

機関番号：15201

研究種目：若手研究（B）

研究期間：2008～2010

課題番号：20730555

研究課題名（和文）英語口頭運用能力を測定するための共通参照尺度の開発

研究課題名（英文）Developing a Common Reference Scale of Oral English Proficiency

研究代表者

猫田 英伸（NEKODA HIDENOBU）

島根大学・教育学部・講師

研究者番号：80452598

研究成果の概要（和文）：本研究全体の最終的な目的は、学校教育現場で教員が生徒のスピーキング技能の熟達度を簡便に評価する際に使用可能な評価尺度を開発することである。本研究期間においては、上記目標の達成に向けて以下の成果を挙げることができた。

- (1) 評価尺度に含めるルーブリック（試案）の文言の妥当性について検証し、必要な修正を行った。
- (2) 新規にルーブリックの開発を行い、評価尺度に追加した。
- (3) 項目応答理論と潜在ランク理論の併用の可能性について検討することができた。

研究成果の概要（英文）：The final aim of this study is to develop common reference scales of oral English proficiency, which can be easily used among English teachers in school. During the three-year granted research period, the following short-term goals have been achieved:

- (1) The rubrics included in the tentative scales have been examined and revised;
- (2) New rubrics have been developed and added to the scales;
- (3) The possibility of combining Item Response Theory and Latent Rank Theory has been discussed.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008年度	300,000	90,000	390,000
2009年度	500,000	150,000	650,000
2010年度	700,000	210,000	910,000
年度			
年度			
総計	1,500,000	450,000	1,950,000

研究分野：社会科学

科研費の分科・細目：教育学・教科教育学

キーワード：英語教育、パフォーマンス評価、主観的評価、評価尺度、ルーブリック

1. 研究開始当初の背景

近年、日本の英語教育においてはコミュニケーション能力の中でも特に英語を「聞く」「話す」能力の育成が重視されるようになってきている。そしてこれに伴い、「聞く」「話

す」能力を適切に評価するための方法の開発が急がれている。このことは、平成15年3月に文部科学省が示した『「英語が使える日本人」の育成のための行動計画』の中で、(1)中学校では「聞くこと」「話すこと」を重視した指導

を行い、これらの能力を目標準拠評価（いわゆる絶対評価）に基づいて評価すること、および(2)大学、高校入学試験において「聞くこと」「話すこと」を含むコミュニケーション能力を適切に評価することの必要性が明記されたことから見て取れる。

そんな中、「聞くこと」「話すこと」の能力の評価については外部試験を積極的に利用することがしばしば提案されてきた。しかしながら、日々の教育実践における指導、学習、評価のサイクルの中で外部試験を形成的評価あるいはミニ総括評価（松沢，2002）の方法として利用することは不可能である。このような現状認識のもと、本研究は日本の学校教育現場において英語教員が学習者に「聞くこと」「話すこと」を実際に指導する中で、学習者の英語口頭運用能力を形成的に評価する際に利用することのできる共通参照尺度の開発を目指す。そして最終的には、**Bachman & Savignon (1986, p. 388)** が提案しているような「言語使用場面や内容領域を超えて、あらゆるスピーチ・サンプルを評価するための共通基準」の開発を果たすことを意図する。このような尺度が依拠する具体的な理論的基盤についてはここでは割愛するが、当然のことながらここで想定されている共通基準とは素点をそのまま得点とするような単純な尺度ではなく、評定者の特性 (*rater leniency*)、被評定者の特性 (*performance quality*)、評価項目の特性 (*item difficulty*) の相互作用を数理的に処理することで一般化された確率論的得点 (*probabilistic score*) を算出するような尺度である。もしも、このような共通参照基準が開発され、英語教員が日々の教育実践の中でインターネットを通じてこの尺度を容易に活用できるようなシステムを構築することができたならば、(1)中学校、高等学校、大学といった教育段階、(2)評定者間の評価の厳しさの

差、(3)言語タスク間の認知的負荷の差などを越えて、学習者の産出したあらゆるスピーチ・サンプルの質についての得点を一般化された値で得ることが可能となる。

2. 研究の目的

このたび科学研究費の交付を受けた期間において扱った具体的な研究課題は以下の2つの点である。

一点目はルーブリックの精緻化である。すでにこれまでの研究において、世界中の既存の尺度に含まれるルーブリックから使いやすいものを選別するようターゲット・ユーザである日本人英語教員に依頼し、トップ・ダウン的にルーブリックを開発している。そこで今度は逆に、日本人英語教員集団が持つ職業的日常生活言語を分析することでボトム・アップ的にルーブリックの妥当性を検証する。

第二点目は、第一点目で述べた検証結果に基づいてルーブリックの修正を行うとともに、必要に応じてルーブリックの新規作成を行うことである。その際、日本人英語教員が評価に際して実際に用いる生の「ことば」を評価項目であるルーブリックとしてまとめることを目指した。

3. 研究の方法

上記2つの目的を達成するため、具体的には以下の方法を用いた。次図は本研究で扱ったデータの全体像を示している。ここでは便宜上、図中の右側から順に本研究の第一、第二、第三の柱と呼ぶこととする。第一、第二の柱では「No. 6 言語行動」(**Council of Europe, 2004, p. 221**:「能力記述文と典型的な言語行動とがレベル帯の中で対置され、記述されている内容と実際に行われたこととの一致を確認する。」)を用いる。具体的には、すでに**Nekoda et al. (2007)**によって試作された評価尺度内のルーブリックに示されてい

る各種の言語行動の特徴について、(項目応答理論による分析結果に鑑みて) 対応するレベルのパフォーマンスのビデオ映像を3名の現職英語教員が描写した際に言及があるか否かに着目して分析する。第一の柱の部分に対応する研究では、既存のルーブリック群の中で英語教員の評価の観点、評価言語から乖離していると思われるものを特定し、削除、修正の見込みを検討する(猫田, 2010)。続いて第二の柱の部分にあたる研究では、英語教員の評価の観点や評価言語をほぼ反映していると思われるルーブリック群に関して、意味の変わらない程度に(英語教員にとって)より自然な言い回し、表現になるよう一つひとつ文言の修正を行う。これが本研究の目的の第一点目に対応する方法である

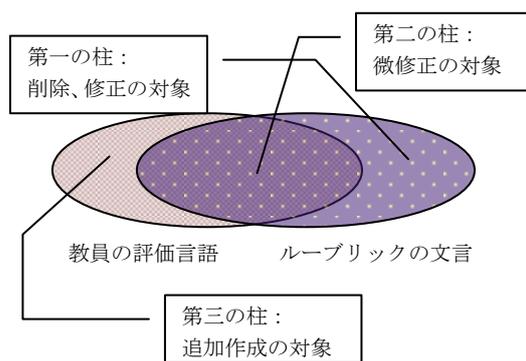


図 本研究の全体像

第三の柱は本研究の目的の第二点目に対応する研究である。方法としてはCouncil of Europe (2004)が示すところの「No. 9 比較判定法」を用いる。具体的には言語行動(パフォーマンス映像)を2つずつセットで準備し、それらを(先ほどと同じ)3名の現職の英語教員が視聴、比較しながら「良い点」「悪い点」について顕著な特徴を記述するというものである。これによって、英語教員が生徒たちの英語能力を評価する際に日常的に用いている

評価言語を引き出し、得られた教員の生の声を反映する形でルーブリックの追加作成を行うことを試みる。また、図より明らかなように、本論文は基本的に第一の柱、第二の柱の研究と同じデータの異なる部分に着目して分析することになるため、研究同士の横の相互補完も同時に果たすことになる。

4. 研究成果

○第一の研究目的: ルーブリックの精緻化

Nekoda et al. (2007)によって開発された評価尺度の中には現職の英語教員の評価観がある程度うまく取り入れられていたことが明らかになった。全44個のルーブリックのうち、特に問題のあったルーブリックとしては、(1)3名の教員の中で誰もまったく言及しなかった要素のみで構成されているルーブリックが1つ、(2)3名の教員の中で1名のみが一部の要素に言及したルーブリックが2つ、(3)3名の教員のうち2名が一部の要素に言及したルーブリックが1つ、以上4つが特定された。全体としてはすべてのルーブリックに含まれるパフォーマンスの特徴について3名の教員が広く共通して言及していることが明らかになった。その一方、3名の教員から得られた具体的な評価言とルーブリックの文言とをつき合わせる作業を行う中で、評価尺度内のルーブリックの表現には改良の余地があることも分かった。

○第二の研究目的: ルーブリックの新規作成

本研究では3名という非常に限られた数の現職英語教員から得られたデータをもとに分析を行った。そのため、収集されたデータ自体の信頼性について疑念が残る。しかしながら、3名から得た評価言語であっても猫田(2007), Nekoda et al. (2007)において作成された評価尺度内の言語特徴記述をかな

りの程度カバーしており、追加作成されたルーブリックの数は 16 個にとどまった。本論文において扱ったデータそのものは非常に限られているものの、このような結果が具体的なデータに基づいて明らかにされたことの意義は決して小さくない。むしろ、①任意に取り出された 3 名の英語教員の日常的な評価言語でさえ本研究が開発中の評価尺度と一貫して整合性を持つことが示され、②さらに両者の間のズレを明示的に特定するとともに、③新規にルーブリックを作成、追加することでそのズレを埋めることができたことは、本研究の当初の目的が達成されたことを意味していると言える。

○研究目的以外の成果

本研究を遂行する中で、新しいテスト理論である潜在ランク理論について講演を聞く機会があった。今後、本研究が評価尺度の開発を行っていくうえで、この潜在ランク理論を無視することはできないと感じるに至った。本研究課題を終えるにあたって、今後の研究の新たな視点を得ることができたことは極めて大きな成果である。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 3 件)

①猫田英伸, 英語教員の評価言語に基づく記述子の開発, 『教育学研究ジャーナル』7 巻, 11-20, 2010, 査読有り

②猫田英伸, 英語高等運用技能の熟達度に関する評価尺度の質的妥当化, 『日本語テスト学会研究紀要』, 13 巻, 55-70, 2010, 査読有り

③Hidenobu NEKODA, An investigation on the rating tendencies of Japanese English teachers, *International Journal of Curriculum Development and Practice*, 11, 13-23, 2009, 査読有り

[学会発表] (計 1 件)

①猫田英伸, 学習者の熟達度と教師の質的判断の傾向: スピーキング・パフォーマンス評価で一たから, 第 34 回全国英語教育学会, 2008 年 8 月 9 日, 昭和女子大学 (東京都)

6. 研究組織

(1) 研究代表者

猫田 英伸 (NEKODA HIDENOBU)

島根大学・教育学部・講師

研究者番号: 80452598

(2) 研究分担者

()

研究者番号:

(3) 連携研究者

()

研究者番号: