

学習者間で行うペア・インタビューテストにおける「聞き手」評価項目の表面妥当性

徳間 望・森川 舞子

要 旨

表面妥当性検証法は、そのテストが許容されていたかどうかをテスト使用者の直感的な判断から検討する。表面妥当性の確保は、学習の動機づけを左右するという点で非常に重要である。本研究では、学習者間で「聞き手」の役割を担う形式のインタビューテストを実施し、教師の提案した「聞き手」評価項目の表面妥当性を検討した。研究課題は、(a) 学習者は「聞き手」評価項目を適切だと感じていたか、(b) どのような理由から適切、あるいは適切でないと感じていたかである。韓国の大学生 35 人を対象に自由記述式アンケートを実施し、KJ 法を参考に分類・分析を行った結果、(c) 多くの学習者が「聞き手」評価項目を適切だと捉え、表面妥当性は支持されていた。(d) 表面妥当性を支持する理由として、i) インタビューという課題の特性からインタビュアーの役割(会話の進行、インタビュー内容の充実や相手への配慮、正確な理解と発信)を果たすのは当然、ii) インタビュアーに必要とされる技能習得(相づち、言い換え、聞き返し)は、日本語会話能力伸長にも有用、iii) 対話のスムーズな進行には聞き手の積極的な会話参加が必要、などが挙げられた。

【キーワード】学習者間、ペア・インタビューテスト、「聞き手」評価項目、表面妥当性

1. 問題の所在

日本語会話クラスを担当する教師なら誰もが一度は直面する問題の一つに、会話パフォーマンスの評価がある。教師が何十人もの学習者を一人ひとりテストするのは時間がかかり、疲労にもなって全ての学習者に対し均一に試験を実施することは難しくなってくる。特に、リアル・タイムで評価する場合は、評価の観点が徐々にぶれていくことに対する不安がどうしてもつきまとう。また、Face-to-face で会話パフォーマンステストを行う場合、多くは教師が対話者となる。しかし、それにより対話者間の教師-学習者という力の不均等が生じ、会話の主導権も教師が握ってしまいがちになるという問題もある。

この問題への解決策の一つに、学習者をペア、あるいはグループにして評価する方法が考えられる。現に EFL 教育では、近年のコミュニカティブ・アプローチ教授法への移行に合わせるように、評価もペアやグループ形式が採用されるようになっている (Taylor & Wigglesworth 2009)。日本語教育でも、学習者間で行うペアやグループ形式の会話パフォーマンステストの取り組みが報告され始めているが(堀川 2006, 堀川・徳間 2008, 久野・齊木 2009, 徳間 2009 等)、まだ一般的とはいえないようである。パフォーマンステストは従来から、実施方法や評価

方法が難しいテストであると言われているが (Brooks 2009, Taylor & Wigglesworth 2009 等)、会話クラスの評価を会話パフォーマンスによって行うのは、妥当性の面から自然な流れであろう。

2. インタビュー形式採用のメリット

会話クラスのパフォーマンステストは、インタビューの他にもロールプレイやスピーチなど、様々な形式が取り入れられている。通常、日常生活でインタビューそのものを行う人はあまりいない。そのため、授業で行うインタビュー活動も、一見日常生活とは関係が薄くなりがちに思われる。しかし、現実のコミュニケーションを考えたとき、インタビューの形式になっているやり取りは案外多い。例えば、道を聞く、イベントの問い合わせをするなどである。これらは、自分が知りたい情報を得るために、情報を持っていると思われる人に質問して情報を得るという点で、インタビューとよく似ており(国際交流基金 2007)、授業でも現実感を持った活動としてアレンジがしやすい。また、インタビュー形式は評価に際して、「聞き手」「話し手」の役割がはっきりと分かれるため、学習者はテスト準備やパフォーマンスがしやすいという利点もある。

3. 「聞き手」評価項目検討の意義

現実のコミュニケーションは、双方向のやり取りから成立している。堀口(1988)で「話し手が話を進めていくためには、聞き手からの反応や働きかけや助けが必要であり、話しことばによるコミュニケーションは聞き手の積極的な参加によって成立する」(p.14)と指摘されているように、会話では対話者である聞き手が重要な役割を担っている。同様に、インタビューテストでも、聞き手がどのように会話を進めていくかは、話し手の話しやすさを大きく左右するポイントになる。

徳間(2009)は、教室内テストとして学習者間で行うペア形式のインタビューテスト(以下、ペア・インタビューテスト)を実施し、緊張が緩和され学習者の情意的側面にプラスに働くこと、テストの真正性向上に有効であることなど、肯定的な波及効果があったことを報告している。しかし一方で、35人中15人の学習者がペア・インタビューテストでは「聞き手(インタビュアー)¹の役割が難しい」と感じていた。その理由として徳間は、「聞き手」評価項目の配点が低く、準備に意識が向きにくかったことや、聞き手の言語機能を特に意識させる活動や評価項目の不足を挙げており、配点も含めた「聞き手」評価項目の見直しが課題となっていた。一つのテストは複数の側面から評価できるが、教師、学習者双方にとって明確な「聞き手」評価項目の検討は、前述した評価の観点がぶれることへの根本的解決にもつながる先決問題と言える。

4. パフォーマンステストの妥当性検討

テストの三要素には、妥当性、信頼性、実行可能性がある。妥当性は何を測定しているか、信頼性は正確に測定しているかであり、実行可能性を含め、これら三要素のバランスが大切であるとされている(マクナマラ 2004)。

一方、オルダーソン他(2010)は、測定したい能力を測定しているかどうかは、正確に測定しているかどうか優先する、つまり妥当性あっての信頼性という立場をとる。オルダーソン他では、妥当性を、内的妥当性(internal validity)および外的妥当性(external validity)に区別し²、内的妥当性の検証方法としては、一般に、表面妥当性検証法(face validity)、内容妥当性検証法(content validation)、応答妥当性検証法(response validation)があるとしている³。本研究

では、このうち表面妥当性検証法に着目する。

表面妥当性検証法は、テストの専門家ではない学習者やテスト使用者がテストに関して行ったコメントをデータとして、テストが使用者から許容されているかどうかを検討する。検定は、テストを受けた学習者等の直感的な判断を基に行われるために、テスト研究の分野では非科学的であるとして否定的に受け取られる傾向もあるが、オルダーソン他(2010)は重要な概念であるという立場をとっている。その理由として、「第一に、表面妥当性に欠けるテストは使用者からすれば目的に合っていないテストということになるので、真摯に受け止められない可能性があるからである。第二に、テストの妥当性が高いと見えるならば受験者もそれだけ真剣に解答に取り組むであろうから、適切な解答を行う、すなわち応答妥当性も高くなると思われるからである」(p.158)と述べている。

本研究でも、学習者自身がその評価項目を学習の指針として有用だと認識すれば、学習や指導に肯定的波及効果が生じると考えられることから、「聞き手」評価項目の表面妥当性を検討する。評価項目の達成を目指すことが、すなわち実際の会話場面でも役に立つと学習者に認識されるかどうかは、学習者のテスト準備のみならず、会話学習全体に対する動機づけも左右する。その点から、表面妥当性の確保は重要である。

5. 研究課題

以上から、本研究ではペア・インタビューテストを実施し、「聞き手」評価項目の表面妥当性を検討する取り組みを報告する。研究課題は、1) 学習者は「聞き手」評価項目を適切だと感じていたか、2) どのような理由から適切、あるいは適切でないと感じていたか、の2点とした。

6. 実践の概要

6.1 対象者

対象者は、2009年後期に韓国の大学で外国語として「中級日本語会話」の授業を履修した学習者2クラス計35人である⁴。ただし、このコースではプレースメントテストを行わないため、受講生の会話レベルには初級後半程度から上級までかなりのばらつきがあった。コースは学部2年生を対象とした第一専攻選択科目で、対象者の専攻は二人を除き日本語である。

6.2 授業の概要

調査対象としたクラスでは、実際のコミュニケーションを意識しながら、よい「話し手」と「聞き手」の役割について学ぶことを目指し、学習目標は、①相手の質問に答えるだけでなく、自ら会話を積極的に進めていく力をつける、②日本語での「相づち」の打ち方、わからない表現や語彙があった時の「聞き返し」の仕方などを学ぶ、③相手にわかりやすい話し方、発音を学ぶ、の3点とした。学習者は普段から授業で2~3人が組になりインタビュー形式の活動を多く行っていた。その際、インタビュアーには1問1答で話を進めるのではなく、話し手が話す内容をよく聞き、話の流れに合わせて会話を進めるよう指導した。

6.3 テストの概要

ペア・インタビューテストは、2009年後期の期末テストとして行われた。このクラスでは、中間テストも同じ形式でテストが実施されたため、全ての学習者は少なくとも1回はこの形式での試験を受けた経験があった。期末テストの一週間前には、学習者にテストの方法と、評価項目表(表1)を提示した。

表1 評価項目表

聞き手 (インタビュアー)	配点
A 開始、終了、相手の話の促しなど、インタビューの進行ができたか	4×3
B 相づちを打ったり、理解を示すことができたか	4×3
聞き手・話し手共通	
C 言い換えて説明したり、聞き返したりできたか	4×3
話し手 (インタビュイー)	
D 発音(聞き取りやすいか)	4
E 語彙(話題に合った語彙を使っているか、語彙量)	4
F 文法(大きな文法の間違いはないか)	4
G 流暢さ(不自然に長く沈黙していないか)	4
合計	52点

テストの流れを説明する。試験の時間はそれぞれのクラスで120分間あり、十分な会話を引き出すため1つのペアにつき合計約12分間とした。テストは3つのセッションに分かれている。第1セッションでは、学習者Aがインタビュアー、学習者Bが話し手の役割でインタビューをする。学習者Aは学習者Bに対して5分間インタビューし、5分後、教師の合図でインタビューを終了する。第2セッ

ションでは、学習者Aと学習者Bは役割を交代してインタビューをする。同じく、5分後に教師の合図でインタビューを終了する。第3セッションでは、第1、第2セッションで教師が気付いた点について口頭で2分ほどフィードバックした⁵。

インタビューのテーマは、期末テストまでに授業中扱ったものの中から、「私の性格」「将来就きたい仕事」「最も感銘を受けた本/映画」「友情」の4つとし、事前に学習者に告知した。学習者は話し手としてパフォーマンスする際、この4つの中から自分で話したいテーマを1つ選んでテストを受けた。また、ペアの組み合わせは、試験の公平性を保つため、レベルが違いすぎないこと、そして普段から親しくしている学習者同士は避けることの2点を考慮して、クラスの担当教師が決定した。組み合わせは学習者が事前にスクリプトを作り演技することができないよう、当日まで知らせなかった。

6.4 評価項目の設定と評価方法

前述した表1の評価項目表は、A~Gまで7項目あり、国際交流基金(2007)、堀川・徳間(2008)、徳間(2009)等を参考に、学習目標に沿うよう設定した。具体的には、学習目標①「自ら会話を積極的に進めていく力をつける」は、表1のAに対応し、②「相づちの打ち方、聞き返しの仕方などを学ぶ」は、B及びCに対応している。③「相手にわかりやすい話し方、発音を学ぶ」は、A~G全てと関係するが、D~Gの評価項目として重点的に評価することとした。

それぞれの項目は、4点~1点の4段階(4:大変良い、3:良い、2:普通、1:もう少し)で評価した(稿末資料1参照)。学習者は「聞き手」「聞き手・話し手共通」「話し手」評価項目の合計点を最終的な得点とした。ただし、「聞き手・話し手共通」評価項目は、第1セッション、第2セッションを通して1つの得点が与えられた。

「聞き手」「聞き手・話し手共通」評価項目は、重み付け(weighting)⁶を行い、それぞれの素点を3倍することにした。配点を変えることによって信頼性や妥当性が増すことはあまりないと言われているが、受験者にカリキュラム上どこがより重要かを認知させるなど、教育上の理由から重み付けが必要になることがある(オルダーソン他2010)。本研究でも、クラスでの重点指導項目を際立たせ、「聞き手」役割に対する十分なテスト準備を喚起するため重み付けを行った。なお、テストは学習者に説明の上すべ

て録音された。評価はクラスの担当教師1名がテスト当日および後日、録音を聞き返して行った。

7. データと分析方法

テストの翌週、ペア・インタビューテストの「聞き手」評価項目（以下、「聞き手・話し手共通」評価項目も含む）を学習者がどのように捉えていたかを調査するため、アンケートを実施した(稿末資料2)。質問および回答は、学習者の負担を減らし率直な回答を得るため韓国語で行った。質問内容は、「ABCの『聞き手』評価項目の設定は適切だったと思いますか。理由を教えてください」である。

回答の日本語訳は2名の日本人日本語教師が担当し、不明な箇所は韓国人日本語話者に確認した。分類・分析は2名の日本人日本語教師で行った。アンケートの自由記述回答をデータとし、KJ法を参考に以下の手順で分析を行った。

- 1) 自由記述回答の1つのアイデアにつき1枚のカードを書く。1つの自由記述回答に2つ以上のアイデアが含まれている場合には、分割し、それぞれを1枚のカードに書く。
- 2) カードの中で似通ったものをいくつかのグループにまとめ、それぞれに見出しをつける。
- 3) 2人の分析者が各自の分析結果を持ち寄り、比較しながら最終的なカテゴリーを決定する。

8. 結果と考察

8.1 研究課題1

研究課題1では、学習者は「聞き手」評価項目を適切だと感じていたかについてまとめた(表2)。

表2 「聞き手」評価項目は適切だったか

	「聞き手」評価項目	適切/適切でない
A	開始、終了、相手の話の促し	35人/0人
B	相づちを打つ、理解を示す	32人/3人
C	言い換え、聞き返し	30人/5人

以上に示すように、ABCいずれの評価項目に対しても、多くの学習者は適切であったと回答していたが、Cの項目中に、適切でないという回答が見られた。Bの項目中にも適切でないと答えた学習者がいたが、分析対象外としている。理由は「聞き手」評価項目が適切であったか尋ねた質問に対し、「よくできた」「難しかった」など、テストでの学習者自身のパフォーマンスを振り返る回答を行っている

ためである。続く研究課題2では、なぜ適切、あるいは適切でないと感じたかについて詳述する。

8.2 研究課題2

研究課題1で得られた結果を受け、研究課題2では、どのような理由から適切、または適切でないと感じていたかについて分析を行った。その結果、表3のようにa~jの合わせて10の理由に分類された。表中の人数は全て延べ数である。また、人数欄の括弧には、その項目中「適切でない」と答えた人数を示した。

表3 「聞き手」評価項目が適切だったと思う理由

A. 開始、終了、相手の話の促しなど、インタビューの進行ができたか	人数
a インタビュアーの役割	10
b 会話能力の伸長に必要	11
c 実社会での有用性	2
その他(適切にパフォーマンスできたか等)	8
無回答	4
合計	35人
B. 相づちを打ったり、理解を示すことができたか	人数
d インタビュアーの役割	9
e 会話能力の伸長に必要	9
f 聞いていることの明示が必要	7
その他(適切にパフォーマンスできたか等)	10(3)
無回答	2
合計	37人
C. 言い換えて説明したり、聞き返したりできたか	人数
g インタビュアーの役割	6
h 会話能力の伸長に必要	6
i 正確・確実な理解のために必要	8
j 必要なし	6(5)
その他(適切にパフォーマンスできたか等)	10
無回答	1
合計	37人

まず、Aの項目では、「a. インタビュアーの役割(10人)」「b. 会話能力の伸長に必要(11人)」「c. 実社会での有用性(2人)」という3つの要素が抽出され、表面妥当性を支持する結果が得られた。以下、学習者の自由記述例を示しながら考察する。(括弧内の数字は、学習者番号)

「a. インタビュアーの役割」自由記述例

・インタビュアーはインタビューを主導する能力が当然必要だ。だから、会話が続くようにしなければならない義務がある(23 番)

「b. 会話能力の伸長に必要」自由記述例

・能動的に話を進行する能力を育てることができる(15 番)

「c. 実社会での有用性」自由記述例

・インタビューの進行を通して質問の仕方(私的ではなく公的に)などの練習ができる(25 番)
・社会に出てもグループの代表になったらこのような能力は必要だからいい練習になったと思う(32 番)

例に示すように、学習者はインタビューの進行をインタビュアーが当然果たすべき役割であると認識していた。この評価項目達成を目指すことで、能動的に会話を進める力も身に付き、会話能力の向上につながることで、さらに、インタビューを実社会で運用する場合を想定し、公的な場でインタビュアーの役割を担うよい練習になるという点から、この項目が設定されることは適切であると考えられていた。

このように、この項目ではインタビューという課題の特性からそれを当然の役割であると考えたり、将来起こりうる実際のコミュニケーション場面と関連付けて、インタビュー進行の技能が役に立つと意義付けされていた。

次に、B の項目では、「d. インタビュアーの役割(9 人)」、「e. 会話能力の伸長に必要(9 人)」、「f. 聞いていることの明示が必要(7 人)」という 3 つの要素が抽出され、同じく表面妥当性が支持されていた。

「d. インタビュアーの役割」自由記述例

・インタビュアーはあるテーマについて聞いて、情報を集めるものなので、聞き手の役割は重要だ(26 番)
・インタビューを効果的にするためには相手が話しやすいようにしてあげなければならないから(10 番)
・相手の話をもっと引き出すのに必要(9 番)

「e. 会話能力の伸長に必要」自由記述例

・相づち、理解を示す能力は会話や聞き取りの能力を伸ばせるより重要な能力だと思う(19 番)
・日本語会話という科目にふさわしく、相づちを打つことを重要視して話す日本人の特性に接することができる機会になる(30 番)

「f. 聞いているということの明示が必要」自由記述例

・相づちを打ったり、理解をしているのか表現をしなければ、相手の話を注意深く聞いているのかわからないから(1 番)
・聞き手が話す内容を理解したかを話し手が把握できなければ対話がスムーズに進行しないから(4 番)

この項目でも、相づちを打ったり理解を示すという聞き手の役割は、あるテーマについて聞いたり情報を集めたりするインタビュアーの果たすべき役

割であると認識されていた。これに加えて特徴的なのは、インタビュー内容を充実させるために、相手が話しやすいよう配慮することが大切だと認識している点である。このような聞き手の働きは、対話相手の話しやすさに直接関係しており、相手の話を豊かに引き出すためにも、話し手だけでなく聞き手の役割を学ぶことが必要だという気付きがあった。また、相づちの打ち方や頻度などは言語によって差があり、韓国人日本語学習者と日本語母語話者では、相づちの頻度や相づちに対する価値観に相違が多いことが指摘されている(任・李 1995, 内藤 2003 等)。相づち、理解を示せる能力が、日本語会話能力の向上に結びつくという点からも、この評価項目の設定が適切であると認識されていた。

以上のように、B の項目では、話を聞いているという合図を相手に送ることが聞き手の役割として大切であることや、対話がスムーズに進行するには、聞き手が積極的に会話参加する必要があることへの気付きが見られた。

最後に、C の項目は、「g. インタビュアーの役割(6 人)」、「h. 会話能力の伸長に必要(6 人)」、「i. 正確・確実な理解のために必要(8 人)」という理由から表面妥当性が支持された一方、「j. 必要なし(6 人)」と表面妥当性を支持しない記述も見られた。

「g. インタビュアーの役割」自由記述例

・インタビューだったらインタビューを受ける人の話をできるだけ詳しく伝えなければならないので、その面からこの項目は必要だったと思う(10 番)

「h. 会話能力の伸長に必要」自由記述例

・実際、日本人ではない以上、100%理解できない状況があるだろう。聞き返しの練習は必須だと思う(22 番)

「i. 正確・確実な理解のために必要」自由記述例

・会話なのだから自分が分からない内容が当然出てくる可能性があり、それを明確に理解しなければならない(23 番)
・分からない単語は、やさしい単語で説明しなおしてもらえるから(29 番)

「j. 必要なし」自由記述例

・相手の話を全部理解できた場合は、聞き返す必要はないから(4 番)
・私にとって一番難しい部分でした。相手の話を全部理解している状況で無理に意識して聞き返す必要はないと思います(27 番)
・これは B(相づちを打つ、理解を示す)と似ている項目のようだから、強いてもう一つの項目がなければならぬのかわからない(8 番)
・ただ聞いていれば話を聞いていることが確認できるから(31 番)

例のように、実社会でのインタビュアーであれば、話し手の話を正確に理解し正しく伝えなければならないという役割があり、それが C の項目における表面妥当性支持の理由として挙がっていた。また、実際のコミュニケーション場面で、学習者は相手の話が 100%は理解できないという状況もある。そのため、聞き返しなどを練習し、明確に理解できるようにする必要があると認識されていた。

テスト実施にあたって、ペア・インタビューテストでは相手とのレベル差がどうしても生じる。ペアでレベル差が小さいほうが良いのか、あっても問題にならないのかは、今後さらなる検討が必要である。しかし、ここでは少なくとも言い換えや聞き返しを評価項目に入れることによって、学習者は相手にわからない箇所をわからないと安心して言うことができ、しかもそれがパフォーマンスとして評価してもらえるという利点があると言えよう。

さて、本研究で表面妥当性を支持しないという記述が見られたのが C の項目であった。理由は大きく 3 つに分けられる。1 つめに、聞き返しの必要性である。対象としたクラスには、日本で育った帰国子女や交換留学を終えた学習者等が数名含まれていた。このような学習者は、他の学習者より未知語などに遭遇する可能性が低く、聞き返しをする必然性が薄いと予想される。ところが、「必要なし」と答えたのはむしろ中間層の学習者であった。4 番や 27 番の学習者に、今回のテストでたまたま未知語などがなかったことも考えられるが、聞き返しの必要性に対する認識と、学習者レベルとがあまり関係しない可能性もある。むしろ、自然なコミュニケーションを意識すると、する必要のないパフォーマンスを評価項目達成のためにあえて行わなければならないという矛盾に学習者自身が気づき、問題提起する記述だと解釈できるだろう。

2 つめに、B と C の評価項目は類似しており、これらの項目が別々に立てられている必要性が感じられないという意見があった。教師側は、B は「聞き手として、相手の話に対して適切に反応できるか」を測り、C は「相手の話を正しく理解できるよう、ストラテジーを適切に使用できるか」を測る項目として設定していたが、その差が学習者によく伝わっていなかったことがわかった。表面妥当性は学習者が直感的に判断するものであり、それを確保するためには学習者自身が評価項目について明確に理解していることが絶対条件だと言える。

3 つめに、言語化して表現するのが理解を示す唯一の方法ではないという意見があった。例えば、頷くなどの非言語的な方法によっても相手が理解しているかの確認はできる。そのため、あえて言語化して表現しなければならないことに対して、否定的に受け止められたと考えられる。

以上から、「言い換え」「聞き返し」などのストラテジー能力を、限られた試験時間内の口頭パフォーマンスに現れたかどうかで評価することには限界があることが示唆された。このため、改善策として、現行の評価項目「言い換えて説明したり、聞き返したりできたか」を、「相手の話を正しく理解できるよう、ストラテジーを適切に使用できたか」など明確な書き方に改善する。もしくは、さらに具体的に①「聞き返すべき重要なポイントは逃さず聞き返すことができた。または、聞き返す必要がなかった」、②「言い換えて説明ができた。または、言い換える必要がなかった」のように 2 つに分けて説明することも効果的であろう。これにより、教師側が意図する評価項目内容が学習者に伝わり易くなると同時に、必要性や言語化の問題にも対処することができる。

9. まとめと今後の課題

本研究では、ペア・インタビューテストを実施し、「聞き手」評価項目の表面妥当性を検討した。その結果、教師が設定した「聞き手」評価項目を多くの学習者が適切だと感じており、表面妥当性が支持されたことがわかった。

評価項目ごとに表面妥当性を支持する理由には違いがあったが、まとめると、①インタビューという課題の特性から、インタビュアーの役割(会話の進行、インタビュー内容の充実や相手への配慮、正確な理解と発信など)を果たすのは当然、②インタビュアーに必要とされる技能習得(相づち、言い換え、聞き返しなど)は、日本語会話能力伸長にも有用、③対話のスムーズな進行には、聞き手の積極的な会話参加が必要、などが挙がっていた。学習者はこれまでの会話テストで、話し手としてのパフォーマンスを評価されることが多かったと考えられるが、今回のテストで聞き手の役割が評価項目に入ったことによって、聞くことも含めて会話能力に必要だと認識されていた。一方、問題点としては、C の評価項目「言い換えて説明したり、聞き返したりできたか」で、聞き返しの必要性、評価項目間の類似性、

ストラテジー能力を言語化する必要性の3つが浮かび上がった。評価項目を学習者にわかりやすく納得のいくものに改善していくことが肝要である。

以上のように、学習者は学習目標として掲げた、「実際のコミュニケーションを意識しながら、よい『話し手』と『聞き手』の役割について学ぶこと」を目指し、課題における聞き手の役割の重要性をよく理解していた。ただし、学習者の中にはインタビューそのものができるようになることを目標と考え、授業やテストに臨んでいたと考えられる認識もあった。インタビューはあくまで形式であり、インタビューの形式を利用して、自ら積極的に会話を進める力をつけてもらうことを意図していた教師と、学習者の認識とは隔たりがある。学習者には現実の会話場面をより意識してもらうため、教師側がどのような意図でインタビュー活動を取り入れているのか明示的に説明したり、課題の出し方や練習の仕方をより現実に近いものに改善するなどの工夫が必要であろう。本研究では「聞き手」評価項目の表面妥当性について検討したが、今後はペアの組み合わせやテーマ設定等、より多角的に検討を続け、ペア・インタビューテストの表面妥当性についてさらに明らかにしていきたい。

謝辞：本研究の調査にご協力くださった筆者の元勤務校である韓国外国語大学の学生のみなさんと中村八重先生、そして貴重なアドバイスをくださった査読者の先生方に心よりお礼申し上げます。

注

1. 本研究での「聞き手」と「インタビュアー」、「話し手」と「インタビュイー」はそれぞれ同じ意味として扱っている。
2. 内的妥当性(internal validity)は「認識されたテストの内容や、認識されたテストの影響」であり、外的妥当性(external validity)は「ある特定のテストから得られた言語能力とテスト以外の方法で得られた能力とを比較検証」する方法である(オルダーソン他 2010, 156-157)。
3. 内容妥当性検証法(content validation)は「テストの専門家、言語学者、各科目の専門家などの判断」をデータとする。検証は「専門家—判断が自分のものと異なっているとしてもそちらを信頼できる人たち—の判断をもとに行う」。また、応答妥当性検証法(response validation)は「テスト受験者が自己報告や自己観察などのさまざまな定性的(qualitative)方法を使って得られたデータをもとにして、テストにどのようなプロセスで解答しているか、解答の理由は何かなどを判断する」ものである(オルダーソン他 2010, 157-159)。

4. 学年は2年生23人、3年生3人、4年生8人、1年生1人である。
5. 口頭フィードバックの内容は、よくできたところや今後の課題である。ただし時間の制約があったため、翌週個々に筆記によるフィードバックも行った。
6. 「あるテスト項目は別の項目よりも重要であり、それゆえより多くの得点を与えたいという場合がある。付加的に得点を与える場合、これを重み付け (weighting) という」(オルダーソン他 2010: 139)。

参照文献

- 任榮哲・李先敏(1995)「あいづち行動における価値観の韓日比較」『日本語教育論集 世界の日本語教育』5, 239-251.
- 久野由宇子・斉木ゆかり(2009)「学習者同士の対面テスト」の試み『東海大学紀要留学学習者教育センター』29, 63-77.
- 国際交流基金(2007)『国際交流基金日本語教授法シリーズ 第6巻「話すことを教える」』ひつじ書房
- 徳間望(2009)「学習者が相互に「インタビュー進行」を担う会話テストの試み」『言語文化と日本語教育』38, 38-45.
- 内藤真理子(2003)「あいづちのスピーチレベルとそのシフトについて—日本語母語話者と韓国人学習者の相違—」『日本語教育論集 世界の日本語教育』13, 109-125.
- 堀川有美(2006)「日本語会話テストにおいてテスト形式が受験者の発話に与える影響—グループ形式とインタビュー形式の比較—」『人間文化論叢』9, 255-264.
- 堀川有美・徳間望(2008)「日本語グループ・オーラル・テストの学習への波及効果に関する予備的研究—教室内テストの場合—」『JLTA Journal 研究紀要』11, 139-155.
- 堀口純子(1988)「コミュニケーションにおける聞き手の言語行動」『日本語教育』64, 13-26.
- Brooks, L. (2009) Interacting in pairs in a test of oral proficiency: Co-constructing a better performance. *Language Testing*, 26(3), 341-366.
- Alderson, J. C; Clapham, C. & Wall, D. (1995) *Language Test Construction and Evaluation*. Cambridge University Press. (チャールズ・オルダーソン他(2010)渡部良典(編訳)『言語テストの作成と評価—あたらしい外国語評価のために』春風社)
- McNamara, T. (2000) *Language Testing*. Oxford Introductions to Language Study. (ティム・マクナマラ(2004)伊東祐郎他(監訳)『言語テストング概論』スリーエーネットワーク)
- Taylor, L. & Wigglesworth, G. (2009) Are two heads better than one? Pair work in L2 assessment contexts. *Language Testing*, 26(3), 325-339.

稿末資料1 評定尺度

評価項目 得点	聞き手 (インタビュアー)		話し手・聞き手 共通	話し手(インタビュイー)			
	A	B	C	D	E	F	G
	開始／終了／ 相手の話の促 し	相づちを打つ ／理解を示す	言い換え／聞き 返し	発音	語彙	文法	流暢さ
4点	インタビュー 開始や終了が できる。相手 の話の流れに 合わせ、相手 に話を促すこ とができる。	相手の発話に 合わせ、適切 に相づちを打 ったり、感想 を示している。	相手の発話に合 わせ、適切な表 現で自分の理解 が正しいかどう か確かめたり、 相手の話が理解 できないとき、 うまく聞き返す ことができる。	全体的に発音 が明瞭で 聞き取りやす く、理解を妨げ ない。	話題内容や文 脈に合わせて、 様々な語彙を使 用している。	発話内容に合 わせて、文法を 効果的に使用し ている。軽微な誤 用が見られるこ ともあるが、意 味の理解を全く 妨げない。	全体的に発話 がスムーズで、 発話が中断する 際にはフィラー を適切に使用し ている。
3点	インタビュー 開始や終了は できるが、相 手に話を促す ことがほとん どない。	相手の発話に 合わせ、相づ ちを打つ、ま たは感想を示 している。	相手の発話に合 わせ、自分の理 解が正しいかど うか確かめたり 、相手の話が理 解できないとき 、うまく聞き返 すことができる。	発音が不明 瞭な部分も あるが、会 話の流れの中 で理解を妨げ ない。	難しい語彙で は不適切な使 用が見られる が、多くの語 彙は適切に使 用している。	基本的にシン 플な文法を 使用している。 誤用があるが、 意味の理解を 妨げない。	発話途中でポ ーズが入ったり 、言いよど んだりするこ とがあるが、 コミュニケーション の流れを阻害し ない。
2点	インタビュー 開始や終了は できるが、相 手の話を促す ことがない。	相手の発話に 合わせ、相づ ちを打ったり 、感想を示す ことがほとん どない。	相手の話の流れ は考慮してい ないが、自分の 理解が正しいか どうか確かめ たり、相手の話 が理解できな いとき聞き返 すことができる。	発音の不明 瞭さが目立 ち、会話進 行中に理解を 妨げる。	基礎的な語彙 を中心に発話 している。	シンプルな文 法のみ用いて いる。意味の 理解を妨げる ような誤用が 目立つ。	不自然に長い ポーズや言い よどみ等によ って、コミュニ ケーションが阻 害されることが ある。
1点	インタビュー 開始や終了が できない。相 手の話を促す ことがない。	相手の発話に 合わせ、相づ ちを打ったり 、感想を示す ことがない。	自分の理解が正 しいかどうか確 かめたり、相手 の話が理解でき ないとき、聞き 返すことが全 くない。	全体的に発音 が不明瞭で、 理解が 難しい。	少数の基礎的 な語彙のみを 使用している。	ごく基礎的な 文法知識をも っているが、 運用できない。	ポーズや言い よどみが極端 に多く、発話 が途中で止ま ってしまう。

*A, B, Cは素点を3倍する。

☆このアンケートは会話テストの改善のために行うものです。成績には一切関係しません。

☆韓国語で教えてください。

1. 評価項目とテスト準備についてお聞きします。

<評価項目>の表を見て教えてください。評価項目の設定は適切だったと思いますか。

<評価項目>

聞き手(インタビュアー)	
A	インタビューの進行ができたか(開始、終了、相手の話の促しなど)
B	聞き手の役割ができたか(相づちを打つ、理解を示すなど)
話し手・聞き手共通	
C	自分の理解が正しいかどうか確かめたり、相手の話が理解できないときうまく聞き返すことができたか(言い換えて説明する/聞き返しするなど)

A. インタビューの進行ができたか < 適切だった/適切でなかった >

→理由をできるだけ詳しく教えてください ()

B. 聞き手の役割ができたか < 適切だった/適切でなかった >

→理由をできるだけ詳しく教えてください ()

C. 自分の理解が正しいかどうか確かめたり、相手の話が理解できないときうまく聞き返すことができたか

< 適切だった/適切でなかった >

→理由をできるだけ詳しく教えてください ()

Face Validity of Interviewer's Criteria in Paired Interview Test

TOKUMA Nozomi, MORIKAWA Maiko

Abstract

Face validity judges usefulness of a test from the students' impressions about it. As a test influences students' motivation, maintaining face validity is of importance. The purpose of this research is to study the face validity of the criteria proposed by the researchers criteria in paired interview tests. Data for this research was collected from 35 Korean undergraduate students who took part in paired interview tests. The objective of the research was to (a) see if the students found the criteria appropriate, and (b) to know the reasons behind their judgment. Results of the current study showed that: (c) face validity of the criteria suggested by the researchers was approved and (d) there were three primary reasons why the students approved of the criteria – firstly, it is fairly reasonable to play the part of an interviewer in an interview test, secondly, students' verbal ability improved by playing the part of the interviewer and thirdly, positive participation of the interviewer is essential to have a smooth conversation.

【Keywords】 between students, paired interview test, interviewer's criteria, face validity

(TOKUMA: The Japan Foundation, New Delhi)

(MORIKAWA: The Yamasa Institute)