

博士學位論文

内容の要旨及び審査結果の要旨

令和2年度

京都外国語大学

はしがき

これは学位規程（平成 25 年文部科学省令第 5 号）第 8 条による公表を目的として、令和 3 年 3 月 15 日に本学において博士の学位を授与した者の論文内容の要旨及び論文審査の結果の要旨を収録したものである。

氏名	笠巻 知子
学位の種類	博士（言語文化学）
学位記番号	甲第24号
学位授与の日付	令和3年3月15日
学位授与の条件	本学学位規程第3条3号該当
学位論文題目	学生による相互評価力とプレゼンテーション力 に及ぼす要因に関する実証的研究 — 学生の相互評価と教員による評価との相関と 評価者トレーニングに基づいて —
論文審査委員	主査 専任講師 鈴木 寿一 副査 教授 相川 真佐夫 副査 教授 平井 明代（筑波大学）

論文内容の要旨

本博士学位請求論文は、学位請求者である笠巻氏が非常勤講師として勤務する大学で行われていた「プロジェクト発信型英語プログラム」の授業内で行われた一連の研究をまとめたものである。「プロジェクト発信型英語プログラム」は、学生各自が取り組むプロジェクト活動を通して、リサーチ、ディスカッション、プレゼンテーションのスキルを学び、また、自分自身のプロジェクトの成果を英語で発信する力を身につけることを目標とした授業である。

一般に、学生が行うプレゼンテーションの評価者は教員であることが多いが、これまでの研究で明らかにされてきた「学生による相互評価」の利点を活かして、「学生による相互評価」を成績評価に組み入れる前の段階として、本研究は、教員による評価と学生による相互評価の相関を調べるという手法を用いて、以下の3点を検証することを目的に行われた。

①学生がお互いの発表を評価するために必要な、その内容を理解できるだけの英語力、その内容に対する予備知識の有無、プレゼンテーション力、そして評価者トレーニングの4つの要因が評価者としての学生の評価力にどの程度の影響を及ぼすのかを調べ、学生による相互評価の信頼性を検証すること。

②プレゼンテーションの指導の一環として、評価項目別に評価者トレーニングを行うことは学生の評価力に影響を及ぼすか。

③評価者トレーニングは、学生の評価力だけではなく、学生のプレゼンテーション力そのものを伸ばすか？

本研究は以下の9つの章から成る実証研究である。

第1章 はじめに

第2章 相互評価に関する先行研究

第3章 研究1：学生の「プレゼンテーション力」が評価者としての学生の評価力に影響を及ぼすか？

第4章 研究2：学生の「英語力」が評価者としての学生の評価力に影響を及ぼすか？

第5章 研究3：学生の「予備知識」が評価者としての学生の評価力に影響を及ぼすか？

第6章 研究4：「評価者トレーニング」が評価者としての学生の評価力に影響を及ぼすか？

第7章 研究5：「評価項目別評価者トレーニング」が評価者としての学生の評価力に影響を及ぼすか？

第8章 研究6：「評価者トレーニング」が学生の「プレゼンテーション力」に影響を及ぼすか？

第9章 おわりに — 研究結果のまとめと今後の課題 —

以下、各章の内容について概要を述べる。

第1章では、教員による学生のスピーキング力の評価を補うための方法として学生による相互評価があるが、その信頼性を検証する必要があることを指摘している。その上で、学生の評価力に影響を及ぼすと思われる4つの要因（プレゼンテーションの内容を十分に理解できるだけの英語力とその内容に対する予備知識の有無、プレゼンテーション力、そして評価者トレーニング）を取り上げ、その4つの要因が学生によるプレゼンテーションに対する学生の相互評価にどのような影響を及ぼすか調べるために、第3章から第8章の各章で検討される6つのリサーチ・クエスチョンが挙げられている。

第2章では、スピーキングの評価の観点、評価方法、評価に影響を及ぼす要因についての先行研究に加えて、「学習者に学習目標を明確にさせる」、「学習者の学習に対する動機づけが高まる」、「相互評価を通して学生同士が互いに学び合うことができる」といった相互評価の利点と、教員による評価よりも若干甘くなる傾向がみられ、極端な評価を避ける傾向があるという問題点を指摘した先行研究がまとめられている。また、評価の信頼性に関する先行研究として、学生による評価は信頼でき、成績の一部に組み入れることが可能であるとする先行研究と、教員による評価との一致度が低く、信頼性に欠けるとする先行研究もまとめられている。

第3章では、評価者としての学生の評価力に影響を与えるとされる4つの要因のうち、学生の「プレゼンテーション力」が学生の評価者としての評価力に影響を及ぼすかを調べてい

る。学生が行ったプレゼンテーションに対する学生による相互評価と教員による相互評価との相関を成績群別に調べたところ、①「評価の合計点」では、すべての成績群において中程度の相関 ($r = .529 \sim .609$) が得られ、どの成績群の評価力も高くないこと、②上・中・下の3つの成績群の中で、プレゼンテーション力が一番高いとされる上位群の学生による相互評価と教員による評価との相関が最も低いことが確認されたことから、学生のプレゼンテーション力は評価者としての学生の評価力には影響を及ぼさない可能性が高いとしている。また、全体としては、教員の評価と比べると学生は甘めの評価をすること、成績群別にみると、上位群の学生は低めの評価をし、下位群の学生は高めの評価をする傾向があると報告している。

第4章では、学生の「英語力」が評価者としての学生の評価力に影響を及ぼすかを調べている。その結果、①評価の合計点では、全成績群において中程度の相関 ($r = .551 \sim .589$) が得られ、どの成績群の評価力も高くないこと、②上・中・下の3つの成績群の中で、英語力が高い上位群ではなく、中位群の評価力が一番高いことから、学生の英語力は評価者としての学生の評価力には影響を及ぼさない可能性が高いとしている。また、学生の評価傾向としては、教員による評価と比べると甘目の評価を行う傾向があることを報告し、その要因として、クラスメートに対して評価することに気まずさを感じることから、高めの点数をつけてしまうといった心理的要因や、他者との関係性で評価を変えてしまうといった評価者の性格が、評価に影響を及ぼしているのではないかと述べられている。

第5章では、学生の「予備知識」が評価者としての学生の評価力に影響を及ぼすかを調べている。学生の予備知識の程度を2から10の5段階に分け、予備知識の程度別に学生一人一人の発表に対する教員による評価点と学生による相互評価点の差を調べた結果、①学生の予備知識の多少は、学生の評価傾向に少し影響を及ぼす可能性があること、②学生の評価傾向として、教員による評価と比べると、学生による評価は甘目になること、その要因として、学生の評価経験の浅さのために、評価の際に自分自身のプレゼンテーションの出来と比較してしまうこと、さらに、クラスメートに対して評価することへの気まずさといった心理的なものがあるとしている。学生による相互評価を活用するためには、学生に評価力をつけ、評価の信頼性を高めることが必要であり、そのためには、判定基準および評価の観点が明確に示されたルーブリックを用いて評価者トレーニングを行う必要があると主張している。

第6章では、学生を、45分間の評価者トレーニングを行ったグループと、約15分間のトレーニングを行ったグループに分けて、評価者トレーニングが評価者としての学生の評価力に影響を及ぼすかを調べている。教員による相互評価と学生による相互評価の相関係数を算出した結果、評価の合計点では、両グループともに、弱い相関が得られ ($r = .383$)、トレーニングが1度だけでは学生の評価力に影響を及ぼすとは言えないことがわかった。続いて、トレーニングの回数が学生の評価力に影響を及ぼすかを調べた結果、2回目のトレーニング後では、評価の合計点では両グループとも中程度の相関 ($r = .619 \sim .625$) が得られ、評価力が向上したことが確認された。トレーニング回数では両グループの間に差はな

かったことから、トレーニングの内容や質よりも実施回数の方が学生の評価力に影響を及ぼし、相互評価の信頼性が高くなる可能性があることも指摘している。

第7章では、指導する評価項目別に評価者トレーニングを行うことで、評価者としての学生の評価力に影響を及ぼすかを調べるために、相互評価を行う前の3週間をかけて授業時に「姿勢」・「アイコンタクト」・「ジェスチャー」・「声の大きさ」・「明瞭さ」・「スライド」の6項目について項目別評価者トレーニング（1回30分、合計3時間）を行う詳細グループと、相互評価を行う前週の授業時に6項目すべての評価項目を1回のみでの評価者トレーニング（45分間）を行う簡易グループに受講者を分けてトレーニングを行った。その結果、評価の合計点において、両グループともに中程度の相関（ $r = .698$ 及び $.699$ ）が得られ、学生の総合的な評価力は、両グループとも比較的高いことがわかったが、評価項目別にみると、中程度の相関から相関なしまで、さまざまであることがわかった。学生の評価力を高めるにはさらなるトレーニング方法の改善や回数を増やす必要があることなどが今後の課題として残った。

第8章では、評価者トレーニングによって、学生の「プレゼンテーション力」（学生のプレゼンテーションの成績）が向上するかを調べている。「トレーニングなしグループ」（トレーニング開始前の2015年度の学生のグループ）と「トレーニングありグループ」（トレーニングを開始した2016年度の学生のグループ）の各グループの中間発表と最終発表の成績を用いて分析が行われた。その結果、「トレーニングなしグループ」では、中間発表から最終発表の成績の間には変化は見られなかったが、「トレーニングありグループ」では、最終発表の成績が中間発表の成績を若干ではあるが上回っていることから、評価者トレーニングは学生のプレゼンテーション力を向上させる可能性があるとして述べている。さらに、詳細なトレーニングを受けた学生と簡易なトレーニングを受けた学生のプレゼンテーションの成績を比較した結果、簡易的なトレーニングを行う方が学生のプレゼンテーション力に影響を及ぼす可能性があることがわかった。以上の結果から、評価者トレーニングが学生のプレゼンテーション力に影響を及ぼすためには、評価項目およびトレーニングの方法、実施回数、および実施する時期を見直し、手順を簡易にしたトレーニングを継続的に実施し、さらに発話トレーニングも併せることで、学生のプレゼンテーションの指導の一つとして活用できる可能性を指摘している。

第9章では、本研究であきらかになったこと（各章の概要ですでに述べたので省略）と、本研究から示唆されること、本研究の問題点と今後の課題がまとめられている。まず、本研究から示唆されることとして、評価者トレーニングの意義は、①教員による評価との一致度を向上させる、言い換えれば、学生による相互評価力を向上させる可能性があること、②トレーニングによって、クラスメートの良い発表からも、そうでない発表からも学ぶことができるようになり、学生の批判的思考が高まるとともに、将来学生が自律した学習者になるために必要な判断力を身につけることにつながることで、③適切なトレーニングを行えば、学生自身のプレゼンテーション力の伸ばせる可能性があること、④学生にとって評価が難しい

評価項目を知ることは、教員自身による評価を見直すことにもなり、教員自身の評価力を向上させる可能性が指摘されている。

次に、本研究の問題点として、①「プロジェクト発信型英語プログラム」は毎回の指導内容・進度・指導法・評価法など、全クラス統一シラバスで行われた授業であったため、毎時間の指導法や進度の変更ができなかったことと、受講学生は半年間で入れ替えがあったため長期的な指導ができなかったこと、各学期の最終プレゼンテーションでは、本研究の参加者の学生が他のクラスの学生のプレゼンテーションを評価することになっていたため本研究で必要なデータが得られなかったこと、全クラスで統一されていた相互評価項目が十分に適切なものではなかったこと、学生の英語力は TOEIC IP だけで測定されていたことなど、本研究の研究結果に影響を及ぼす要因をコントロールできなかったこと、②各学生の TOEIC IP のスコアを非常勤講師は知ることができなかったため、学生の自己申告に頼るしかなかったこと、③評価者トレーニング方法に改善の余地があること、④学生が互いのプレゼンテーションを評価するためのルーブリックに改良の余地があること、⑤本研究では同じ科目を担当する教員の協力が得られなかったため、教員自身の評価者内信頼性のみで、評価者間信頼性を確保できなかったことなどが挙げられている。

今後の課題として、①評価者としての学生の評価力に影響を及ぼす要因を明確に突き止めるには至らなかったことから、その要因を探るべく、引き続き研究を続ける必要があること、②問題点で挙げられた研究結果に影響を及ぼす要因をコントロールして研究を継続すること、③本研究で明らかになった評価者トレーニングを見直し、手順を簡易にしたトレーニングを継続的に実施することで、学生の評価力を高められるかどうかを検証する必要があること、④評価者トレーニングが、学生のプレゼンテーションの指導の一つとして活用できるようにするために、発話トレーニングも併せて行うことで、学生のプレゼンテーション力を向上させることができるかどうかを検証することなどが挙げられている。

口述試問及び審査結果

口述試問では 3 名の審査委員と発表者の間で質疑応答が行われた。審査委員からは以下のような質問が出された。

- ①担当授業科目が全クラス統一シラバスで行われた授業と言う制約があったため、コントロールできなかった本研究の研究結果に影響を及ぼす要因が挙げられているが、その制約のためにコントロールできなかったもう一つの重要な要因が論文には書かれていないが、それは何か？
- ②今回コントロールできなかった点をコントロールできる実践の場を持つことができるようになった場合に、評価者トレーニング前と学生の評価力とその後の評価トレーニングを継続して行く過程で学生の評価力が向上するようにするにはどのような評価者トレーニングを行えば良いか？

- ③評価者トレーニングが評価者としての学生の評価能力を向上させるだけでなく、自身のプレゼンテーション力を向上させるためには、具体的にどのような指導を行えば良いか？
- ④評価者トレーニングだけでなく、英語の発話トレーニングを行う必要があると論文に書かれていたが、そのトレーニングとは具体的にはどんなものか？
- ⑤プレゼンテーションを聴いて相互評価を行うには、英語による説明を聴いて理解できるだけのリスニング力が必要で、そのリスニング力を測定するのに TOEIC IP が用いられているが、TOEIC はこのようなプレゼンテーションを聴いて理解する英語力を測定するテストとして適切なテストと考えられるか？
- ⑥「プロジェクト発信型英語プログラム」は英語力を伸ばす事が第一の目標ではないが、プレゼンテーションの内容を理解するには、別の授業で英語力を伸ばす指導が行われる必要がある。その授業ではどんな指導が行われていたのか？また、そこで行われている指導が適切でない場合は、どんな指導を行う必要があるか？
- ⑦今回の研究は単独で行われたため、評価者である教員の評価力の信頼性を担保するための評価者間信頼性を調べることができず、次善の策として評価者内信頼性を調べているが、本研究における教員の評価力をさらに担保するには、教員自身が評価者トレーニングを受けた経験があることが望ましい。評価する教員である笠巻さん自身は評価者トレーニングを受けた経験はあるか。
- ⑧p. 14 の正規性の検定結果（表 2）の数値と p. 34 の正規性の検定結果（表 1）の数値が同じだが、写し間違いではないか？
- ⑨この研究では、学生の英語力を測定するために TOEIC のリスニングとリーディングのスコアが用いられているが、この授業で必要な英語力として発信能力は必要ないのか？
- ⑩評価者トレーニングの効果に関する研究では、「詳細トレーニング・グループ」、「簡易トレーニング・グループ」、「トレーニングなしグループ」の 3 群で比較する方が良かったのではないか？
- ⑪学生の相互評価力に影響を及ぼすのはこの研究で取り上げられた 4 つの要因以外に、学習者が持っている belief もあるのかではないか？また、発音の項目のループリックでは英語らしく聞こえるかどうかだけでなく、聴き手側から見たわかりやすさ (intelligibility) も発音の評価基準にいれるべきではないか？
- ⑫研究 1 のプレゼン力を偏差値化して上中下の 3 群に分けているが、その元になっているデータは何か？
- ⑬研究 2 のオリジナリティは、p. 71 のループリックのオリジナリティに基づいていないようだが、研究 2 のオリジナリティはどんな基準で学生は評価したのか？
- ⑭p. 71 のループリックは独自に作成したものか？また、オリジナリティの基準として書かれている文言が「評価」の一部となり得るか？オリジナリティの中で「『内容がとて

もよくわかった』、『ほとんど初めて聞いた内容だった』という基準は、評価者の学生の知識基準ではないか？ルーブリックの基準の文言には主語は省略されているが、『内容』と『オリジナリティ』の基準の主語は評価者であるのに対して、『アイコンタクト』や『スライド』などの主語は発表者であると考えられるが、どうか？

⑮『内容』と『オリジナリティ』の2つは鍛えるべき評価力になり得るのか？

⑯p. 15に「相関の程度が中程度 ($r = .605$) からマイナスの関係 ($r = -.193$) までと3段階におよび、」と書かれているが、3段階とはどういう意味なのか？

以上の質問に対する応答は、一部に不十分な点があったが、概ね適切であった。

次に、次の8つが改善すべき点として審査委員から指摘された。

- ①この研究で用いられている「評価力」の定義が書かれていない。
- ②「プロジェクト発信型プログラム」に絞ってこの研究が行われた理由を明記しておく必要がある。
- ③学生の相互評価力に及ぼす要因として、英語力、予備知識、プレゼンテーション力、評価者トレーニングの4つの要因が挙げられているが、その4つを取り上げた理論的背景を述べる必要がある。
- ④上中下のグループ分けをした際の偏差値のレンジだけでなく、実際の成績のレンジも記載すべきである。
- ⑤図表がたくさんあるので、すべての図表に適切なタイトルをつける必要がある。
- ⑥全体として統計処理結果の記述を修正すべきである。統計処理結果（特に検定の結果で、p値が.05を越えている場合）の記述が適切でないので修正する必要がある。例えば、マンホイットニーのU検定結果のp値については何も記されていないので、p値に基づいて「有意差はなかったが、効果量は〇〇であった」のように記述すべきである。
- ⑦参照された先行研究や統計処理結果からは断定できないことを断定してしまっている箇所がある。
- ⑧研究1～3ではルーブリックなしで相互評価を行っていたことと、研究4～6はルーブリックを用いて学生が相互評価を行っていたことを明記しておく必要がある。

また、今後の課題として、審査委員がさらなる研究を笠巻氏に期待するのは次の3点である。

- ①評価者トレーニングを2回行った結果、1回目の相関より2回目の相関が高くなっているが、評価者トレーニングを増やしていくことで学生の相互評価力が徐々に向上していくことが予想される。評価者トレーニングの方法とともに回数を増やすことで、学生の相互評価力がどこまで伸びるかを検証してもらいたい。
- ②評価者トレーニングを行うことで学生のプレゼンテーション力そのものが向上する可能性が本研究でも指摘されているが、トレーニングの方法を改善することで、学生のプレゼンテーション力がどこまで伸びるかを検証してもらいたい。

③本研究が行われた授業は全クラス統一シラバスによるもので、特に学生による相互評価に大きく影響を及ぼす各学生のプレゼンテーションのテーマがすべて異なっているという、研究結果に大きく影響を及ぼす要因をコントロールできなかったことが惜しまれる。しかし、「授業で受講者全員が同じ教材を学習すれば、共通の内容スキーマを持つことができ、そのテーマについて各学生がさらに調べたことをまとめるとともに、自分の考えを述べることで、今回コントロールできなかった要因をコントロールすることが可能である」と口頭試問時の応答で述べられたが、それを実際に行うことができる授業を近い将来担当できるとのことなので、さらに研究が発展することを期待したい。

学生の相互評価力とそれに影響を及ぼす要因を取り上げた研究はまだ多くないことに加えて、本研究で検討された点をきちんと押さえて行われた研究は非常に少ないことから考えても、本研究が優れたものであると言える。本研究が特に高く評価されるべき点は次の5つである。

- ①様々な制約がある中、行われた6つの研究はその実施から統計処理まで、適切に行われていること（ただし、統計処理結果の解釈については問題がある）。
- ②6つの研究が密接につながっていること。
- ③学生の評価力に影響を及ぼすと考えられる4つの要因（プレゼンテーションの内容を十分に理解できるだけの英語力、その内容に対する予備知識の有無、プレゼンテーション力、そして評価者トレーニング）を取り上げ、その4つの要因が、学生によるプレゼンテーションに対する学生の相互評価にどのような影響を及ぼすかについて、様々な観点から検証していること。
- ④学生の相互評価力の向上の鍵となる評価者トレーニングをきちんと行っていること。
- ⑤評価者トレーニングは、学生の相互評価力を高めるだけでなく、学生のプレゼンテーション力そのものを向上させることができるかを調べた本研究は、この分野のこれまでの先行研究では見られなかった評価者トレーニングの教育的有効性という新しい視点を加えたこと。

本研究は優れた研究であるが、口頭試問時に出されたこれらの改善すべき点のほかに、各審査委員から改善すべき点が指摘されている。しかし、どれもみな加筆または修正することで改善できる点ばかりである。また、今後の課題としては、上で挙げた審査委員からの3点以外に、笠巻氏自身が論文の中で挙げている課題も残されているが、それは本研究の価値を低めるものではなく、むしろ研究の発展の可能性を示すものである。本研究は、学生による相互評価に関するまだ十分に進んでいない研究分野に新しい視点を加えた研究であるとともに、英語による学習者の発信能力を向上させるための指導と評価の改善に貢献できる研究である。

本審査委員会での審議の結果、審査委員全員一致で、本論文が博士（言語文化学）の学位に値するものであるという結論に達した。