Book Review

Assessment in Second Language Pronunciation. By O. Kang and

A. Ginther (Eds.), London and New York: Routledge, 2018, 190

pp., $39.91 (paperback). ISBN: 978-1-138-85687-5 (pbk)

Dunlai Lin[1*] & Sen Liu[1]

[1] School of Foreign Languages and Literature, Beijing Normal University, Beijing, P.R. China

[*] Dunlai Lin, School of Foreign Languages and Literature, Beijing Normal University, Beijing, P.R.
China

The pendulum of L2 pronunciation research has swung back to a time when there is a revitalization in
research into assessment of second language pronunciation, with two edited volumes published within
two years, Isaacs and Trofimovich (2016) and Kang and Ginther (2018), the book reviewed here.
Assessment in Second Language Pronunciation takes pronunciation as a crucial aspect of oral
production and delves into important constructs like accentedness, intelligibility and comprehensibility
and the assessment of these constructs and sub-constructs in international communication and the
related validity issues. The volume also discusses application of speech technology in pronunciation
assessment, making it a highly-welcomed addition to the literature.

The volume is divided into two parts with an introduction by the two editors. In the introductory
chapter, Kang and Ginther firstly give a brief historical account of research on L2 pronunciation and its
assessment in particular, and moves on to an overview of this edited volume and its intended
readership.

The first part (Chapters 1-5) focuses on the current issues in pronunciation assessment. In the first
chapter, Thomson reviews how applied linguists have come to understand and evaluate L2
pronunciation, focusing specifically on definition and confusion of accentedness, intelligibility and
comprehensibility in the literature. Thomson then probes into features that affect accentedness,
intelligibility and comprehensibility ratings and measures that have been used to assess these constructs.
He also brings forward acceptability as a dimension of pronunciation.

In Chapter 2, Harding explores the validity issue in pronunciation assessment by quoting a
tongue-twister designed for the recruitment of call-center workers. He then gives a historical account of

the conceptualizations of validity, and works out a validity framework for pronunciation assessment based on the most recent view of validity (evaluation, generalization, explanation/extrapolation, utilization), pointing out unique questions relating to pronunciation assessment, namely construct definition, listener effect, automated scoring and social identity issue. In terms of current contributions, Harding firstly discusses evaluation, arguing that existent scales often mix nativeness principle with intelligibility principle, and uncertainty in factors affecting intelligibility also renders potential weakness for pronunciation assessment. In generalization, Harding mainly focuses on rater variability in view of rater's familiarity with a speaker's accent and attitudes towards it and the impact of tasks. Research supporting an explanation inference aims to identify criterial features of pronunciation at different score levels or explore correlations between pronunciation scores and other related measures, while research in utilization is scarce. In rethinking constructs, Harding puts forward the idea of phonological accommodation as potential future research. He concludes this article by proposing effect-driven approach to pronunciation assessment design.

Dimova (Chapter 3) addresses pronunciation assessment in the paradigm of World English (WE) and more recent English as Lingua Franca (ELF). She takes a balanced view on the recent criticism of failure to adopt the WE perspective so as not to "realistically represent the variation of pronunciation norms in international communication" (p. 50). She characterizes taking WE or ELF perspectives on pronunciation as a daunting task, though great progress has been made in test construction and design. A non-exclusive, pluricentric orientation is endorsed with special emphasis on phonological intelligibility, though more research is needed to examine the speaker, the interlocutor, the context and their interaction.

In Chapter 4, Yan and Ginther review research on listener background characteristics that influence evaluation of L2 accented speech, including familiarity with the accent, familiarity with the speech topic, attitude towards the accent, and language proficiency of the listener and investigate how these characteristics factor into listeners' and raters' evaluation of L2 speech. They suggest that future research should focus on investigating the impact of rater interaction with L2 accents.

Chapter 5 deals with assessing pronunciation for research purposes with listener-based numerical scales. Isbell firstly presents a brief review of numerical scales with a special focus on the appropriateness of particular measurement techniques for valid interpretation of L2 pronunciation scores, including scale function, rater variation, and difficulties associated with the rating task. The author uses an illustrative example of L2 Korean pronunciation to explore scale function of 9-point scales to judge comprehensibility and accentedness by adopting many-facet Rasch analysis and debriefing questions. The result shows better function of comprehensibility scale than the accentedness scale and discrepancy in listeners' judgments and use of the scales.

The second part (Chapters 6-8) is on the role of technology in pronunciation assessment. Chapter 6 deals with pronunciation features in rating criteria. Ghanem and Kang point out that recent studies focus more on the importance of suprasegmental features in contributing to a listener's perception of a

speaker's intelligibility or comprehensibility. The authors describe segmental features and suprasegmental features before illustrating selected pronunciation features and their measurement with screen shots from PRAAT to illustrate how the measurement is attained. Finally, the authors suggest incorporating certain features that have been proven significant in L2 speaking and assessment research into the band descriptors, selecting task-specific constructs, employing automated speech recognition (ASR) system and ensuring consistent use and definition of pronunciation constructs.

In Chapter 7, Van Moere and Suzuki firstly describe the process for developing an automated system for assessing pronunciation, including four components, namely, the acoustic model, the language model, the pronunciation dictionary and the scoring model. For the acoustic model, the authors have an in-depth discussion about the reference speakers with different intended purpose of pronunciation assessment. Other considerations in the use of ASR include the system being speaker dependent/independent and speech being constrained/unconstrained. Recent developments in pronunciation scoring include spectral match scores, fundamental frequency and energy, and duration features. With constrained tasks like read aloud, the authors illustrate how features of speech are extracted and weighted to derive scores for pronunciation sub-constructs.

Loukina et al. cover automated pronunciation assessment in spontaneous speech in Chapter 8. The authors firstly give an account of early attempts to assess pronunciation automatically, which generally referred to highly-constrained tasks, and then introduce two major groups of measurements applicable to either constrained or unconstrained tasks. One group is model-based approach, which compares learner pronunciation to the existing reference model for the same segment, word or sentence to detect similarities or discrepancies. The other group is generic approach, which aims to find out general patterns of pronunciation discriminating between different levels of proficiency. As the authors point out, most systems now combine these two models for automated pronunciation assessment. Loukina et al. also bring forward challenges in automatic assessment of pronunciation in unconstrained speech in technical and conceptual terms. One limitation is errors in automatic transcription of ARS. Another challenge is to select appropriate reference models in international communication. The authors also explore validity issues relating to automated assessment of pronunciation in the final section by referring to Xi (2010, 2012). The chapter ends with suggestions for future development: "Big Data" may boost ASR accuracy, providing promising future for automated assessment of unconstrained speech. Crowdsourcing also promises to improve automated assessment of pronunciation. More concerted efforts by technologists, phoneticians, and second language acquisition researchers are expected.

This edited volume is laudable in the following aspects: Firstly, it provides readers with authoritative and updated review of important concepts like accentedness, intelligibility, and comprehensibility. Clear definition would dispel confusion, and is conducive to the healthy development of research on pronunciation assessment. Secondly, a discussion about the unique validity issues in pronunciation assessment is timely. Thirdly, putting pronunciation assessment in the perspective of international

communication is highly relevant. Fourthly, having a whole part focusing on technology and pronunciation assessment quenches avid readers' thirst for issues in automated scoring of pronunciation, especially as applied in unconstrained tasks.

This volume does share some common focuses with Isaacs and Trofimovich (2016), such as the discussion about constructs in pronunciation, the native speaker norm and ELF, factors influencing rater/listener judgments of pronunciation, etc. But it is distinctive from the latter in that while the latter includes technology and automated scoring in future direction, this volume has detailed discussion about this issue. The discussion about validity in pronunciation assessment is invaluable too. Other gaps that have been filled are such as the inclusion of listeners' language proficiency in rating pronunciation. With 100 pages thinner than the latter, this volume is more accessible to wider audience. Language teachers would find this book helpful in developing their assessment literacy and knowledge about criterial features of pronunciation that should be of concern to them in classroom instruction. Language testers would benefit from the book about clarified definition of important constructs in pronunciation assessment and speech technology in automated scoring of pronunciation. Intending researchers would find factors affecting pronunciation rating extremely interesting to them. Despite all the merits, it is argued that the volume would be a perfect reference if there were a conclusion chapter in the end. It is true that in every chapter, the authors provide new directions and recommendations for each topic, but a conclusion chapter for the whole volume would bring the issues and future on pronunciation assessment into better perspective.

Funding

References

Isaacs, T., & Trofimovich, P. (2016). *Second Language Pronunciation Assessment: Interdisciplinary Perspectives*. Bristol, UK: Multilingual Matters. https://doi.org/10.21832/ISAACS6848

Xi, X. (2010). How do we go about investigating test fairness? *Language Testing*, *27*, 147-170. https://doi.org/10.1177/0265532209349465

Xi, X. (2012). Validity and the automated scoring of performance tests. In G. Fulcher, & F. Davidson (Eds.), *The Routledge Handbook of Language Testing* (pp. 438-451). New York: Routledge.