論文の和文要旨

題目：	スピーキングテストとしてのグループ・オーラルテストの妥当性について－インタビューテストとの比較を中心に－
名前：	曲　明

言語教育の指導とテストは表裏の関係である。テストの波及効果を考えれば、スピーキング能力の育成には、スピーキングテストを行うことが効果的である。今まで、対面式スピーキングテストと言えば、インタビューテストが主流であったが、しかし、一般教育現場の大規模教室における現状として、スピーキングテストを従来のインタビューテストで行う場合、時間がかかりすぎるというマイナス面があり、多くの教育現場では敬遠されている。そこで、本研究では、中国語教育におけるスピーキングテストの多様化を目標として、大規模教室において、学習者同士の話し合いを評価するグループ・オーラルテスト（以後グループテストと略す）について分析を試みる。日本の中国語教育分野においては、スピーキングテストに関する研究はほとんど行われていないため、本研究では、英語教育分野の先行研究の知恵を借り、多くの積み重ねがあるインタビューテストとの比較を通して日本人大学生の中国語学習者を対象にしたグループテストの妥当性について分析及び考察を行う。具体的にこの二種類のテストの形式の違いが受験者の得点と発話にどのような影響を及ぼすかを明らかにしようとしたものである。

本研究は6章構成である。

第1章では研究の要旨を述べる。研究背景、研究課題、主な結果、研究意義について簡単に紹介し、最後の節では、本研究の構成を紹介する。

第2章は二つの先行研究の概観を行った。一つはスピーキングテストで測る能力、いわゆるスピーキング能力とは何かについて時系列的に言語の能力観を考察し、その後本

研究の研究対象であるグループテストとインタビューテストの違いを考察した。グループテストとインタビューテストの大きな違いは、受験者の「対話者」にある。インタビューテストの評価対象は母語話者または母語話者に近い言語熟達者である試験官と受験者との一対一の会話であるのに対して、グループテストは受験者同士の会話が評価対象である。この「対話者」の違いがテストパフォーマンスに何らかの影響を与える可能性があるとすれば、その違いを量的、及び質的に精査していく必要があると思われる。第二章で行ったもう一つ先行研究の概観はグループテストについての実証的な研究結果のまとめである。1980年代以降、グループテストについて、実証的な研究はいくつか行われた。しかしその多くは得点に関する研究であった。テストの得点には、採点手続き、つまり評価尺度、採点者の主観的な判断による人為的な要素などの影響を含んだものであり、誰がどのような評価尺度を用い、どのような評価方法で評価したかによって受験者の得点は変化し得るものである。したがって、テストの形式の影響を明らかにするには、受験者の得点のみならず、彼らの発話についての研究が必要であると考える。

第3章では、本研究の研究方法、具体的にテストの概要、採点方法、分析方法について述べた。本研究は中国語を学ぶ日本人の大学生96人を対象にスピーキングテストとしてグループテストとインタビューテストの両方を行った。テストの形式の違いが受験者に及ぼす影響を「得点に及ぼす影響」と「発話に及ぼす影響」と二つの研究課題を立てた。採点は分析的採点にし、発音、語彙、文法、流暢さ、コミュニケーションスキルと6つの尺度を決めた。また、得点の採点尺度に合わせて、発話も以下の三つ、①発話の発話量、ターンの長さ、ターンの数　②発話の正確さ、流暢さ、複雑さ　③発話におけるコミュニケーションストラテジーの使用を中心に分析することにした。

本研究では、「得点の分析」を通して、テストの形式の影響のおおよその傾向を明らかにし、「発話の分析」の部分では問いを深め、より細かく分析するために、テストの形式の効果のみならず、テストの形式と言語習熟度の組み合わせによる効果も分析した。尚、

得点の分析では統計手法として相関分析と t 検定を用い、発話の分析は二元配置分散分析と $\chi 2$ 検定を用いた。

第 4 章では、研究結果を述べた。まず、両テストの形式の違いが受験者の得点にどのような影響を与えているのかを、両テストの総合得点及び下位項目得点の相関分析と平均値の差の検定（t 検定）を行い検討した結果、すべての項目において有意な相関が見られた。「総合得点」と「発音」では両テストの得点には高い相関があるが、他の各下位項目では中程度の相関が見られた。また、t 検定の結果では、「発音」、「語彙」、「文法」の 3 項目は、両テストの得点の平均に有意差がなく、一方「流暢さ」では、インタビューテストよりグループテストの方が平均得点が有意に高く、「コミュニケーションスキル」項目では、インタビューテストよりグループテストの方が平均得点が有意に低いことが分かった。

次に、両テストの形式の違いが受験者の発話に及ぼす影響を検証した結果、まず両テストの形式の違いが発話量、ターンの数、ターンの長さに及ぼす影響を検証したところ、テストの形式の違いは発話量に影響しないが、ターンの数に影響を及ぼし、インタビューテストより、グループテストのターンの数の方が有意に多いことが分かった。1 ターンの長さについては、テストの形式と言語習熟度の間に交互作用の効果が認められ、成績の上位群にのみ有意傾向が見られ、グループテストよりインタビューテストの方が 1 ターンの長さが長かった。

次に両テストの形式の違いが発話の正確さ、複雑さ、流暢さに及ぼす影響を検討したところ、テストの形式の違いは発話の正確さの全指標、複雑さの全指標、流暢さの五指標のうちの三指標、つまり「繰り返しの数」、「自己訂正の数」、「出だしの言い間違いの数」に影響を及ぼさなかった。テストの形式の違いが有意の影響を及ぼしたのは流暢さの二指標である「有声ポーズ数」と「無声ポーズ数」であった。テスト形式と言語習熟度の組み合わせの効果は「有声ポーズ数」に影響を及ぼし、成績の下位群にのみ有意傾

向が見られ、グループテストよりインタビューテストの方が多かった。「無声ポーズ数」の項目において、テスト形式と言語習熟度要因の組み合わせの効果による影響は有意ではないが、テスト形式要因のみの効果は「無声ポーズ数」に影響を及ぼし、グループテストよりインタビューテストの方が「無声ポーズ数」が多かった。

最後に両テストの形式の違いがコミュニケーションストラテジーの使用に及ぼす影響を検証した結果、成績の下位群にのみ、両テストでのCSの使用頻度に有意差があった。「直接アピール」の項目に関して、グループテストでの使用頻度は有意に低く、インタビューテストでは有意に高かった。「話題回避」項目において、インタビューテストでの使用頻度は有意に低く、グループテストでは有意に高かった。

両テストで測っている内容に異なる部分があることは、グループテストの実施意義が示されたと言える。

第5章では第4章で得られた研究結果の要因を探った。まず得点分析の結果から両テストで測っている5つの下位能力の中、「発音」、「語彙」、「文法」の3項目は、両テストの得点の平均に有意差がなく、一方「流暢さ」と「コミュニケーションスキル」の2項目においては、得点の平均に有意差が見られたため、テストの形式の違いによって難易度が異なると考えられる。

インタビューテストよりグループテストでの「1ターンの長さ」が短いことから、ターンが極端に短い場合には、文法の複雑さや談話レベルの評価が難しくなる可能性があり、グループテストで文法の複雑さや談話構成能力などを評価対象とする場合には、発話量、ターンの短さなどの発話特徴を考慮した評価尺度を作る必要があるであろう。

グループテストではインタビューテストより流暢である理由として、以下のことが言えるであろう。インタビューテストでは、インタビュアーが発話内容の大部分をコントロールしており、次に何が話されるかを受験者は予測することが難しい。それに対してグループテストでは、受験者同士が協働で会話を作り上げていく中で、受験者自身が話題を提

供したり、相手に質問したりするため、グループテストの方が相手の発話内容を予測しやすいと考えられる。言い換えれば、グループテストはインタビューテストより相手の発話内容を予測しやすいことが流暢さに繋がったと考えられる。

最後に、下位群の受験者はインタビューテストでは「直接アピール」というコミュニケーションストラテジーを有意に多く使い、グループテストでは「話題回避」を有意に多く使うことに関して、両テストにおける対話者の支援の有無がコミュニケーションストラテジーの使用状況の違いに繋がったと考えられる。インタビューテストでは、会話の相手は言語能力が高い母語話者であり、 受験者はコミュニケーション挫折に遭遇する際に、「アピール」さえすれば、ほぼすべてのケースにおいてインタビュアーからの助けがあったため、「話題回避」をせずに会話を続けることができたと思われる。一方、グループテストを受ける時、「間接アピール」で言語運用力の限界を示しても、助けてもらえない時「話題回避」または「伝達回避」になるケースが多かった。この問題を解決するには、成績の下位群の学生たちに様々な発話文脈においても支援をもらえるように「直接アピール」の言い方を初め、また「対話者に対して理解の可否を自分から確かめる」、「簡単な言い換え」など様々なコミュニケーションストラテジーの使用を学生たちに慣れさせるべきではないかと思われる。初級者だからこそ、このような教育上の工夫がスピーキング能力の向上に繋がると考える。

第6章では、これまでの研究のまとめを行い、本稿の示唆と限界について述べた。特に、教育と評価の一貫性、中国語教育におけるコミュニケーション教育への示唆について詳細に論じた。最後に、インフォームドアセスメントという概念を提唱し、教師と学習者の間で、評価の意図や目的を共有することの重要性について示唆を与えた。本研究はグループテストとインタビューテストの形式が受験者のパフォーマンスに及ぼす影響を明らかにすることによって、テストの目的に合わせて適切なテストの形式を選択することの意義と根拠を示すことができたと考えられる。