

ロールプレイを用いた口頭試験の評価 —「産出」と「やりとり」の視点から—

倉沢 郁子

要旨

本稿では、初級後半の日本語の授業で行っているロールプレイを用いた口頭試験のルーブリックの改定を試みることを目的に、現行の評価について振り返り、JF日本語教育スタンダードの枠組みから評価基準を考察した。試験の内容やタスクの難易度には妥当性があると考えられるものの、課題の提示の仕方やタスクの性質を考慮にいれ、「やりとり」、そして「産出」の二つの観点からの評価基準を設ける必要があることがわかった。

【キーワード】 口頭能力の評価、JF 日本語教育スタンダード、タスクの性質、「やりとり」、「産出」

1. はじめに

以前に学習項目を教室の外での実際使用へとつなげる学習活動の一環として、ロールプレイカードを指導に取り入れた実践報告をした(鹿浦ら 2018)。学習者がこの教室活動になれていることもあり、口頭試験にもロールプレイカードを用いたことも先の実践報告で述べたが、どのように学習者の能力を適切に評価できるのかという点において多くの課題があった。産出における文法的な正確さや流暢さ等を評価項目に含めていたが、タスク遂行の視点から評価を行った場合、言語的挫折が見られたとしても、タスクは達成されたケースがほとんどである。授業では正確さやポライトネスについての概念を説明することはあっても、タスク達成を強調することはあまりない。文法シラバスでタスク達成を目標としてしまうと、学習項目の練習にならなくなってしまう可能性があるためである。しかし、接触場面を考慮に入れるのであれば、カリキュラムを考える上でタスク達成は大きなポイントにもなる。

また、もし「入り口」であるコースの目標設定にタスク達成という視点を含めるとなると、「出口」である評価にも反映させる必要がある。では、どのような評価を行えばいいのだろうか。

本稿では、ヨーロッパ言語共通参照枠 (CEFR - Common European Framework of Reference for Languages 以下 CEFR) をもとにし、コースデザインや授業設計、評価を考えるために利用できる枠組みとして開発された「JF 日本語教育スタンダード」(国際交流基金 2017、以下 JFS) の視点から、評価基準について考察を加える。JFS は「相互理解のための日本語」を理念とし、そのために「何ができるかという『課題遂行能力』とさまざまな文化に触れることでいかに他者の文化を理解し尊重するかという『異文化理解能力』」が必要であるとしている。コミュニケーションのための日本語を目指す授業の口頭試験で、文法シラバスとタスクを掛け合わせたタスク(課題) 遂行をどのように評価したらいいのか、JFS の「課題遂行能力」、特に「コミュニケーション言語活動」の視点から振り返り、考察を加える。

2. 口頭試験の概要

対象となる口頭試験は、筆者が関西外国語大学留学生別課で担当しているコースの期末口頭試験である。「初級日本語 げんき II」(坂野ら 2011) の第 18 課～第 23 課までをカバーし、初級後半、CEFR の A2～B1 と考えられるレベルである(表 1 参照)。

表 1 「学習、教授、評価のためのヨーロッパ共通参照枠」(CEFR 共通参照レベル)

全体的な尺度

B1	<ul style="list-style-type: none"> ▪ 仕事、学校、娯楽で普段出会うような身近な話題について、標準的な話し方であれば主要点を理解することができる。 ▪ その言葉が話されている地域を旅行しているときに起こりそうな、たいていの事態に対応することができる。 ▪ 身近で個人的にも関心のある話題について、単純な方法で結びつけられた、脈絡のあるテキストを作ることができる。経験、出来事、夢、希望、野心を説明し、意見や計画の理由、説明をうまくすることができる。
A2	<ul style="list-style-type: none"> ▪ ごく基本的な個人的情報や家族情報、買い物、近所、仕事など、直接的関係がある領域に関する、よく使われる文や表現が理解できる。 ▪ 簡単で日常的な範囲なら、身近で日常の事柄についての情報交換に応ずることができる。 ▪ 自分の背景や身の回りの状況や、直接的な必要性のある領域の事柄を簡単な言葉で説明できる。

試験にかかる時間は一人につき約7分程度で、全体の流れは、①ウォームアップ、②4つのロールプレイ、③意見陳述のタスク、最後に④クールダウンという形で進む。②のロールプレイのカードは6つあり、その中の3つは全員が必ず行い、もう一つは残りの3つの中から一つ各自で選び、コース担当教師（筆者）と会話をするという形にした。この口頭試験ではロールカードを先に学生に共有している。それは、学習不安が出やすいテスト場面（Horwitzら 1991, 倉沢 2002）において、事前に準備・練習ができるようにということが理由である。③の意見陳述では、近年外で遊ばない子供が増えているが、それについてどう思うか、意見を述べる課題を出した。この話題にした理由は、抽象度が高くなく、学習者自身の経験からも意見が述べやすいと考えたからである。

この試験は成績全体の30%にあたる期末試験の内訳の30%、つまり成績全体の9%に値する内容である。評価は①、③、④をPart A、③をPart Bとして各パート15点、合計30点とした。評価項目は、1) 表現のバラエティ（3点）、2) 正確さ（3点）、3) 流暢さと発音（2点）、4) コミュニケーション・スキル（2点）、5) タスク完了（タスクにつき各2点）である。

1) 表現のバラエティや2) 正確さは、タスク遂行にあたって、試験中にどのぐらい新出文法を用いたか、間違いがどのぐらいあったのか、3) 流暢さや発音は、言い淀みがないか、もしくは違う音に聞こえるために意味まで変わってきてしまうような極端なケースがいくつあったか、4) は日本語で会話を遂行できたか、フィラーや相槌なども評価の視点とした。5) タスク完了は、意味のあるやりとりの結果、タスクが完了できたかどうかを見た。

3. 信頼性と妥当性

3.1. 本試験の信頼性について

口頭試験は、評定者によって評価の結果が異なる可能性が大きい主観テストに分類され、その信頼性や妥当性を高めることは非常に重要なことである。評価の信頼性を高める対策として、小野塚ら（2008）は、1) 評価項目及び評価基準を用意する、2) すべての評定者が評価項目及び評価基準について共通理解を得る、3) 練習の段階で評価結果が評定者によって異なる場合があったら、評価項目及び評価基準の理解を深め、再度採点の練習を行うという3点を挙げている。この口頭

試験における評価項目・評価基準は先述の通りであり、学習者と試験に先立って評価基準を共有している。また、筆者は一人で担当しているが、評定者内信頼性を高めることが今後必要である。

3.2 内容の妥当性について

次に試験の妥当性であるが、内容に関しては、今回の試験全体で評価の対象となっているものは第 18～23 課までの文法や語彙、表現で、その項目が使用できるロールカードを考えた。Part B の意見陳述のタスクも、「げんき 2」第 22 課で「賛成」、「反対」という単語が導入され、また同課には意見を述べる練習などもあるため、文法シラバスから離れたタスクであるが授業活動と一致している。

3.3 「意見陳述」のタスクの難易度

では「意見陳述」というタスクの難易度の妥当性はどうか。「意見を述べる」という言語活動を JFS で見てみると、表 2 ⁽¹⁾ のようなものがある。

表 2 「意見を述べる」に関する言語活動

種別 整番 種類 言語活動 レベル	カテゴリー	第一 トピック	Can-do 本文
CEFR 238 活動 やりとり B2	共同作業中 にやりとりを する		他人に仲間に入るように誘ったり、意見を述べるように促したりすることによって、作業を先に進めることに貢献できる。
JF 033 活動 産出 B1	論述する	学校と 教育	大学のゼミのディスカッションで、簡単な説明や理由を交えて自分の意見を述べるができる。
JF 034 活動 産出 B1	論述する	社会	増税など、自分の生活に密接にかかわる政策について、友人に簡単な説明や理由を交えて自分の意見を述べるができる。
CEFR 235 活動 やりとり A2.2	フォーマルな 場面で議論 する		直接自分に向けられた質問ならば、実際の問題についての関連情報をやり取りし、自分の意見を示すことができるが、自分の意見を述べる際には、 <u>人の助けを借り、必要に応じて鍵となるポイントを繰り返してもらわねばならない。</u>

CEFR 235 は、A2.2 のレベルで「やりとり」の言語活動に分類される Can-do であり、タスクの難易度という観点からも妥当性があると考えられる。しかし、ここで

注意を向けたいのは、このタスクは「やりとり」であること、そして「自分の意見を述べる際には、人の助けを借り、必要に応じて鍵となるポイントを繰り返してもらわねばならない」という点である。評価基準を考える際に考慮に入れるポイントとするべきところである。

4 JFS からの評価基準の再考

4.1 「産出」と「やりとり」の視点から

では、実際にどのように評価基準を改善していけるだろうか。まず、表2にある言語活動に注目すると、この4つの中に「産出」と「やりとり」が混在している。JFS の言語活動には、「受容」「産出」「やりとり」の3種類があり、その性質から Part A の①、③、④は「やりとり」の性質に近く、Part B の②ロールプレイは「やりとり」を想定していたが、表3⁽²⁾にある特徴にあるように「練習しておけば」という「産出」の「条件」の特徴があり、このことから評価の観点や評価項目を Part A と Part B で変える必要があると考えられる。具体的には Part A では、タスク遂行を大きなポイントとし、Part B では、正確さに重きをおいた評価に変更したほうが妥当性が高くなるのではないだろうか。

表3 「産出」と「やりとり（相互行為）」の「条件」「話題・場面」特徴の比較

レベル	やりとり		産出	
	条件	話題・場面	条件	話題・場面
B1	<ul style="list-style-type: none"> ▪ 時には言いたいことは言えないこともあるが ▪ ディベートに参加するのは難しいが ▪ 時には繰り返しを求めるところもあるが ▪ 相手が標準的な言葉遣いで発言してくれれば ▪ 話が自分に向けられていれば 	<ul style="list-style-type: none"> ▪ 一般的な話題 ▪ 音楽や映画のような抽象的な文化的話題 ▪ 興味のある話題 ▪ 身近で個人的関心のある事柄 ▪ 日常生活に関連する話題(家族、趣味、仕事、旅行) ▪ 自分の専門分野に関する話題 ▪ 身近な話題 ▪ 具体的・抽象的な話題 ▪ あまり日常的では起きない状況(気に入らなかった品を返品するなど) ▪ 旅行中に起きそうなこと 	<ul style="list-style-type: none"> ▪ <u>練習しておけば</u> ▪ アクセントとイントネーションにはかなり耳慣れない部分もあるが 	<ul style="list-style-type: none"> ▪ 自分の専門範囲の日常的・非日常的な事柄 ▪ 日常的な事柄 ▪ 意見、計画、行動 ▪ 自分の関心のあるさまざまな話題 ▪ 現実や想像上の出来事、経験 ▪ 事故などの予測不能の出来事 ▪ 夢や希望

レベル	やりとり		産出	
	条件	話題・場面	条件	話題・場面
A2	<ul style="list-style-type: none"> 必要がある場合に相手が助けてくれれば 議論がゆっくりとはっきりとなされれば はっきり、ゆっくりと、自分に直接向けられた発話ならば 必要な場合に鍵となるポイントを繰り返し返してもらえらば 地図や図を参照しながら 時々繰り返しや言い換えを求めることが許されるのであれば 	<ul style="list-style-type: none"> 自分の周りで議論されている話題 旅行、宿泊、食事、買い物のようなミア日の生活での普通の状況 日常の課題に関して 仕事や自由時間に関する身近な毎日の事柄 直接必要なこと 予測可能な日常の状況で 身近な話題 興味ある話題 	<ul style="list-style-type: none"> 聞き手が集中して聞いてくれれば、練習した上で 	<ul style="list-style-type: none"> 自分の毎日の生活に直接関連のある話題 計画、準備、習慣、日々の仕事、過去の活動や個人の経験 家族、住居環境、学歴、現在やごく最近までしていた仕事 人物や生活、職場環境、日課、好き嫌いなど 予測可能で身近な内容の事柄

4.2 「意見陳述」の評価例

③の「意見陳述」では、何について賛成か反対かを述べるか、まず状況説明や意見自体が理解できないケースがある。また、以下の試験の会話例では、賛成・反対を述べることは問題がなく、文法上の間違いがあるにせよ、理由を述べてきているため、タスク達成とした。しかし、S4 をみると、既習単語が思い出せず、リソースである評価者に助けを求めている。

口頭試験の会話例

Teacher (T)1: ちょっと意見を聞きたいんですけど、最近ね、こういう意見があります。子供が外で遊ぶのは危ないから、外に出ないほうがいいです…

Student (S)1: すみません、もう一度いってくれませんか。

T2: はい。外で遊ぶのはとっても危ないから、外で遊ばないほうがいいです。うちにずっといなさい、うちにいなさい。そういう意見があります。

S2: えっと、私の意見？

T3: うん、賛成ですか、反対ですか。

S3: えーとー、反対です。

T4: うん

S4: 私は、私は、えーと、こどもが、こどもが、えーと、外で遊び、遊びの方がいいと思います。運動は、運動は、とても…あ、あ、日本語で、すみません、日本語で **Important** は何と言いますか。

T5: 大切

S5: 大切、はい、運動は、運動はとても大切です。

T6: うん、うん、だから、外で遊んだほうがいいと思うんですね。

S7: はい、えーっと、えーっと、あー、あー、あー、う、家だけに、いる、いる、いるは、えーと、元気、子どもは元気にな、な、なり、なり、ならないです。なりません。あ、なれません。

T8: だから、外で遊ぶと元気になる、んですね。

S8: はい。

表 3 によると、「やりとり」である意見陳述のタスクでは「時には繰り返しを求めることもある (B1)」、「必要がある場合に相手が助けてくれれば (A2)」、「必要な場合に鍵となるポイントを繰り返してもらえるならば (A2)」の条件の上でコミュニケーションが成り立つとしている。この点は「産出」の言語活動であれば減点対象にすることも考えられるが、助けを求めるストラテジーを活用できている、つまりコミュニケーション・スキルは評価できると解釈した。授業では、正確に文法を使いこなせることを強調していたが、「やり取り」に焦点があるタスク遂行の評価の観点として、この条件をどのように設定するかをもう一度考える必要があることがわかった。

5. まとめと今後の課題

以上、JFS の枠組みを用いて、現行の評価方法を振り返り、どのような評価を行えば、学習者の運用能力を適切に図ることができるのか、評価基準を改定することを目的に考察を加えた。現行のものは、タスクの難しさは妥当性があると考えられるものの、タスクの性質を考慮にいれ、「やりとり」、そして「産出」の二つの観点

からの評価基準を設けることで、学習者の運用能力をもっと適切に測ることができそうだということがわかった。今後の評価に役立てていきたい。

建設的なフィードバックができるようにという意味でも、評価は非常に大きな意味がある。学習者が自律的に日本語学習を進めていけるように、どのような点において練習すればいいのかを知らせる指標にもなると考えられる。

今後の課題としては、「学習者の学習可能性に注目して行うダイナミック・アセスメント」(小村 2009)での評価や、『言語化される前の潜在的なスピーキング能力』を把握することもスピーキング能力の理解には必要」であり、自己評価を言語化以前の潜在的な話す能力と捉えている佐藤ら(2015)を参考に、現在同コースで行っている学習ポートフォリオの自己評価をどのように活用していけるか等を考えながら、学習者が教室という空間を超え、自信をもってコミュニケーションできるように支援していきたい。

注

- (1)表 2 は「みんなの Can-do サイト」(<https://jfstandard.jp/cando/top/ja/render.do>) から抽出したものを紙面に合わせて表を加工したものである。表内の網掛け、点線は筆者による。
- (2)表 3 は『JF 日本語教育スタンダード 【新版】 利用者のためのガイドブック』(国際交流基金 2017) から部分的に抜き出して表にしたものである。

参考文献

- 小野塚若菜、島田めぐみ (2008) 「日本語教師のための Excel でできるテスト分析入門」スリーエーネットワーク
- 倉沢郁子 (2002) 「日本語教育における学習者の不安とその処理」『日本語教育学会秋季大会予稿集』160-165
- 小村親英 (2009) 「日本語口頭試験におけるダイナミック・アセスメントの試み」『関西外国語大学留学生別科日本語教育論集』第 19 号 1-19
- 国際交流基金 (2017) 『JF 日本語教育スタンダード 【新版】 利用者のためのガイドブック』
- 佐藤千秋、関敦彦、伊藤玲子、川口裕司 (2015) 「初級フランス語におけるスピーキング能力可視化の試み - CEFR-J 自己評価、タスク評価、学習ストラテジー調

査一』『ふらんぼー=Flambeau : 東京外国語大学フランス語研究室論集』第 41 号
165-183

鹿浦佳子・倉沢郁子 (2018) 「教室から実際使用へ ～ロールプレイカードを用いた口頭能力育成の試み～」『関西外国語大学留学生別科日本語教育論集』第 27 号
49-58

坂野永理・池田庸子・大野裕・品川恭子・渡嘉敷恭子 (2011) 『初級日本語 げんき II [第 2 版]』The Japan Times

Horwitz, E., Horwitz, B., & Cope, J. (1991) Foreign Language Classroom Anxiety, In Horwitz, E. & Young, D., *Language Anxiety from Theory and Research to Classroom Implication*. New Jersey, USA: Prentice Hall, Inc. 27-36

(kurasawa@kansai-gaidai.ac.jp)