

グループ・オーラル・テストにおける 相互作用発話機能

堀川 有美

要 旨

本研究は、学習者間の相互作用を評価するテストとしてのグループ・オーラル・テスト(以下、グループテスト)の妥当性を、テスト談話の分析により検証したものである。テストは、中級レベルの会話コースにおいて、学習者の会話力診断及び学習者へのフィードバックを目的として実施された。4グループ12名の学習者の発話を対象に、相互作用発話機能(He & Dai 2006)の分類を用いて、学習者間の相互作用が実際のテスト場面において引き出されているかどうかを調べた。分析の結果、教室内テストとしてのグループテストにおいて、相互作用発話機能が幅広く引き出されていたことがわかった。さらに、各発話機能の出現した談話を分析し、グループテストでは受験者間の自然な会話を評価できる一方で、教回のやりとりを必要とする相互作用は評価対象としにくいことを示した。

【キーワード】 会話能力、 グループ・オーラル・テスト、 相互作用発話機能、 教室内テスト、 内容妥当性

1. はじめに

第二言語・外国語の会話テスト形式の 1 つに、少数の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(以下、グループテスト)がある。グループテストでは学習者同士で話し合いが進行されることから、学習者間の能動的な相互の働きかけや学習者による会話の管理等も含めて評価ができると期待できる。しかし同時に、学習者のみで進行される会話では、評価者の意図した発話が引き出せない可能性や話し合いが停滞する可能性もある。

本研究では、実際のグループテストにおいて学習者による相互作用発話が引き出されているかどうかを調べることで、学習者間の相互作用的な発話を引き出すテスト形式としてのグループテストの妥当性を検証する。

2. 先行研究

2.1 グループテスト研究の概観

グループテストの実施報告は 1970 年代から英語教育の分野でなされてきた(Folland & Robertson 1976 等)。グループテストの長所として、複数の受験者を評価することにより効率が高まること(Swain 2001)、受験者自身が好むテスト形式であること(Fulcher 1996)等が挙げられている。

一方、グループテスト実施の際に懸念されるこ

ととして、学習者の組み合わせによる影響がある。組み合わせにより各受験者のパフォーマンス及び評価結果に違いが生じるとすれば、評価の公平性が損なわれる恐れがある。そこで近年、グループテストにおける対話者の影響についていくつかの観点から研究が行われてきている(Berry 2004 性格; Kobayashi, Johnson, & Van Moere 2005 発話量; Van Moere 2006 言語レベルと性格)。これらの研究では、組み合わせが各受験者のパフォーマンスや評価結果に影響を与える可能性が示唆されている。

しかし、テスト全体の信頼性に関しては、大半のグループテスト実施報告において十分な信頼性が確認されている(Fulcher 1996; Bonk & Ockey 2003 等)。信頼性は個別のテスト実施方法や受験者の特性等により変わってくるため、「グループテストは常に信頼性が高い」と言うことはできないが、学習者に会話の進行を任せるグループテスト形式でも、テスト手続きや採点方法を整備することにより信頼性の高い評価が可能であると言えよう。

それではグループテストではどのような会話を引き出すことができるのだろうか。Skehan(2001)は「グループ形式の口頭テストは、共同構築の観点から相互作用を描くことを可能にする。そこでの談話は、評価者によって前もって決められているものではない」と述べている。また、Bonk & Ockey

(2003: 90)はグループテストが学生に「会話を導入したり管理したりする機会を与える」と述べている。グループテストは、他の口頭テストと同様、発音・語彙・文法等の言語面の評価が可能であるが、それ以外に、相手とどう関わるのか、会話にどう参加するのか、という相互作用的な力も評価の対象とすることができると言える。

近年、多くの第二言語教育現場において、学習者同士の相互作用を重視したペア又はグループワークが取り入れられており、評価場面においてもグループテストを導入することによって学習や指導に肯定的な波及効果(washback)¹を与えることができると考えられる。Bonk & Ockey(2003: 106)も、「今日の言語教室におけるタスクのタイプを考えると、第二言語口頭能力テストにピアのディスカッションや受験者自身が管理する会話が含まれていないことは、重大な欠点である」と述べている。

2.2 テスト談話の分析による内容的妥当性の検証

あるテストが測定しようと意図した能力を適切に測定しているか、また、テスト得点をその能力の指標として解釈できるか、という妥当性(validity)の度合いを確かめることはテスト研究における重要な課題である。バックマン・パーマー(2000: 24)は「ある得点解釈を正当化するためには、テスト得点が測定したいと思っている言語能力の分野を反映しているという証拠を提供する必要がある。」と述べている。

妥当性は複数の側面から成る包括的な概念であるため、テストの妥当性が高いことを示す「証拠」は様々な角度から集められる必要があり、近年はテスト得点の統計分析だけでなく質的分析による妥当性検証も行われてきている(Davies & Elder 2005)。特にスピーキングテストの妥当性検証においては、当該テスト課題によってテスト作成者が意図したような発話を実際に引き出されているかどうかや、構成概念として定義した能力が受験者が実際に使用しているかどうかを検討するために、テスト談話及び受験者のパフォーマンスを詳細に見ていく必要があり、談話分析的手法が多く用いられている。このような分析はテストの内容が評価を行う基礎として適切かどうかという内容的妥当性(content validity)を扱っており(O'Sullivan, Weir, & Saville 2002)、全体的な妥当性の重要な一側面を明らかにすると考えられる。

談話分析的手法を用いてグループテストを分析

した研究に He & Dai(2006)がある。彼らは中国の大学生向けの英語スピーキングテスト CET-SET(College English Test-Speaking English Test)のディスカッション・セクションの談話を分析し、妥当性検証を行った。相互作用的コミュニケーションの評価を目的として導入された CET-SET には、3名の受験者による4分半のディスカッションのセクションが含まれている。He & Dai は、48グループ144名のディスカッションを対象に、相互作用に関わる発話機能(interactional language functions)の出現を調べた。採用された発話機能は、次の8つである。

- a) 賛成・反対する((Dis)agreeing)
- b) 意見や情報を求める(Asking for opinions or information)
- c) 挑戦する(Challenging)
- d) 支持する(Supporting)
- e) 修正する(Modifying)
- f) 説得する(Persuading)
- g) 発展させる(Developing)
- h) 意味交渉する(Negotiating meaning)

このうち a)は他者の発話に賛成か反対かを短く簡単に表明するものであり、明示的には述べられていないが、相互作用の度合いが低いものとみなされている。一方、c)から g)までの5つはディスカッションにおいて他者の発話に反動的(reactive)²であるために必要な機能とされ、相互作用の度合いが高いものとみなされている³。分析の結果、全ての相互作用的発話機能のうち a)が約半数、b)が約4分の1を占め、他者の発話に反動的な発話や意味交渉の出現頻度が少なかったことから、同テストにおける談話が相互作用的になっていなかったと述べている。

He & Dai は、原因として次の4つの可能性を述べている。(1)テストであることを受験者が強く意識していた、(2)正確さや流暢さを重視する受験者はディスカッションに参加するほど言語に自信がなかった、(3)話題に関心がなかった、(4)発話量の多さが評価されると誤解した受験者が多かったため、ターンが長くなり、各自がスピーチを行っているような会話になった。このうち、(1)と(2)は CET-SET が利害関係の大きいテスト(high-stakes test)⁴であることと関連していると考えられる。グループテストを採用したからといってどのような状況でも受験者間の活発な相互作用が生まれるわけではないことを示した点で、He & Dai(2006)の研究は啓発的である。

では、どのような状況であれば相互作用が生まれるだろうか。本研究では、利害関係の小さい(low-stakes)教室内のテストとしてのグループテストを対象に、受験者間で相互作用的な発話が見られるかどうかを調べる。それにより、相互作用を含めて会話能力を評価する手段としてのグループテストの妥当性検証に貢献し、今後のテスト実施に関する示唆を得ることを目的とする。

3. 研究課題と研究方法

3.1 研究課題

以下の二つの研究課題を設定する。

研究課題 1 教室内のグループテストにおける受験者間の会話は、相互作用的吗。

研究課題 2 相互作用的发話機能(He & Dai 2006)は、どのような会話の流れの中で現れるか。また、現れにくい発話機能があるとすれば、その理由は何か。

3.2 対象者とテストの位置づけ

対象としたのは、発音・会話学習を目的として設置された約 3 ヶ月間の日本語クラスにおいて、コース後半の時期に実施された会話テストである。

クラスは 2 クラスあり、2 クラスとも各々週 1 回 2 時間行われ毎回テーマを決めてディスカッションやロールプレイ等の活動を行った。クラスの参加者は合計 15 名であった。そのうち今回分析対象としたのは 12 名である。全員、日本に滞在している韓国人学習者であり職業は主婦・学生・教師・アルバイトであった。日本語学習歴・滞日期間にはばらつきがあるが、コース開始前に実施した SPOT(A 版)⁵の点数及び事前面談における日本語でのやりとりから中級以上の日本語レベルと判断された。

テストは学習者の会話能力診断及びフィードバックを目的として実施された。後日、各受験者の長所と弱点に関する情報をフィードバックしたが、本クラスが研究の一環として実施されたクラスであったことから、テスト結果を成績決定やコース修了認定のために用いることはなかった。従って、結果が何か重大な決定に使われることはない、利害関係の小さいテスト(low-stakes test)であったと言える。

3.3 テストの手続き

評価対象能力つまりテストの構成概念は「相手とのやりとりにおいて自分の考えや意見を日本語で述べる力」とした。評価項目は、Bachman & Palmer

(1996)の言語能力のモデル及び母語話者へのインタビューをもとに発音・語彙・文法・談話構成・流暢さ・自発性・応答の適切さの 7 項目とした⁶。He & Dai(2006)のように発話機能の形では構成概念を定義しなかったが、本研究のテストも単に自己の意見を主張するだけではなく他の受験者との相互作用をしながら会話に参加する力を評価するため、「相互作用のコミュニケーション」を評価する CET-SET と構成概念の内容が重なっていると考えられる。

テストは 1 グループ 3 名で実施した。組み合わせは、コース中の学習者の様子も考慮した上で、レベルや学習者の背景が比較的近い者同士が同じグループになるよう、教師が決定した。

話題は、(1)「仕事を選ぶ基準」、(2)「住むところを選ぶ基準」の 2 つであった。4 グループ中 2 グループは話題(1)、残りの 2 グループは話題(2)で受験した。話題は話し合い促進のための 2、3 の質問を併記したカードにより提示し、開始と終了の合図以外は学習者に進行を任せた。万一会話が停滞したり発話が極端に少ない受験者がいたりした場合には教師が介入するという対処方法を予め決めておいた⁷。時間は Fulcher(1996)に倣い、約 15 分間と設定した⁸。

3.4 分析方法

テスト談話の文字化資料 4 グループ 12 名分を分析データとし、各受験者のターン数及び相互作用的发話機能(interactional language functions 以下、ILF)をカウントした。

ターンは、金(2000)の"turn"の定義⁹を参考に「他の会話参加者の発話やポーズで区切られた発話」を 1 ターンとした。他者の発話中に発せられる理解表示のあいづちはターンとみなさなかった。

ILF の分類は、He & Dai(2006)で用いられた 8 つの機能をそのまま用いたが、各分類の定義は He & Dai(2006)を参考に本研究のデータに合わせて一部修正した。各 ILF の定義は表 1 の通りである。

ILF は、統語上・談話構造上の単位ではなく、機能的な単位である。そのため、1 ターンの中に複数の ILF が現れたり (例:「私もそう思います (賛成)。A さんはどう思いますか? (意見情報要求)」)、2 ターン以上にわたって 1 つの ILF が観察されたりする場合もあった。表 1 の定義に照らし合わせて該当する ILF が出現すれば発話の長さに関わらず 1 回とカウントした。全く ILF が現れないターンもあった。

コーディングでは、まず半数のデータについて 2

名の協力者と協議しながらコーディングを行った。始めに定義を確認した後、別々に作業し、一致しない箇所についてはお互いが納得できるまで話し合い、一致させた。このようにして分類基準を確固にした後、残り半数のデータを筆者1名で行った。

なお、(a)賛成・反対、(c)挑戦、(d)支持の発話は、本研究のような3名以上の談話においては、参加者の間で意見が分かれた場合、どの発話者に対して賛成(または反対)なのかという判断が複雑である¹⁰。本研究では、中田(1991: 300)の提案にしたがい、

表1 相互作用発話機能の定義

<p>(a)賛成／反対する：他の話者の意見に対して賛成か反対かを短く述べる。賛成・反対表示に続けてその理由や証拠を述べた場合は、「支持」「挑戦」に分類する。</p> <p>(b)意見や情報を求める：新しく意見や情報を求める。</p> <p>(c)挑戦する：他の話者の意見や主張に反する理由や証拠を挙げて、異論を唱える。</p> <p>(d)支持する：他の話者の意見や主張を支持する理由や証拠をさらに挙げる。</p> <p>(e)修正する：(直前またはそれ以前の)他の話者の発話を受けて自己の意見や議論を修正する。</p> <p>(f)説得する：他の話者の発話内容に直接は基づかない自己の主張を、相手に受け入れてもらえるように代替案を提示したり表現を工夫したりしてはたらきかける。</p> <p>(g)発展させる：(直前またはそれ以前の)他の話者の発話内容に基づいて、新しい視点を盛り込んだ自己の主張を行う。</p> <p>(h)意味交渉する</p> <p>h1 明確化要求：他の話者の発話の不明な点について説明を求める。</p> <p>h2 明確化：他の話者の要求に応えたり誤解を訂正したりするために、自己の発話を説明し直す。</p> <p>h3 確認要求：自分の理解が合っているかどうか相手に確認を求める。</p> <p>h4 理解確認：聞き手が理解しているかどうかを確認する。</p>
--

4. 結果

4.1 相互作用発話機能(ILF)の割合

分析の結果、受験者のターン数は全体で 531 ターンであり、合計 260 の ILF が見られた。

(1) 全発話(ターン)に占める割合

全発話(ターン)に占める ILF の割合を表 2 に示す。

(a)賛成／反対は相互作用の度合いが低い機能とされているため(He & Dai 2006)、(a)を除いた ILF の割合について検討する。個々の受験者の発話には平均 18.8 個の ILF(a を除く)が見られた。1 ターン当たりでは、平均 0.42 個であった。仮に、カウントされた ILF は全て 1 ターン毎に現れた(1 ターンに 2 個以上の ILF は現れなかった)と単純化して考えると、およそ 4 割のターンに ILF が見られた計算となる。

ILF が見られない発話というのは、相手の発話内容と直接関わりなく自己の主張を述べているもの、質問に応えるだけのもの¹²、あいづち、等である。これらを除いた部分にある程度の頻度で ILF が現れたということは、「各自がスピーチを行っているような会話」ではなかったと言える。

「発話時に文脈において主となっている方向性」に照らし合わせて分類を決定した。

以上の分類を行った後、まず研究課題 1 を明らかにするために、各受験者の全発話(ターン数)に占める ILF の割合と、全 ILF に占める 8 つの ILF それぞれの割合を調べた。8 つの ILF の割合については、He & Dai(2006)の結果との比較も行った。

研究課題 2 に関しては、ILF が現れた箇所の談話例を通して分析を行った。補助資料として、フォローアップインタビュー¹¹の文字化資料を用いた。

(2) 各 ILF の割合

次に、全 ILF 中の各 ILF の割合を He & Dai(2006)の結果と共に表 3 に示す。

He & Dai(2006)では、全 ILF のうち約半数を(a)賛成・反対が占めていた。また、その次の(b)意見情報要求が約 4 分の 1 であり、それ以外の(c)~(h)はどれも低い割合であった。それに対し本研究では、(b)意見情報要求(25%)が最も多く、続いて(h)意味交渉(22.7%)と(c)挑戦(16.5%)が多かった。(a)賛成・反対はその次で 13.5%であった。

つまり、本研究では He & Dai(2006)で相互作用の度合いが低いとされた(a)賛成・反対の占める割合がより低いこと、He & Dai(2006)よりも ILF の種類の偏りが少ないことがわかった。特に、(h)意味交渉と(c)挑戦の機能が多く現れた。反対に割合が低いのは(e)修正と(f)説得であった。12 名中、修正は 5 名の発話に、説得は 7 名の発話に観察された。

研究課題 1 の結果は以下の 3 点にまとめられる。

- ・平均して全体の約 4 割のターンに ILF が見られた。
- ・He & Dai(2006)に比べて、(a)賛成・反対の割合が

低かった。

・He & Dai(2006)に比べて、比較的バランスよく複数種類のILFが現れた。

これらの結果から、He & Dai(2006)と比べて本研究のグループテストのほうが、より相互作用的な話になっていたと言える。

表2 各受験者のターン数及び発話中のILF数(n=12)

	ターン数	ILFの数		ILF/ターン	
		全ILF	(a)を除いて	全ILF	(a)を除いて
平均	44.3	21.7	18.8	0.49	0.42
標準偏差	23.3	8.8	7.2	0.11	0.12

表3 各ILFの割合(%)

	a)賛成反対	b)意見情報要求	c)挑戦	d)支持	e)修正	f)説得	g)発展	h)意味交渉
He&Dai(2006)	49.5	24.0	7.4	4.9	2.0	4.4	4.4	3.4
本研究	13.5	25	16.5	7.3	1.9	5.8	7.3	22.7

4.2 グループテストの相互作用の特徴

次に、グループテストにおける受験者間の相互作用の特徴を探るために、4.1で確認したILFが受験者間の会話の中にどのように現れていたかを見ていく。ここでは、本研究において出現率の高かった4つのILF(意見情報要求、挑戦、意味交渉、賛成反対)と、出現率の低かった2つのILF(修正、説得)に注目し、その前後の発話と会話の流れから相互作用の特徴を分析した。

会話例の文字化規則は、以下の通りである。

[発話の重なり	?	上昇調の抑揚
=	切れ目なく続く発話	hh	笑い等の呼気音
()	括弧内の秒数の沈黙	##	聞き取れない発話
:	音声の引き延ばし	{ }	理解表示のあいづち

4.2.1 出現率の高かったILF

分析の結果、出現率の高かったILFは(1)直前の発話者の発言に対する反応として現れる場合と、(2)直前の発話者や直前の発話内容とは関わりなく現れる場合があることがわかった。

(1) 直前の発話者の発言への反応

これは、直前の相手発話に何らかの反応を示す現れ方であり、相互作用の通常形であると考えられる。意見情報要求、挑戦、意味交渉、賛成反対、の4つのILF全てにこの現れ方が見られた。

さらに、意見情報要求、挑戦、意味交渉の3つのILFは、やりとりの中で同じILFが連続して現れる場合があった。このことも、これらのILFの出現率の高さに影響していると考えられる。以下では、どのように連続して現れているのかを見ていく。

会話例1は、意見情報要求が連続して現れた例

である。話題は「住むところを選ぶ基準」であり、この会話例の前の箇所で受験者Bが自分の住んでいる寮について説明していた。会話例1では、まずBに対してCが質問を行い(1C、3C)、それに対する答えを受けて今度は受験者Aが質問を行っている(5A)。さらにその後も7A、9Cで、AとCが交互にBに対して質問をくり返している。AとCがBの発話内容に関心を持ち、さらなる発話を引き出すとしている様子が窺える。

会話例1 受験者A, B, C

→1C	飲食はつく、作って食べますか?	意見情報要求
2B	え?	
→3C	できない?	意見情報要求
4B	うん、でも、	
→5A	じゃ、ただ寝るだけ?	意見情報要求
6B	(0.4) はい、そうです。バイトするときは、ちゃんと2回 {L ああ、ああ} 出るから、まかない。(1.2) なんか、バイト入ってないときは、近くのコンビニとかで、弁当。{L ふうん} (1.0) はい。	
→7A	(0.8) じゃ、そしたら、あの、バイトの寮に住んでるから、近いんですよ、あの、仕事場所。	意見情報要求
8B	うん、そうです。	
→9C	交通費とかは、いらない。	意見情報要求
10B	すぐです hhh. 1階。	

また、ここではBが情報を提供する側にまわっているが、他の場面ではBがAやCに対して質問を行っている例も観察された。グループテストにおいては複数の参加者がいるため、質問者という役割が一人に固定されにくく、お互いに質問をし合いながら会話を展開していると考えられる。

会話例2は、挑戦の発話の後にさらに挑戦が続

く例である。「自分の好きな仕事、楽しい仕事を探したい」という D と E のやりとりを受けて、F が「仕事はストレスがあるから楽しいものではない」(6F)と反論を行うが、それに対して D と E が同時に「(仕事には)楽しいこともある(7E)」「満足感もある(8D)」と、再反論を行っている。F はそこで終わらずに 9F で「どんな仕事でも満足感はある」と再々反論を行っている。他にも、このようなやりとりのあとに、もとの話者が自己の発話を修正するというパターンも見られた。

会話例 2 受験者 D, E, F

- 1E なんかも、仕事: やりながら、あ、ほかの自分がほんとに好きなこと探して、10年後とか 15 年後には、あ::これ、は、俺に合うって、それを探して。
(略)
- 6F あ、でも、仕事するときには、なんでも、なんなか、ストレスがあるから、楽しいものじゃないと思うんですけど。 **挑戦**
- 7E [でも、楽しいこともある。 **挑戦**
- 8D [でも、たまにあの、その仕事を(0.4) あの: 完成したら、あの、自分なりに満足感もあるし、そんなこともあるんじゃないですか。 **挑戦**
- 9F どんな仕事しても、そんな満足は(0.2)できるんですよ。自分が一所懸命したら。 **挑戦**

次に、意味交渉の発話がが続いている例を見る。意味交渉には、4 つの下位機能(明確化要求、明確化、確認要求、理解確認)が含まれているが、会話例 3 は、明確化要求と明確化が行われている例である。まず I の明確化要求(2I)を受けて H が明確化をしているが(3H)、なかなか発話意図が伝わらないため、明確化をくり返している(5H, 8H)。

明確化は、意味交渉の 4 つの下位機能の中で、最も出現頻度が多かった。相手からの明確化要求がない場合でも、自己の発話意図が誤解されていると気づいたときに自らすすんで明確化を行う例も見られた。

一方で、明確化要求をくり返す例は見られず、相手の発話に不明確な部分があっても聞き手の側がそれを追求するという事は起こりにくいことがわかった。フォローアップインタビューで「分からない言葉があっても、聞き直さなかった」と述べた受験者(K)もあり、自己の発話に関する意味交渉に比べて、相手の発話に対する意味交渉はそれほど積極的に行われない可能性が示唆された。

会話例 3 受験者 G, H, I

- 1H (0.3) あの、あれあるんじゃないですか、お金がたくさんもらえる(;)ですけども、働く時間が長いとか、働く時間が長い###、短いけどお金が少ないとか、こ、両方あるんじゃないですか。(0.4) 両方あったら、どっち選びます? **意見情報要求**
- 2I (2.2) だ、時間が長く、て;、 **明確化要求**
- 3H 働く時間が長く(;)て、お金がたくさんもらうとか、で、そんなに仕事はしなくても、お、も、お金はちょっとあんまりもらえないけど、ないけど、自分、自身の自由時間がほしいとか、(1.6) お金のほうがいいか、時間のほうがいいか、2 つの中で選ぶならどっちがいい? **明確化**
- 4G (0.4) えそれはそのときの都合によってちがうんです。=
- 5H =えだから、そんなこと関係なしで、もし、同じ状況だったら。 **明確化**
- 6G (0.3) 同じ状況だったら。(4.8) ちょっと難しいそれは hhh.
- 7I 私は;。(0.4) 短く働くほうがいいですね。(1.0) 働きに日本に来たんじゃないですか;.
- 8H あ、かんこ、日本だけじゃなくて、自分の。 **明確化**
- 9I うん、自分の国だとしても;。(0.9) こう、はたら、はたら、ばかり、だったら、もうすぐ疲れて、すぐ、やめちゃうと思います。

(2) 直前の発話者や発話内容とは無関係

出現率の高かった 4 つの ILF の中でも、意見情報要求には、直前の発話者や発話内容とは直接関わりをもたずに現れる例が観察された。それらは、新しい話題を導入したり発言権を取得したりするという会話管理的な機能を併せ持っていた。

会話例 4 は、会話例 3 の続きである。「お金か時間か」という話題が続いていたが、3G で G が「大きい会社と小さい会社のどちらを選ぶか」という新しいテーマで H と I に意見を求めている。G はフォローアップインタビューで「私が話したい話題にどういうふうに言って話を戻せばいいのか、そういうのも考えながら話した」「自分が話したい話題に戻すために、逆に質問した」と述べている。

会話例 4 受験者 G, H, I

- II (2.3) で例えば私の趣味って、公演見に行ったり、こう、こう、人; 人に会って、集まりに行ったり、そういういいことが好きですから。(0.9)うん、で、仕事して;。(0.4)こう、自由な時間に、やりたいこと、(1.0)できるほうがいいですね。
- 2H (3.8) 時間のほうがいいんじゃないか。
- 3G (1.2)そうですね、今まではちょっとバイトの?話だけしたんですけども、ほんとも

し？、あの将来、近い将来に自分が？仕事を
選ぶとき、なんといったらいいんでしょう
か、なんかこう、大きな？会社で、なんかこ
う、給料も？なんか、保証されますよ。でも、
ちょっと自分がなんか面白くないな。って
そう思う仕事と、ちょっと小さい会社でもな
んか将来性があるって、(0.4) あ、このしご、あ
の会社だったら、自分が何でもできる、何で
も挑戦できる。そういう、あの 2 つの会社が
あったとき、(0.2) あの会社があったとした
ら、どっちを選びますか？(0.4)若い世代は
hhh. **意見情報要求**

4I 私は、自分が楽しく働ける仕事を選びます。

会話例 5 は、発話していない他の受験者に発言
権を譲る例である。受験者 J、K2 人のみで会話
が進行しており L が参加していない状況であったが、
6J において J が L に向けて情報要求を行い、それ
によってその後 L が発言権を得ている。内容的
には同じ話題であるが、6J の発話はその話題につ
いて話していた相手への直接的な反応ではなく、他
の話者へ発言権を渡すという会話管理的な機能を併せ
持っていることがわかる。

会話例 5 受験者 J, K, L

- 1J で自転車。(0.3) ま、いま持ってるけど、や
あの。冬、冬は寒い、寒いから乗れないし、
夏は暑いから乗れないし、hhh
2K 暑いし雨 hhh **支持**
3J 雨が降ってるときもそうだし。
4K そうですね。 **賛成**
5J あんまり、あの、近く：に、あの、(1.7) 地下
鉄とかバスが、一番便利：なところだから。;
あんまり、hhh 自転車も必要 hhh じゃないで
すよね。
→6J (3.2) L さんも、自転車で、あの、買い物した
りしますか？ **意見情報要求**
7L たまたまですよね、もう、ふだんは歩いて
行くけど、なんか、米、うん、米、米？お米
とか重いでしょうかね。5 キロと 10 キログ
らいだったら、自転車乗るけれど (略)

4.2.2 出現率の低かった ILF

次に、全体における出現率の低かった ILF、修正
と説得が現れた箇所を検討した結果、これらは数回
のやりとりを経た後に出現していることがわかった。

まず、修正の例を見る (会話例 6)。挑戦を含む
やりとりの中で D は「最初はできなくても成功す
る人もたくさんいる(4D)」と述べているが、D の反
対を受け「(たくさんではなく)たまにいる(6D)」と
自己の主張を修正している。

会話例 6 受験者 D, F

- 1F あ：私の場合は、自分が (0.3) できる仕事な
ら、いいと思うんですよ。
2D お金ももらうから。hhh
3F これじゃなくて、どんな仕事にしても、自分
ができることのものであれば、いいと思うんで
す。(1.0) できないこととするときには、なんだ
かストレスもたくさんもらう#####。(1.0) あ
んまりがんばっても、できない仕事もあるん
じゃないですか。 **明確化**
4D でも前の、成功した人を見ると、自分が最初
にはできないとしても、一緒にがんばって(;)あ
の最後には、成功する人もたくさんいます。 **挑戦**
5F (1.2)たくさんじゃないんです。 **反対**
全 hhh
→6D たまにいます。たまに。 **修正**

説得も数回のやりとりを経てから現れていた。会
話例 7 は、「田舎に住みたい」という C に、A と B
が反対している会話である。特に B は「私はにぎ
やかな雰囲気が好き(2B)」 「田舎に住んだら仕事
がない(6B)」と、反対や挑戦の発話をくり返し行っ
ている。しかしなかなか C に受け入れられないため、
「住むんじゃなくて、田舎に別荘があったらいい
(12B、17B)」と新しい案を述べている。より相手
が受け入れやすい形で、自己の主張を理解してもら
おうとしている様子が窺える。

会話例 7 受験者 A, B, C

- 1C 私はなんか、(0.8)にぎやかな？{A うん} (1.0)
場所は、場所はなんかほしくないんです。
2B え。私あの、にぎやかな雰囲気が好きなんで
す。 **反対**
3A 私も。hhh **賛成**
4B (0.2) つらいじゃないですか。なんか何もな
いし。 **挑戦**
5C あつらく、なんか楽に、なんかかいじょう？
かいじょうして。
6B なんか仕事として、やりたいこと、は何です
か？田舎に住んだら、やることないし。
意見情報要求
7A 農業。hhh
8B [職業になんか、
9C [なんかね、じを、なんか、作家になりたい。
10B え。:
11C なんかあとで、あれ、
→12B だから、住むんじゃなくて、hhh 住むんじ
ゃなくて、い、田舎に別荘とか。 **説得**
13A そうそうそう。なんかね。 **賛成**
14C あ：住みたい住みたい。別荘はきらい。 **反対**
15B なんで？ **意見情報要求**
16C 別荘はき、いらないだと思います。
→17B 私はね、なんか、都市に住んで、なんか行く場
所、え、別荘があったらいいと思います。 **説得**

このように、修正や説得の出現率が低かったのは、これらの機能の出現に数回のやりとりが必要とされるためであると考えられる。したがって、意見情報要求等の ILF と比べて修正や説得の ILF の出現が少なくなるのは、自然な相互作用においても同様であると推測される。

5. 考察

以上の結果をふまえて、グループテストの妥当性と実施意義について考察を行う。研究課題 1 の結果として、本研究では He & Dai(2006)よりも多くの相互作用的発話機能が見られたことがわかった。妥当性は一回のテスト毎に検証されるべきものであり、両テストの妥当性の程度は単純に比較できないが、ここでは、各テストの実施条件の違いから実施条件が受験者のパフォーマンスに与える影響を探りたい。

本研究のほうがより多くの相互作用的発話機能が見られた理由として、本研究で実施したテストが教室内で実施された利害関係の小さいテストであり、テスト場面であることを受験者があまり意識していなかったことが考えられる。そのため、自分の力をよりよく示さなければいけないというプレッシャーをほとんど感じることなく、普段の会話に近い形で相互作用がなされたのではないだろうか。

もちろん、教室内のテストであっても成績決定やクラス分けを目的とする場合には、受験者へのプレッシャーが大きくなると考えられる。しかし、普段の教室活動と関連したテスト課題を、クラスメートと一緒に受験できるテスト状況は、そうしたプレッシャーを最小限にとどめることができるのではないだろうか。本研究においても、受験者が同じクラスで学ぶ既知関係の学習者であったことが、より自然な相互作用を生み出すのに影響したと思われる。

ILF の種類別に見た場合、He & Dai(2006)よりは偏りが少ないものの、本研究においても各 ILF の出現頻度に違いがあった。研究課題 2 でそれらの ILF が現れた会話の流れを分析したところ、以下のようなことがわかった。

- ・ 出現率の高い ILF(意見情報要求、意味交渉、挑戦、賛成反対)は、直前の相手発話に対する反応として現れる場合が多い。同じ ILF が連続して現れることもある。
- ・ 意見情報要求の発話に関しては、会話管理的機能を併せ持って現れることもある。

- ・ 出現率の低い ILF(修正、説得)は、数回のやりとりを経たあとの段階で現れる機能である。

受験者間で意見情報要求や挑戦等の発話がくり返されながら活発に話し合いが行われていたことや、受験者自身が会話管理を行っていたことから、学習者同士で進行するテスト形式であっても会話が停滞することなく、学習者間の自然な会話を評価対象とすることが可能であることが確認できたと言える。

一方で、グループテストで受験者の相互作用の仕方を評価する場合、直前の発話者や発話内容に対する即時的な反応を観察することは可能だが、数回のやりとりの中で最終的に意見を修正したり、相手を説得しようとしたりする発話は、評価対象としにくいことがわかった。

お互いの意見をぶつけ合うだけでなく、どのように意見を収斂させていくかということまで含めて評価を行う場合には、テストの話題設定を工夫する必要がある。本研究では、「仕事を選ぶ基準は何か」「住むところを選ぶ基準は何か」という話題を用いたが、「最も重要な基準は何か」というようにグループで 1 つの結論を出すことを課題にすることも考えられる。それにより、また違った相互作用が現れると考えられる。

グループテストには、学習者の組み合わせや採点の仕方等、信頼性に関わる問題が残されている。それでも、学習者同士が話し合う活動をテスト場面に取り入れることによって期待される肯定的波及効果を重視する場合、本研究で示したような受験者間の相互作用が自然に生まれる形式は意義があるだろう。フォローアップインタビューでグループテストに対する感想を述べた受験者 B は、「(グループテストでは)他の人を見て私も勉強になる」「他の人が言うときとか、あんな言い方もあったねとか(気づく)」と述べている。テストも学習者にとっては学習の場の 1 つであることを改めて思い起こさせられる。

6. まとめと今後の課題

本研究では、教室内テストとして実施したグループテストにおいて受験者間の相互作用が引き出されていたことを示した。相互作用を含めた会話力を評価する方法として、グループテストの内容的妥当性は高いと言える。

本研究では、内容的妥当性のみ焦点をあてたため、今後、評価の方法や評価結果を含めて検証を

重ねていく必要がある。また、本研究では2クラス12名の学習者のみを対象としたため、さらに対象を増やして検証を続けたい。今回、受験者が意見情報要求という形で会話管理を行っていることが確認されたが、他の方法で会話管理を行っているかどうか等、グループテストの特徴を反映していると考えられる面についてもさらに分析を行っていきたい。

注

1. 波及効果とは、「テストが受験者、教育・学習、教育システムや社会に及ぼすと考えられる影響」(静等 2002: 189)のことであるが、本稿では主に学習者及び教師への影響を想定してこの用語を用いた。学習者も教師も、テストで出題される内容に、より力を入れて学習や指導を行うのが普通である。従って、教室での学習活動とテスト課題を一致させることにより、学習や指導の効果が高まると考えられる。また、学習内容と一貫性のある評価を行うことで、評価結果とその後の学習を直接関連づけることができる。
2. 「反応的」とは、発話が他の参加者の先行発話に依存している二つの発話の関係を指す(Young & Milanovic 1992)が、He & Daiは「相互作用の度合いが高い」といった意味でこの語を用いていると思われる。
3. b)及びh)の相互作用の度合いに関しては述べられていない。これらはデータを通して慎重に考えていく必要があるためだと考えられる。本研究ではHe & Dai同様、a)は他のILFより相互作用の度合いが低いとみなす。a)は、その他のILFに比べてその後の発話内容に積極的に関与しないと考えられるためである。b)~h)の間の相互作用の度合いの差については区別しないこととする。
4. 当該テストの結果が多数の人々の人生に重要な結果を及ぼすテスト。特定の教育機関への入学、奨学金の給付、資格等の決定に用いられる。(静等 2002: 88)
5. 対象者の得点は、52点以上であった。
6. テストの採点方法及び評価基準については、本稿の目的の範囲外であるため稿を改めて述べたい。
7. 実際に1つのグループでテスト後半に会話が一時ストップしたが、すぐに受験者の発言により会話が開いたため、教師の介入は行われなかった。
8. その他のグループテストの報告でも、ディスカッションのための時間を約10-15分に設定しているものが多い(Bonk & Ockey 2003 等)。
9. 金(2000: 83)は、「turn」を「話し手が話し始めてから他者の発話やポーズで区切られ、話すのをやめるまでのひとまとまりの発話」と定義している。
10. 「直前の発話と規定することもできるが、実際の談話では、対象となる発話の後にコメントなりいくつかの発話のがはさまってから同調・非同調の色合いを持った発話が出てくることもある。」(中田 1991: 300)
11. テスト終了直後に記入してもらったアンケートをもとに、約2~3週間後に個別にインタビューを実施した。

本稿に関連のある質問は、「テスト中に感じたことは何か」「テスト中に気をつけていたことは何か」等。

12. 質問への応答も、相手発話を受けて発話するという意味においては「相互作用性」と考えられるが、相手にはたらきかける要素が少なく消極的な相互作用とみなされることから、本研究においてもHe & Dai(2006)同様、ILFには含めなかった。

参考文献

- 金志宣 (2000) 「turn および turn-taking のカテゴリー化の試み—韓・日の対照会話分析—」『日本語教育』105号, 81-90.
- 静哲人, 竹内理, 吉澤清美 (2002) 『外国語教育リサーチとテストの基礎概念』関西大学出版部
- 中田智子 (1991) 「発話分析の観点—多角的な特徴記述のために—」国立国語研究所報 103『研究報告集 12』秀英出版, 279-306.
- バックマン L. F., パーマー A. S. (著) 大友賢二, ランドルフ・スラッシャー(監訳) (2000) 『実践言語テスト作成法』大修館書店 原著 Bachman, L. F. & Palmer, A. S. (1996) *Language testing in practice*. Oxford: Oxford University Press.
- Berry, V. (2004) A study of the interaction between individual personality differences and oral performance test facets. Unpublished PhD thesis, Kings College, University of London.
- Bonk, W. J. & Ockey, G. J. (2003) A many-facet Rasch analysis of the second language group oral discussion task. *Language Testing*, 20(1), 89-110.
- Davies, A. & Elder, C. (2005) Validity and Validation in Language Testing. In Hinkel, E. (ed.) *Handbook of research in second language teaching and learning*. New Jersey: Lawrence Erlbaum Associates, 795-813.
- Folland, D. & Robertson, D. (1976) Towards objectivity in group oral testing. *English Language Teaching Journal* 30, 156-167.
- Fulcher, G. (1996) Testing tasks: issues in task design and the group oral. *Language Testing*, 13(1), 23-51.
- He, L. & Dai, Y. (2006) A corpus-based investigation into the validity of the CET-SET group discussion. *Language Testing*, 23(3), 370-401.
- Kobayashi, M., Johnson, K. & Van Moere, A. (2005) Effects of Quantity and Quality of Students' Output in Group Oral Tests. 神田外語大学言語教育研究所『言語教育研究』16号, 275-295.
- O'Sullivan, B., Weir, C. J. & Saville, N. (2002) Using observation checklists to validate speaking-test tasks. *Language Testing* 19(1), 33-56.
- Skehan, P. (2001) Tasks and language performance assessment. In Bygate, M., Skehan, P., & Swain, M. *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing*. London: Longman, 167-185.

Swain, M. (2001) Examining dialogue: another approach to content specification and to validating inferences drawn from test scores. *Language Testing*, 18(3), 275-302.
Young, R. & Milanovic, M. (1992) Discourse variation in oral

proficiency interviews. *Studies in Second Language Acquisition*, 14, 403-424.
Van Moere, A. (2006) Validity evidence in a university group oral test. *Language Testing*, 23(4), 411-440.

ほりかわ ゆみ／お茶の水女子大学大学院 応用日本語論講座
yumih@mub.biglobe.ne.jp

稿末資料 1

相互作用発話機能の発話例 (※本文中に例がないもののみ)

(g)発展させる (他の話者の発話内容に基づいて、新しい視点を盛り込んだ自己の主張を行う)

E なんか、これ、テレビで見たら、ほんとに自分の仕事が、ほんとに子どものときから、好きになって、この、このために、なんか 20 年ぐらい勉強し、とか、いろんなことやって、それで仕事やり-やっている人もたくさんいるんじゃないですか。

(中略)

D その人は特別じゃないですか。

E うん。

→ D 子どものとき自分がしたい仕事を見つけて、ずっとこの仕事のためにいっしょけんめい勉強する人は、たまに、会うけど、いるけど、ふつうの人はそうできないんですね。(0.4)それで私は、今はコンピュータの仕事してますけど、(0.3) 次の仕事、あの、移転するときに?とか、基準は、仕事しながら、ほかの人もできる、(1.2)レジャーとか、旅行とか、そんなこともできる仕事なら、仕事しながら私も、あの休める、(0.8)仕事なら、それが一番いいと

(h)意味交渉する

h3 確認要求 (自分の理解が合っているかどうか相手に確認を求める)

J で今 2 階なんで、すぐ、あの、ん、(,)ペランダの下で、あの:毎日、朝ね、(,)あの郵便局の、(0.4)宅急便とかねそういう車置き場があるから、

K あ:.

J ちょっとうるさいですよ。

→ K あ:2 階な、ですね。

h4 理解確認 (聞き手が理解しているかどうかを確認する)

→ A あ、私はあの川口のほうなんですけど、あの、今住んでいる:家は、あの、(0.4) 公団住宅って知ってます?

B あ知らない。

Interactional Language Functions in Group Oral Test

HORIKAWA Yumi

Abstract

This study investigates the content validity of a group oral test by analyzing the test discourse. Twelve students enrolled in an intermediate Japanese course formed four groups and underwent the group oral test; the discourse was analyzed using interactional language functions(ILFs)(He & Dai 2006). The results showed that a variety of ILFs were elicited, and suggested that the group oral test can assess students' conversation ability during interaction. On the other hand, some functions involving exchanges over several turns were observed only a few times; these functions are difficult to assess by the group oral test.

【Keywords】 speaking ability, group oral test, interactional language functions, classroom-test, content validity

(Department of Applied Japanese Linguistics, Graduate School, Ochanomizu University)