</tr>
<tr>
<td>F0</td>
<td>Fundamental Frequency</td>
</tr>
<tr>
<td>F1</td>
<td>Formant One</td>
</tr>
<tr>
<td>F2</td>
<td>Formant Two</td>
</tr>
<tr>
<td>F3</td>
<td>Formant Three</td>
</tr>
<tr>
<td>F4</td>
<td>Formant Four</td>
</tr>
<tr>
<td>IEC</td>
<td>International Electrotechnical Commission</td>
</tr>
<tr>
<td>ISO</td>
<td>International Organization for Standardization</td>
</tr>
<tr>
<td>GAE</td>
<td>General American English</td>
</tr>
<tr>
<td>GMU</td>
<td>George Mason University</td>
</tr>
<tr>
<td>LIC</td>
<td>Liquid Intelligibility Criterion</td>
</tr>
<tr>
<td>LOR</td>
<td>Length of Residency</td>
</tr>
<tr>
<td>MOA</td>
<td>Manner of Articulation</td>
</tr>
<tr>
<td>MOAPP</td>
<td>Mother of all Pronunciation Problems</td>
</tr>
<tr>
<td>MSDP</td>
<td>Minimal Sonority Distance Parameter</td>
</tr>
<tr>
<td>NIC</td>
<td>Nasal Intelligibility Criterion</td>
</tr>
<tr>
<td>OCP</td>
<td>Obligatory Contour Principle</td>
</tr>
<tr>
<td>PNC</td>
<td>Primary Nasal Consonants</td>
</tr>
<tr>
<td>POA</td>
<td>Place of Articulation</td>
</tr>
<tr>
<td>PSLLT</td>
<td>Pronunciation Second Language Learning and Teaching</td>
</tr>
<tr>
<td>SAA</td>
<td>Speech Accent Archive</td>
</tr>
<tr>
<td>SNR</td>
<td>Sound-to-Noise-Ratio or Signal-to-Noise-Ratio</td>
</tr>
<tr>
<td>SSP</td>
<td>Sonority Sequencing Principle</td>
</tr>
<tr>
<td>TESOL</td>
<td>Teaching English to Speakers of Other Languages</td>
</tr>
</tbody>
</table>
Relevant Acoustic Phonetics of L2 English: Focus on Intelligibility

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TESL</td>
<td>Teaching English as a Second Language</td>
</tr>
<tr>
<td>TEFL</td>
<td>Teaching English as a Foreign Language</td>
</tr>
<tr>
<td>TM</td>
<td>Template Model</td>
</tr>
<tr>
<td>TTS</td>
<td>Text-to-Speech</td>
</tr>
<tr>
<td>UPSIP</td>
<td>UCLA Phonological Segment Inventory Database</td>
</tr>
<tr>
<td>VOT</td>
<td>Voice Onset Time</td>
</tr>
</tbody>
</table>
Intelligibility is the ultimate goal of human communication. However, measuring it objectively remained elusive until the 1940s when physicist Harvey Fletcher pioneered a psychoacoustic methodology for doing so. Another physicist, von Bekesy, demonstrated clinically that Fletcher's theory of Critical Bands was anchored in anatomical and auditory reality. Fletcher's and Bekesy's approach to intelligibility has revolutionized contemporary understanding of the processes involved in encoding and decoding speech signals. Their insights are applied in this book to account for the intelligibility of the pronunciation of 67 non-native speakers from the following language backgrounds – 10 Arabic, 10 Japanese, 10 Korean, 10 Mandarin, 11 Serbian and Croatian "the Slavic Group," 6 Somali, and 10 Spanish speakers who read the Speech Accent Archive elicitation paragraph. Their pronunciation is analyzed instrumentally and compared and contrasted with that of 10 native speakers of General American English (GAE) who read the same paragraph. The data-driven intelligibility analyses proposed in this book help answer the following questions - Can L2 speakers of English whose native language lacks a segment/segments or a suprasegment/suprasegments manage to produce it/them intelligibly? If they cannot, what segments or suprasegments do they use to substitute for it/them? Do the compensatory strategies used interfere with intelligibility?

The findings reported in this book are based on nearly 12,000 measured speech tokens produced by all the participants. This includes some 2,000 vowels, more than 500 stop consonants, over 3,000 fricatives, nearly 1,200 nasals, about 1,500 approximants, over 1,200 syllables onsets, as many as 800 syllable codas, more than 1,600 measurement of F0/pitch, and duration measurements of no fewer than 539 disyllabic words. These measurements are in keeping with Baken and Orlikoff (2000:3) and in accordance with widely accepted Just Noticeable Difference thresholds, and relative functional load calculations provided by Catford (1987).

Ettien Koffi, Ph.D., linguistics, teaches at Saint Cloud State University, Minnesota. He is the author of four books and author/co-author of several dozen articles. His acoustic phonetic research is synergetic, encompassing L2 acoustic phonetics of English (Speech Intelligibility from the perspective of the Critical Band Theory), sociophonetics of Central Minnesota English, general acoustic phonetics of Anyi (West African language), acoustic phonetic feature extraction for application in Automatic Speech Recognition, Text-to-Speech, Voice Biometrics, and Intelligent Systems.