

FACETS 分析による英語音声評価における採点者バイアス検証
Investigating Rater Bias in Measuring English Pronunciation: A FACETS Analysis

吉田弘子 (大阪女学院大学・大阪女学院短期大学)

井上球美子(大阪女学院大学・大阪女学院短期大学)

上田洋子 (大阪女学院大学・大阪女学院短期大学)

大塚朝美 (大阪女学院大学・大阪女学院短期大学)

第二言語学習者にとって、対象言語の適切な発音を習得することが円滑なコミュニケーションを進める上で重要であることは、コミュニケーション能力 (communicative competence) の構成要素として発音の役割が認識されていることから明らかである (Bachman, 1990; Bachman & Palmer, 1996; Canale, 1983; Canale & Swain, 1980; Celce-Murcia, Dörnyei, & Thurrell, 1995)。しかし日本の英語教育において発音指導はこれまで積極的に取り入れられることはなかった。その理由のひとつとして、第二言語英語話者である教師が英語音声を客観的に指導評価する手法が確立しておらず、音声評価を筆記試験による発音知識の評価で代用してきたことが考えられる。Yoshida (2004) の研究では、高い英語力を持つ日本語を母語とする英語教師は、英語を母語とする英語教師と同等に英語の発音評価を行うことができることが示された。しかしこの研究では、採点者は経験豊かな大学英語教員ではあったが、いわゆる発音を教える専門家 (expert) ではなかった。本研究は、Yoshida (2004) の研究を発展させ、phonetics を担当する教員の採点者バイアス (rater bias) を、パフォーマンス評価に適した項目応答理論を応用した統計的手法 (FACETS 分析) を用いて検証した。英語発音評価における採点者のバイアスを分析することにより、今後、日本における信頼性の高い英語発音の評価の普及に寄与することを期待する。

目的

本研究の目的は、日本語を母語とする英語学習者の発音評価する採点者 (phonetics 担当者) の評価の信頼性と採点者バイアス (rater biases) を、採点項目及び被験者の能力別にパフォーマンス評価に適した項目応答理論を応用した統計的手法 (FACETS 分析) を用いて検証することである。

方法

(1) 被験者

被験者は2003年度に入学し、phonetics のコーススタート時 (2003年4月) と修了時 (2004年1月) に *Phonetics in context* (大塚, 今井, 井上, 上田, 米田, 2004, p. 2) の同一英会話文 “Diagnosis” をオーディオカセットに録音した短大1年生262名から無作為に抽出した学生38名である。

(2) 評価項目と評定尺度

アメリカ標準英語 (General American English: GAE) を基にして評価基準とし、GAE の重要な構成要素である音素 (segmentals) 超音素 (suprasegmentals) 準言語要素 (paralinguistic features) に属する15項目について6段階の尺度 (1: Very poor, 2: Poor, 3: Fair, 4: Good, 5: Very good, 6: Excellent) を用いて採点した。

9月9日(金) 研究発表2 第2室(249)

音素	超音素	準言語要素
母音 (Vowels)	単語ストレス (Word stress)	声の大きさ (Loudness)
二重母音 (Diphthongs)	文ストレス (Sentence stress)	速さ / テンポ (Rate / Tempo)
子音 (Consonants)	リズム (Rhythm)	スムーズ度 (Smoothness)
子音連結 (Consonant Clusters)	イントネーション (Intonation)	エネルギーレベル (Energy)
アスピレーション (Aspiration)	弱形 (Weak forms)	明瞭度 (Clarity)

(3) 採点者と評価手順

採点者は、大学と短期大学で phonetics を担当する 5 人の日本語母語話者である。これらの採点者は phonetics の授業でテープ録音した学生の発音を定期的に評価しているが、今回の研究にあたり、統一性を保つために採点者ガイダンスを行い、採点項目の擦り合わせ並びにデモテープと評定尺度を用いて採点練習を行った。なお、採点は 2004 年 8 月に各自個別に実施した。

分析

項目別、被験者能力別に採点者の信頼性とバイアスを FACETS ソフトウェア (Linacre, 1996) を用いて検証した。採点者バイアスは、評価のスコアに影響を与える複数の要因に見られる特別な相互関係を分析するものである。

結果

FACETS 分析の Infit 数値は、5 人の採点者はすべて一貫して採点者内統一性があることを示した。また、採点者間においても 15 各項目で一定以上の信頼性 (.66 - .89) が認められた。しかしながら、採点者の採点の厳しさは同等ではなかった (Reliability index = 0.99, chi-square = 977.9, $df = 4$, $p = .00$)。さらに各採点者は、項目及び被験者能力に関して独自の採点バイアス (厳しさ/ 寛容さ) を持っていることを示した。これらの結果の詳細および考察については発表時に報告する。

引用文献

- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 2-27). New York: Longman.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 1-47.
- Celce-Murcia, M., Dörnyei, Z., & Thurrell, S. (1995). Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics*, 6, 5-35.
- Linacre, J. M. (1996). FACETS (Version No.3.02). Chicago: MESA.
- Yoshida, H. (2004). *An analytic instrument for assessing EFL pronunciation*. Doctoral Dissertation. Ann Arbor, MI.: UMI.
- 大塚朝美, 今井由美子, 井上球美子, 上田洋子, 米田信子 (2004). *Phonetics in context*. 東京, 北星堂.