



話しことばの評価研究の動向と課題 : コミュニケーション評価の再考に向けて

著者	篠原 亜紀
雑誌名	言語学論叢 オンライン版
巻	10 (通巻36)
ページ	28-46
発行年	2017-12-09
URL	http://hdl.handle.net/2241/00151282

話しことばの評価研究の動向と課題 —コミュニケーション評価の再考に向けて—

篠原 亜紀

要 旨

本稿では、話しことばによるコミュニケーションがどのように捉えられ、評価されてきたかを概観し、今後必要とされる評価研究について考える。

まず、言語教育において、コミュニケーションがこれまでどのように捉えられてきたかを述べる。そして、現在、日本語教育現場において話しことばがどのような方法で評価されているかを整理する。さらに、これまで行われてきた話しことばの評価研究について、「だれが」「どのような発話の」「何を」「どのように」評価するかという視点から先行研究を概観する。そして、社会の変化や、今後の社会において必要となるコミュニケーションを考慮しつつ、話しことばによるコミュニケーションを評価するためのよりよい方法を考えるためにどのような評価研究が必要となるかを考える。

キーワード

話しことば コミュニケーション 課題遂行 評価 評価研究

1 はじめに

近年、日本語教育において、「正確さ」だけでなく、「コミュニケーション上の課題が遂行できるかどうか」(以下、課題遂行)が重視されるようになってきた。課題遂行を目指した教育では、言語知識を身につけることそのものを目標とするのではなく、身につけた言語知識を運用してコミュニケーションを遂行することが最終的な目標となる。課題遂行を目標としたコミュニケーションをどのように教えるかについては、教育現場において様々な試みや実践が行われ、教材開発も進められている。しかし、上述したような新しい言語教育観に準じた評価が十分に行われているとは言い難い。

本稿では、まず、これまで言語教育においてコミュニケーションがどのように捉えられてきたかを述べる。次に、現在、日本語教育現場において話しことばがどのような方法で評価されているかを整理する。さらに、これまで行われてきた話しことばの評価研究について、「だれが」「どのような発話の」「何を」「どのように」評価するかという視点で概観し、今後、どのような評価研究が必要となるかを考える。

2 コミュニケーションとは

2.1 コミュニケーションの特徴

国際交流基金 (2007) は、話し手と聞き手のいるコミュニケーションの特徴を分析し、現実の自然なコミュニケーションは、「目的」(話を始める動機・きっかけ)、「情報差」(話し手と聞き手の間のギャップ)、「選択権」(自分で内容や表現形式を選ぶ)、「反応」(相手の反応を見ながら会話を進める)の 4 つの要素から成立していると述べている。蒲谷 (2013) は、コミュニケーションの枠組みとして、「場面 (場+人間関係)」「意識」「内容」「形式」(いつどこで、だれとだれが、何のために、何を、どのように、コミュニケーションするのか) を挙げ、これらが常に連動してコミュニケーション行為が成立していると述べている。

では、コミュニケーションを遂行するためにはどのような能力が必要か、その能力がこれまでどのように捉えられ、位置づけられてきたかを以下に整理する。

2.2 コミュニケーション能力

外国語教育において、1960 年代ごろまでは、言語能力とは文法知識であると考えられていた。1970 年代以降になって、言語能力が広く捉えられるようになり、言語を「使う」能力についても注目されるようになった。そのような中で、Hymes (1972) は、言語を運用する能力を「コミュニケーション能力」(communicative competence) と名づけた。Canale & Swain (1980) は、そのコミュニケーション能力の要素として、文法能力 (grammatical competence)、社会言語能力 (sociolinguistic competence)、ストラテジー能力 (strategic competence) の 3 つを提示している。その後、Canale (1983) によって、談話能力 (discourse competence) が追加された。国際交流基金 (2007) は、これらの 4 要素をコミュニケーション能力とし、話すことの教育において、これらをバランスよく育成するような授業デザインが必要だと述べている。

表 1 国際交流基金 (2007) によるコミュニケーション能力の定義

	能力の定義	関連する技能や知識
文法能力	言語を文法的に正しく理解して使用する能力	文法規則、語彙、発音、文字 (イントネーション、アクセント)、文字・表記
社会言語能力	相手との関係や場面に応じて、いろいろなルールを守って言語を適切に使用する能力	相手や場面によるスピーチスタイルの選択、話題の選択、非言語行動
ストラテジー能力	コミュニケーションがうまくいかなかったときに、自分や相手の発話をコントロールして修復する能力	自分の発話の調整 (ほかの表現への言い換え、説明、母語使用)、相手への援助要請 (聞き返し)
談話能力	談話を管理し、組み立てることができる能力	会話の開始・継続・終了の仕方、あいづちの打ち方、話題の転換や展開の仕方、発言の順番とり

2001 年には、欧州評議会 (Council of Europe) が「ヨーロッパ言語共通参照枠」“Common European Framework of Reference for Languages” (以下、CEFR) を発表した。CEFR では、行動中心主義を掲げ、言語使用を、「能力 (competences) を方略的 (strategically) に使って遂行する行動 (actions) 」と捉えている (Council of Europe 2004)。

2010 年には、国際交流基金が CEFR を参考に開発した「JF 日本語教育スタンダード」(以下、JF スタンダード) を発表した。JF スタンダードでは、コミュニケーションを「コミュニケーション言語活動」と「コミュニケーション言語能力」のつながりと捉え、1 本の木で表現した (「JF スタンダードの木」)。「JF スタンダードの木」では、言語知識や談話能力などのコミュニケーション言語能力を根の部分に例え、根から幹を伝って枝に伸びた部分を話す、読むなどのコミュニケーション言語活動に例えている。つまり、コミュニケーション言語能力はコミュニケーション言語活動を支えるものとして捉えられている (図 1)。

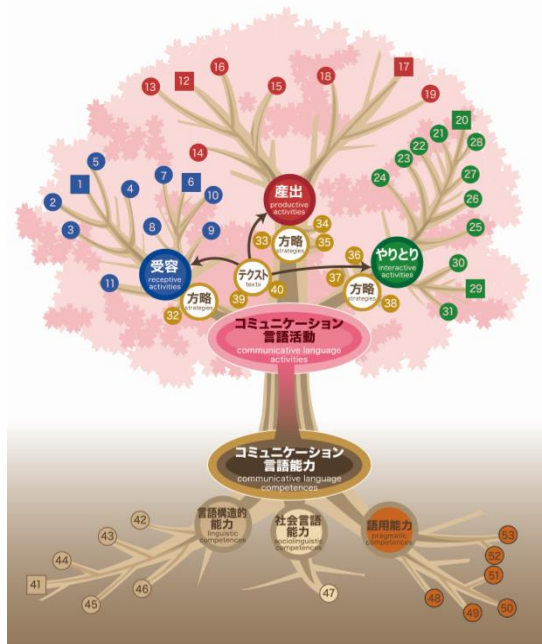


図 1 JF スタンダードの木

コミュニケーション言語活動とコミュニケーション言語能力の構成要素は表 2 の通りである。コミュニケーション言語能力には、大きく分けて、「言語構造的な能力」「社会言語能力」「語用能力」があり、コミュニケーション言語活動には、「受容」「産出」「やりとり」がある。さらに、聞いたことをメモするなど、2 つの活動を同時に行うものなどを「テキスト」としている。言語能力を効果的に使って言語活動を行うための「方略」が「受容」「産出」「やりとり」のそれぞれにあり、コミュニケーション言語能力とコミュニケーション言語活動をつなぐ役割をしている。

表 2 「JF スタンドアードの木」の構成

		活動例		方略
コミュニケーション 言語活動	受容	聞く	指示やアナウンスを聞く、音声メディアを聞く	意図を推測する
		読む	手紙やメールを読む、必要な情報を探し出す	
	産出	話す	経験や物語を語る、講演やプレゼンテーションをする	表現方法を考える 他の方法で補う モニターする
		書く	作文を書く、レポートや記事を書く	
	やりとり	話す	社交的なやりとりをする、店や公共機関でやりとりをする	発言権を取る 展開に協力する 説明を求める
		書く	手紙やメールのやりとりをする、申請書や伝言を書く	
テキスト		メモやノートを取る、要約したり書き写したりする		
コミュニケーション 言語能力	言語構造的な能力		使用語彙領域、語彙の使いこなし、文法的正確さ、音素の把握、正書法の把握	
	社会言語能力		社会言語的な適切さ	
	語用能力	ディスコース能力	柔軟性、発言権、話題の展開、一貫性と結束性	
		機能的な能力	話しことばの流暢さ、叙述の正確さ	

以上、コミュニケーション能力がこれまでどのように捉えられてきたかを整理した。次に、これらのコミュニケーション能力が、これまでどのような方法で評価されてきたを以下に述べる。

3 話しことばの評価

3.1 口頭テスト／会話テスト

口頭運用能力を測るテストとして、教育現場などで使われているものに「ACTFL-OPI (The American Council on the Teaching of Foreign Languages-Oral Proficiency Interview) テスト」(以下、OPI テスト)がある。OPI テストでは、言語を使ってどのようなタスクがどの程度できるかを、テストがインタビューをするという方法で測る。『ACTFL-OPI 試験官養成マニュアル』によれば、OPI テストによる判定尺度のレベルは、超級、上級、中級、初級の4レベルであり、さらに、上級、中級、初級には下位レベル(上、中、下)が設定されている。評価の観点には「総合タスクと機能」「場面／話題」「正確さ」「テキストの型」の4点で、どのような言語活動(総合タスクと機能)を、どのような場面(場面／話題)で、どのように(正確さ)、そして、どのような形式(テキストの型)で遂行できるかを測る。このことから、OPI テストは、言語知識だけでなく、言語知識をどのように運用できるかを測っているといえよう。

また、ヨーロッパでは、CEFR の広まりから、「CEFR 準拠の日本語口頭産出テスト (Oral

Japanese Assessment Europe)」(以下、OJAE) の開発が進められた (OJAE 2010 チーム 2010)。OJAE の評価の指標は、CEFR が「共通参照レベル：話し言葉の質的側面」として提示している 5 領域 (言語表現使用幅、正確さ、流暢性、結束性、やり取り) であり、レベルの決定は、発話がどのような質であったかが重視される。行動中心主義を掲げる CEFR は、言語学習者を社会的成員 (social agents) とみなし、社会において言語を使って何かを遂行することが求められていると考えるが (Council of Europe 2004)、OJAE の指標は、表 2 の「コミュニケーション言語能力」を分析的に測っているにすぎず、それが、「コミュニケーション言語活動」にどのように関わるのか、つまり、評価の指標がコミュニケーションの遂行をどのように支えていると考えて判定するのは疑問である。

一方、2014 年には、「JF 日本語教育スタンダード準拠ロールプレイテスト」(以下、JFS ロールプレイテスト) が開発された。JFS ロールプレイテストでは、「タスクが達成できたか/できなかったか」という言語活動全体から判定し、「文法の正しさ」「語彙の使い方」といった分析的な観点から言語知識を分解して測ることはしない (国際交流基金 2016)。課題遂行のみに着目するという点では、上述した CEFR の行動中心主義の考え方に近いといえるが、OJAE が測っているような質的側面が無視されてしまう危険性がある。たとえば、B1、B2 レベルのテストで、ごく初歩的な語句や表現をつなぎ合わせて何とか課題を達成した場合でも、そのレベルを達成したとテストが判定する可能性もある。

表 3 JFS ロールプレイテストの判定基準

タスクが達成できた		タスクが達成できなかった	
◎	○	△	×
十分に達成できた 難なくタスクをこなしている	何とか達成できた	惜しかったが、達成したとは言えない	全く達成できなかった タスクが難しすぎて全く歯が立たない

また、JFS ロールプレイテストは、OPI テストのような集中的なテストトレーニングを必要としないが (長坂・押尾 2013)、テスト間での判定のズレについても課題が残る。JFS ロールプレイテストの試行段階では、質問の仕方などのテスト行動が判定に影響することが指摘されている (長坂・押尾 2013)。また、鈴木 (2005) は、短時間のテストトレーニングが評価にどのような影響を与えるかを調査したところ、トレーニング後にテストに混乱が生じ、結果的に短時間のトレーニングがマイナスに作用した例を報告している。

3.2 授業における評価

日本語教育において、課題遂行が重視されつつあるが、それにともない、授業内における評価方法も変化してきた。言語知識を問うための筆記テストから、言語活動を直接評価する「パフォーマンステスト」が行われるようになった。パフォーマンステストは、筆記テストとは異なり、正しいか正しくないかという単純な採点はできないため、通常、評価

の観点を表にした評価シート（ループリックとも呼ばれる）が使用される。たとえば、表 4 は、スピーチを評価するための評価シートの例である。

表 4 評価シートの例

	もう少し！	できた！	すばらしい！
スピーチ	簡単な短いスピーチができた	聞き手がよくわかるようなまとまりのある簡単なスピーチができた	はっきりとした、構成のしっかりした詳しいスピーチができた
意見の伝え方	意見を述べることができた	意見や主張したいことについて、理由や根拠を説明できた	述べたい点をうまく強調し、重要な点を詳しく説明して、わかりやすく論じることができた
話の進め方	接続表現を使って、簡単な文を並べて話した まとまりはあまりはっきりしなかった	まとまりがはっきりしていた 構成はシンプルだった	文と文、段落と段落の関係を考えた体系的な構成だった
流暢さ	言い直したり言いかえたりすることがあったが、言いたいことはだいたいわかるように話せた	文法やことばをまちがえたとき、少し時間があいたり言い直したりしたが、全体的によくわかるように話せた	だいたい一定の速さで、止まらずに話すことができた
発音	ときどき聞き取りにくい発音があったが、だいたいわかるように話せた	たまに発音やイントネーションのまちがいがあったが、よくわかるように話せた	はっきりした自然な発音とイントネーションだった

JF スタンドアードのガイドブックには、評価シートの作成方法が記載されており、観点として、課題が達成できたかどうかのほかに、文法、語彙、発音といった言語構造的なものや、談話構成、流暢さなどの語用能力的なものが使用されている（国際交流基金 2017a）。しかし、評価シートに記載する評価の観点は作成者（教師）が主観的に決めるものであり、何らかの研究に基づいたものではない。どのようなコンテキスト（場面・話題）で、何を（内容）を話すかによって、評価の観点やその比重は変わる可能性があり、教師はそれを考慮したうえで評価シートを作成する必要がある。

4 評価研究

非母語話者の発話をどのように評価するかという研究は、これまで数多く行われてきた。これまでの研究を教育現場での評価に活かすため、また、今後どのような研究が必要かを考えるため、以下、「だれが」「どのような発話の」「何を」「どのように」評価するかという視点から、先行研究を概観する。

4.1 だれが評価するか

4.1.1 一般日本人が評価したもの

近年の評価研究では、評価者を日本語教師でない一般の日本語母語話者（以下、一般日

本人) とする調査が多く行われている (関 1989、東間・大坪 1990、小河原 1993、佐藤 1995、原田 1998、織田 1999、石崎 2000、李 2002、小池 2003、渡部 2004 など)。それは、教師が教材や教育方法を考えるうえで、日本語教育的観点をもたない一般日本人がどのように非母語話者を評価するかということにも目を向ける必要性が出てきたためである。

たとえば、原田 (1998) は、教材開発や評価基準の作成において、一般日本人による視点が反映されていないことが問題であると指摘し、一般日本人が非母語話者の会話をどのように評価するかを探った。その結果、文法や語彙などの言語規則より、あいづちや問い返しといったコミュニケーションを円滑にする要素が注目されることがわかった。このような研究の知見を教師が活かし、一般日本人の視点を教育実践に取り入れていくことも必要であろう。

また、多くの場合、録音した音声や録画したビデオを評価者が聞いたり見たりして評価しているが、小池 (2003) は、主婦と学習者の自由会話を収集し、その会話の参加者本人(主婦) が学習者の評価をするという方法をとっている。日常生活において、話しことばを評価するのは会話の参加者であり、ほとんどの場合、会話を外から聞いている第三者ではない。評価者に、会話の相手になったつもりで評価するよう指示することも可能であるが、2.1 で述べた蒲谷 (2013) の「場面(場+人間関係)」「意識」「内容」「形式」の連動に照らして考えると、会話の参加者のほうが、より「場面」「意識」を連動させて評価することが可能であろう。また、相手の「反応」を見ながら会話を進める (国際交流基金 2007) のが自然なコミュニケーションであれば、会話の参加者でなければ評価できない観点もある。その点では、小池 (2003) で得られたような参加者のコメントが貴重である。

4.1.2 評価者の属性による評価の差異を明らかにしたもの

一般日本人が非母語話者の発話をどのように評価するかが明らかになってくると同時に、評価者によって評価の仕方がどのように異なるかを明らかにする研究が多く行われるようになった。

日本語教師と一般日本人の評価の差異を探ったものが、Okamura (1995)、河野・松崎 (1998)、山内 (2004)、渡部 (2005)、野原 (2008)、崔 (2013)、渡辺・松崎 (2014) などである。これらの研究から、日本語教師は一般日本人より評価が厳しいことや、日本語教師と一般日本人は評価の観点が異なること、日本語教師のほうがより分析的に評価すること、コミュニケーション遂行や理解という観点では日本語教師のほうが寛容であることなどがわかっている。山内 (2004) は、OPI テストの評価基準が一般日本人による評価の基準とどの程度合致しているかを探るため、非母語話者の発話 (日本人との会話) を、日本語教育を専門としない日本人大学生に聞かせ、コメントを求めた。その結果、多くの部分が OPI テストの評価基準と合致しながらも、「聞き取り能力」「応答の早さ」「態度」「積極性」「発話の自由性」「話しことばらしさ」「日本語らしさ」「カチンとくる誤り」などの観点が OPI テストに考慮されていないことがわかった。野原 (2008) は、非母語話者の発話を日本語教

師と一般日本人に聞かせ、わかりやすさの要因を探った。その結果、日本語教師は言語規則などを分析的に評価し、一般日本人は内容を重視することがわかった。それに対し、松崎 (2004) は、「日本語教師が一般日本人より厳しいのは一般日本人が言語規則をメタ的に評価できる能力を持たないからであり、また、日本語教師が理解に寛容なのは誤りの傾向がわかるために発話意図が汲めてしまうからである」と二面性を指摘している。一般日本人が日本語教師のような視点や能力をもたないのは当然のことであろう。一般日本人を評価者とした研究は、一般日本人に発話の「採点」をさせようとしているのではなく、一般日本人の視点や印象を明らかにすることが目的であったはずである。したがって、一般日本人の評価を立脚点とした場合、「採点」をさせるのではなく、評価者の「本音」を引き出せるような方法論を確立することが必要であろう。

また、日本語教師による評価について、母語話者教師と非母語話者教師の差異を明らかにした研究もある (渡部 2003、清水 2006、渡辺・松崎 2014 など)。清水 (2006) は、日本人教師と韓国人教師の差異を、渡辺・松崎 (2014) は、日本人教師とロシア人教師の差異を探った。その結果、両者の評価の観点や重みづけが異なることが明らかになっている。現在、海外で教える日本語教師のうち、約 8 割が非母語話者である (国際交流基金 2017b)。また、日本国内においても非母語話者教師は存在する。そのような中で、非母語話者教師が母語を同じとする／同じとしない非母語話者学習者をどのように評価するかという研究をさらに進めていく必要があるだろう。

さらに、学習者の評価傾向を探ろうとしたのが佐藤 (1995) である。佐藤 (1995) は、音声日本人と韓国人に聞かせ、評価の差異を探った。学習者が他の学習者をどのように評価するかといった研究は管見の限りほとんど見当たらないが、グローバル化における非母語話者同士のコミュニケーションの必要性を考えると、このような研究も必要であろう。

また、佐藤 (2001)、松崎 (2002) などのように、学習者が自分自身の発話の評価し、教師の評価と比較したり、自己モニター能力に着目したりする研究も進めていく必要がある。

グローバル化における非母語話者同士のコミュニケーションの必要性を考えると、このような研究も必要であろう。

また、佐藤 (2001)、松崎 (2002) などのように、学習者が自分自身の発話の評価し、教師の評価と比較したり、自己モニター能力に着目したりする研究も進めていく必要がある。

4.2 どのような発話の評価するか

4.2.1 発話データの収集

評価研究の実験／調査において、これまでどのような発話の評価されてきたか、それがどのような方法で収集されたものかを以下に整理する。

発話データ収集の方法は、大きく分けて、決まった文を読み上げたもの (関 1989、東間・大坪 1990、小河原 1993、佐藤 1995、河野・松崎 1998、佐藤 2001、松崎 2002、清水 2006、三木 2008、渡辺・松崎 2014) と、自由に発話したもの (Okamura 1995、原田 1998、織田 1999、

石崎 2000、李 2002、小池 2003、渡部 2003、山内 2004、渡部 2004、渡部 2005、崔 2007、野原 2008) がある。

決まった文を読み上げる場合、意図的に誤りや誤解を与えさせるようなものを含む発話データを作成することが可能となる。小河原 (1993) は、「鈴木さんは事故を起こして病院に運ばれた弟と一緒にいます」のような多義的に解釈できる文を作成し、学習者に読み上げさせて録音した。佐藤 (1995) は、読み上げた文からさらに合成音声を作成し、オリジナル音声と比較して自然になったか不自然になったかを評価させた。渡辺・松崎 (2014) は、ロシア語を母語とする日本語学習者に単語を読み上げさせて録音しているが、使用された単語には、ロシア語を母語とする学習者によくみられる発音上の誤りが含まれる。以上のように、明らかにしたいことがマイクロに絞られている場合、決まった文を読み上げる方法が適しているが、一方で、日常生活における自然なコミュニケーションでもそのような発話が行われているかどうかについては疑問が残る。

より自然なコミュニケーションを評価する場合は、できる限り自然な発話を収集することが必要であろう。自由な発話を収集する方法にもいろいろあるが、原田 (1998) や小池 (2003) は、ロールプレイという方法で発話データを収集している。原田 (1998) は①教師をパーティーに誘う、②宿題の遅れを教師に謝る、③日本語で書いた手紙のチェックを教師に頼むという 3 つのロールプレイを実施した。一方、インタビューという方法をとったのが、Okamura (1995)、織田 (1999)、石崎 (2000) などである。Okamura (1995) は教師が夏休みや将来について学習者にインタビューし、石崎 (2000) は初対面の年長母語話者が学習者に大学生活や勉強についてのインタビューしている。いずれも、ロールプレイの相手やインタビュアーが発話をコントロールしているため、ある程度の必要な発話データを収集することが可能である。また、小池 (2003)、山内 (2004)、渡部 (2004) は、より自由度の高い、おしゃべりのような会話を録音している。

2.1 で述べたように、自然なコミュニケーションには、「目的」「情報差」「選択権」「反応」といった要素が含まれる (国際交流基金 2007)。コミュニケーション全体をどのように評価するかといった評価研究においては、これらの要素を含んだ発話データを収集する必要があるだろう。ロールプレイの場面設定や、会話の相手を工夫することで、蒲谷 (2013) の「場面 (場+人間関係)」をコミュニケーションに連動させることや、社会言語能力に着目することも可能になる。

4.2.2 場面・話題

次に、評価に使用した発話データがどのような場面において、どのような話題で話されたものかという視点で、先行研究を見てみたい。上述した通り、コミュニケーションには、社会言語能力 (相手との関係や場面に応じてスピーチスタイルや話題を変える能力) が必要である。蒲谷 (2013) のコミュニケーションの枠組みにおいても、場面 (場+人間関係) の重要性が指摘されている。

Okamura (1995)、原田 (1998) は教師と学習者の会話場面であり、原田 (1998)のロールプレイでは、教師を誘う、教師に謝る、教師に依頼するといった、人間関係の構築に関わるものである。石崎 (2000) は、インタビュアーを「初対面の年長母語話者」とし、表現や話題により配慮が必要な場面を設定している。さらに、崔 (2007)では、アルバイトの面接場面を設定している。これは、ロールプレイや会話練習ではなく、実際にアルバイトを募集して面接を実施し、面接官との会話を録音したものである。以上のように、誤解を与えることが許されない場面や、自分が不利になっては困るような場面が評価研究において多く設定されているのは、非母語話者が日常生活で困らないような教育に結び付けようとする教育者の意図であろう。

このような特定場面に絞った評価研究について、松崎 (2007) は結果を一般化しすぎることの危険性と限界を指摘し、以下のように述べている。

理詰めで考えてこのような研究が教育に寄与するものは少ないし、仮に、場面別・表現別の評価が明らかにされたとしても、現場への提言としては、発音の正確さが求められる場面の表現をサバイバル的に取り出し、丸暗記で正確・自然な発音ができるように訓練するという教育方針にしか結びつかない。果たしてそれで十分と言えるだろうか。(松崎 2007: 300)

場面や話題によって評価の観点が異なることは明らかであり、多様な場面における評価について今後も研究していく必要はあるだろう。しかしながら、あらゆる場面を想定して実験／調査を行うのはケーススタディにすぎない。また、個々の場面で出た結果から、ある程度普遍的な提言をすることを目指していく必要がある。

4.2.3 課題・タスク

発話データが、どのようなコミュニケーション上の課題／タスクを達成しようとしたものかについても考える必要がある。自由発話を評価材料とした多くの評価研究では、会話、つまり相手のいるインタラクティブなやりとりを使用しているが、野原 (2008) は、4コマの絵を見てその内容を語る「ストーリーテリング」の手法でデータを収集している。会話の評価では、コミュニケーションを円滑にする要素や、流暢さといった点に評価者が注目するという結果が出ているが、野原 (2008) では、内容や意味的つながりが重要視されているという結果となっている。このことから、話者がどのような課題／タスクを達成しようとしているかによって、評価の観点が異なることがわかる。場面や話題における違いだけでなく、どのような課題を遂行するかによって評価の観点や重みが異なることも視野に入れて調査デザインをする必要があるだろう。

4.3 発話の何を評価するか

4.3.1 音声のみを評価したもの

発音評価研究は、大きく分けて、「音声の諸要素のうち何がどれくらい重要か」を明らかにするための調査と、「言語行動の諸要素のうち発音はどれくらい重要か」を明らかにするための調査がある(松崎 2004)が、それに照らして考えてみると、話しことばの評価研究においても、音声のみを取り上げて評価を行ったものと、発話全体を評価したものがある。

関(1989)、東間・大坪(1990)、小河原(1993)、佐藤(1995)、河野・松崎(1998)、佐藤(2001)、李(2002)、清水(2006)、渡辺・松崎(2014)は、評価者がどのように学習者の発音を評価するかを明らかにしている。小河原(1993)は、誤解の原因となる発音がとくに厳しく評価されることや、発音が上手になるにつれて評価が厳しくなることを明らかにした。また、佐藤(1995)では、音声の何が評価に影響するかを調査したところ、「高さ」「長さ」「強さ」の中で、「高さ」が重要だということや、単音より韻律の影響が大きいことがわかった。

このような研究は、JFスタンダードの「コミュニケーション言語能力」、つまり、「JFスタンダードの木」の根の部分をもさらに具現化していくうえで重要だといえる。ただし、根は枝葉を支えるものであり、ある具体的な言語能力が、どのような言語活動をどのように支えているのかを考慮しながら研究を進めていく必要があるだろう。

4.3.2 発話全体を評価したもの

発話全体を評価し、発話全体の中で何がどれくらい重要かを探ったものには、原田(1998)、小池(2003)、渡部(2003、2004、2005)などがある。原田(1998)は、ロールプレイの音声を母語話者に聞かせ、自由に印象を語ってもらった。その結果、文法などの言語規則より、あいづちや問い返しなどの円滑なコミュニケーションの遂行に関する要素が注目されることがわかった。小池(2003)も、インタビューの感想や印象を自由に語ってもらったところ、原田(1998)同様、円滑に流れるコミュニケーションかどうか注目されることがわかった。また、発音面に関しては意味が推測できれば問題にならないことを示唆している。渡部(2004)は、自由会話の音声を評価者に聞かせ、何を評価するかを抽出した。その結果、評価の対象として、「ストラテジー」「表現力」「対話能力」「聴きやすさ」「語彙・表現」の5因子が抽出された。発話全体の評価において、どのような観点があるのか、それはどの程度重要視されるのかといった研究は、コミュニケーション評価の中で語彙、文法、音声などの各要素をどのように位置づけるかを提言するのに役立つであろう。

また、「わかりやすさ」や「流暢さ」といった観点に着目し、その要因を明らかにしたものに、織田(1999)、石崎(2000)、野原(2008)などがある。織田(1999)は、流暢さとは何かを探るため、OPIテストのインタビューを評価させた。その結果、「コミュニケーション能力」「印象」といった因子が抽出され、流暢さとは発話の構成概念ではなく、言語を運用する能力であることが示唆されている。石崎(2000)は、わかりやすさの要因を、質問

紙による主観的評価と、発話そのものを分析した客観的評価によって明らかにした。結果として、語彙が豊富で音の高低が自然であれば意味がわかりやすいことがわかった。一方、野原 (2008) も、わかりやすさの要因を探ったが、「内容」「意味的つながり」「言語使用の適切さ」「音声」「音韻的つながり」という因子を抽出している。このような研究によって、たとえば、教育現場で「わかりやすい発話」を具体的に指導したり評価したりすることが可能になる。

石崎 (2000)、李 (2002)、崔 (2007) は、発話が聞き手にどのような印象を与えるかといった感情面に着目した。石崎 (2000) は、前述した「わかりやすさ」のほかに、「癪に障る」要因についても明らかにした。その結果、音の高低が不自然で区切りが不適切な場合に、感情面の評価に関係する (癪に障る) ことがわかった。李 (2002) は、音声のみを評価対象としたが、ソウル方言の干渉とみられる句末イントネーションに「不快」「失礼」などのコメントが集中した。崔 (2007) は、実際のアルバイトの面接を評価材料とし、何が対人印象に影響するかを探った。その結果、対人印象にはパラ言語や非言語的要素が大きく関わることが示唆された。水谷 (2015) が、「対人関係」という観点から、これからのグローバル社会において、「感じのよい日本語」という観点での研究の必要性を述べている通り、以上のような、感情面に関わる研究も進めていく必要があるであろう。

4.4 どのように評価するか

4.4.1 質問紙調査

最後に、評価研究における実験／調査の方法論について概観する。まず、質問紙を用いて調査を行った研究として、Okamura (1995)、織田 (1999)、石崎 (2000)、野原 (2008) などがある。石崎 (2000) は、「語彙が豊富か」「文法の誤りが多いか」「発音がいいか」などの 20 の質問に 5 段階で答えるという方法で調査を行っている。質問紙は量的調査の方法として用いられ、多数の評価者を対象とした調査が可能である。また、一般日本人のように分析的視点をもたない評価者にとっても、あらかじめ分析的に項目が決められていれば、評価しやすい方法といえるであろう。ただし、質問紙の項目の選択には注意する必要がある。たとえば、評価の観点を明らかにしたい場合、評価者が着目する可能性のある観点を網羅しておかなければならない。先行研究において使用されている項目を使用したり、項目を決定するための予備調査を行うなどして、「だれが」「どのような発話の」「何を」評価するかを考慮しながら、項目を決定する必要がある。

4.4.2 インタビュー・自由記述

質問紙の項目をあらかじめ決めておくのではなく、インタビューや自由記述、自由コメントという方法で、気づいたことを語ってもらう方法をとっているのが、原田 (1998)、小池 (2003) などである。小池 (2003) は、会話の参加者である一般日本人に、会話の直後と、会話を録画したビデオを見せた後に、半構造化インタビューをしてコメントを求めた。得

られたコメントを質的に分析しているが、「話しづらい」「わかりにくい」などの漠然としたコメントが多く、評価者がなぜそのように感じたかというところまでは語られていない。評価者に自由なコメントを求めた場合、質問紙の項目として拾いきれなかったような予期せぬコメントが得られる場合もあることがメリットといえるが、とくに評価者が一般日本人の場合、松崎 (2004) が述べているように、分析的・メタ的に表現する能力をもたないことがあるため、信頼性の面でも疑問が残る。また、分析的・メタ的に表現する能力をもつ日本語教師であっても、「気づいたが指摘しなかった」「指摘することが多すぎたため、その一部しか言及しなかった」という可能性も否めない。加藤 (1997) の教師の非訂正モデルによると、「気がつかないから指摘しない」「気づいたが許容して指摘しない」「気がついて気になったが(たとえば時間的な制限があつて)指摘しない」というパターンがあり、コメントとして表出しなかったものがそのどれに当たるのかはわからない。さらに、小河原 (1993) の「上手になるにつれて評価が厳しくなる」のは、ほかに指摘することがなかったために、重要でないことを指摘してしまうという行動も起こりうる。

表 5 先行研究まとめ

文献	だれが	どのような発話の	何を	どのように
関 (1989)	日本人大学生 (一般日本人)	「桃太郎」読み上げ	発音	発音の各要素の自然さを 7 段階評定 SD 法で日本語らしさを 7 段階評定
東間大坪 (1990)	一般日本人	語・短文読み上げ	発音	何と聞こえたか二択 正しさを 5 段階評定
小河原 (1993)	日本人大学生 社会人 (一般日本人)	留守番電話メッセージ 曖昧文読み上げ	発音	誤解に気づく前後に 5 段階評定、評価が下がったかインタビュー
佐藤 (1995)	一般日本人 韓国人	会話文読み上げ →合成音声	発音 (単音か韻律か)	オリジナル音声と合成音声と比較し、自然になったか/不自然になったかを 7 段階で評価
Okamura (1995)	日本語教師 一般日本人	教師→学習者 インタビュー 夏休み・将来について	評価の観点の重み (文法、流暢さ、適切さ、語彙、理解度、発音)	質問紙
河野松崎 (1998)	日本語教師 一般日本人	会話文読み上げ	発音 日本語らしさ	日本語らしさを 7 段階評定
原田 (1998)	主婦 (一般日本人)	日本人⇄学習者 ロールプレイ ①パーティーに誘う ②宿題の遅れを謝る ③手紙のチェックを頼む	評価者が何に注目するか	印象を自由に語る
織田 (1999)	一般日本人	OPI テスト インタビュー 超級、中級	流暢さとは何か	質問紙→因子分析

石崎 (2000)	一般日本人	初対面の年長母語話者 →学習者 インタビュー (大学生 活、勉強について)	評価の観点 わかりやすさ 痛に障るか	主観的評価 (質問紙) 客観的評価 (発話の分析)
佐藤 (2001)	学習者 一般日本人	会話文読み上げ	発音	発音を評定 学習者は自分の音声を自 己評価
李 (2002)	日本人大学院生	自然談話 自己紹介の部分	発音	自由記述 質問紙 インタビュー
松崎 (2002)	日本語教師 学習者	短文を発音	発音	読んだ後で評価札 (○×) 気をつけたことをメモ 教師→学習者評価 学習者は自己評価 ビデオ再生→インタビュ ー
小池 (2003)	主婦 (一般日本人) *会話参加者	日本人との自由会話 依頼のロールプレイ	評価者が何に注目 するか	インタビュー (感想、印象 を自由に語る)
渡部 (2003)	学習者 非母語話者教師 日本人教師 一般日本人	日本人との自由会話	評価の観点	自評価の観点を自由に語 る
山内 (2004)	一般日本人	日本人との自由会話	評価の観点	自由記述 →OPI の評価基準と比較
渡部 (2004)	一般日本人	日本人との自由会話	評価の対象 全体的評価	評価板 (○/×) を挙げ、 その理由を話す
渡部 (2005)	日本語教師 一般日本人	日本人との自由会話 自己紹介、好きな食べ 物について	評価の観点と重み	質問紙
清水 (2006)	日本人教師 韓国人教師	文章読み上げ	発音	上手さを 5 段階評定 評定理由を自由に語る
崔 (2007)	日本語教育専門 家 大学生、主婦 社会人	アルバイトの面接	印象	質問紙 言語・パラ言語・非言語的 評価 対人印象評価
野原 (2008)	日本語教師 一般日本人	ストーリーテリング 4 コマの絵を見て話す	わかりやすさの要 因	質問紙
三木 (2008)	一般日本人	「これじゃない」を含 む会話文読み上げ	発話意図とイント ネーションの合致	音声を聞いて発話意図を 選択
崔 (2013)	日本語教師 一般日本人	日本人→学習者 インタビュー 食べ物、生活習慣につ いて	評価の観点 印象	質問紙 言語、パラ言語、非言語 対人印象
渡辺 松崎 (2014)	一般日本人 日本人教師 ロシア人教師	単語読み上げ	発音	音声を聞いて 4 段階評定 評価の理由をコメント

5 まとめと今後の課題

以上、コミュニケーションの捉え方や、話しことばの評価方法、これまで行われてきた評価研究を概観した。

CEFR の行動中心主義や、JF スタンドアートの課題遂行の考え方が教育現場に取り入れられつつあり、教材や教授法では新たな試みが行われている一方、評価の方法や評価研究に

においては、いまだ、発話を分析的に「採点」しているような印象を受ける。課題遂行を評価するためのよりよい方法を開発するには、そのための評価研究が不可欠である。そして、課題遂行を評価するにあたっては、コミュニケーション全体を考える必要があり、コミュニケーションを分析的に捉えるミクロの視点での研究を引き続き進めていくとともに、コミュニケーション全体を考えるためのマクロの視点も必要である。

2 節で述べたコミュニケーションの捉え方に沿って考えてみると、まず、コミュニケーションに必要な能力として、「文法能力」「社会言語能力」「ストラテジー能力」「談話能力」の 4 つの能力を総合的に評価するような研究が求められる。さらに、蒲谷 (2013) の枠組みである「場面 (場+人間関係)」「意識」「内容」「形式」について、これらが総合的に連動しているかについて考える必要もある。そのためには、「意識」、つまり、発話者がどのような意図で発話したかということにも着目する必要がある。聞き手がどのような印象を受けたかに着目する研究もあるが、そこでは、発話者が聞き手にどのような印象を与えようとしたかについても含めて考える必要がある。

CEFR の行動中心主義や JF スタンドアードの課題遂行の考え方では、最終的に「コミュニケーション上の課題が遂行できたかどうか」が重要となる。それを評価するには、何がどの程度できれば課題を遂行したといえるのか、コミュニケーションに必要な能力や要素の何がどの程度課題の遂行に関係しているのかを具現化していく必要がある。

「だれが」評価するかについては、日本語教師か一般日本人か、という議論だけでなく、さらにさらに視野を広げていく必要がある。「教師が学習者を評価する」「母語話者が非母語話者を評価する」という研究では、「評価＝採点」という枠組みから離れることが難しいからである。現在、社会は著しく変化し、グローバル化社会、多文化共生社会、複言語・複文化社会への変化を続けている。そういった中で、非母語話者と非母語話者の日本語によるコミュニケーション場面も増えてきている。また、ICT の発達により、世界中の日本語学習者が互いに日本語を使ってコミュニケーションを行う場面も増えていく可能性もある。そのような社会の中でのコミュニケーションとはどのようなものか、第 2 節で述べたコミュニケーション能力やコミュニケーションの要素などにも鑑みながら再考し、研究を進めていく必要がある。非母語話者が母語を同じとする／しない非母語話者を評価するという視点、非母語話者が自己を内省・モニターする自己評価の研究、さらには、非母語話者が母語話者を評価するといった研究も必要だと考えられる。

また、「どのような発話を」評価するかについては、上述したように、社会におけるコミュニケーション場面が多様化していることを考慮しつつも、特定の場面に絞った研究を場当たり的に増やしていくのではなく、場面や話題、タスクによって評価の観点や重みが異なるということを明確にし、学習者が場面を意識できるような教育につながる研究が必要である。

参考文献

- 李惠蓮 (2002)「ソウル方言話者の日本語発話の "end focus"に対する日本語母語話者の評価—日本語教育関係者を対象に—」『広島大学日本語教育研究』12: 115-122.
- 石崎晶子 (2000)「学習者の言語行動に対する母語話者の評価—主観的評価と客観的評価の関係—」『第二言語としての日本語の習得研究』3: 19-35.
- 小河原義朗 (1993)「外国人の日本語の発音に対する日本人の評価」『東北大学文学部日本語学科論集』, 31-12.
- 織田倫子 (1999)「流暢さとは何か—日本語母語話者による評価を通して—」『教育学研究紀要』45(2): 397-402.
- 加藤好崇 (1997)「日本語学習者の誤用に対する教師の非訂正行動」『日本語教育』93, 26-37
- 河野俊之・松崎寛 (1998)「一般日本人と日本語教師の音声評価の差異」『日本語教育方法研究会誌』5 (2): 24-25.
- 蒲谷宏 (2013)『待遇コミュニケーション論』, 大修館書店.
- 小池真理 (2003)「日本語母語話者は第二言語話者との会話をどのように評価するか」『北海道大学留学生センター紀要』7: 16-33.
- 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.
- 国際国流基金 (2016)『JF 日本語教育スタンダード準拠ロールプレイテスト テスター用マニュアル』第2版, 国際交流基金
- 国際交流基金 (2017a)『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.
- 国際交流基金 (2017b)『海外の日本語教育の現状 2015 年度日本語教育機関調査より』, 国際交流基金.
- 佐藤友則 (1995)「単音と韻律が日本語音声の評価に与える影響力の比較」『世界の日本語教育 日本語教育論集』5: 139-154.
- 佐藤友則 (2001)「音声評価基準の習得過程に関する考察」『第二言語としての習得研究』4: 134-149
- 清水寿子 (2006)「日本人教師と韓国人教師による発音評価」『日本語教育学会春季大会予稿集』: 187-192
- 鈴木秀明 (2005)「短時間の評価トレーニングが教師の発話評価に及ぼす効果」『言語科学研究 神田外語大学大学院紀要』11: 77-94.
- 崔文姫 (2007)「日本語学習者の発話に対する日本人の評価—韓国人の日本語学習者に対する印象とその印象に影響を及ぼす要因—」『計量国語学』26(29): 47-63.
- 崔文姫 (2013)「日本語学習者の発話に対する日本語母語話者の評価—日本語教師と非日本語教師の因果モデルを中心に—」『国立国語研究所論集』5: 1-26.
- 東間由美・大坪一夫 (1990)「外国人の日本語発話の日本人による評価」文部省重点領域研究『日

- 本語音声』研究報告 4: 70-75.
- 長坂水晶・押尾和美 (2013)「JF 日本語教育スタンダードに準拠した口頭テスト—課題遂行能力測定のためのロールプレイタスクと評価指標の作成—」『言語文化と日本語教育』46: 70-73.
- 野原ゆかり (2008)「発話の「わかりやすさ」を判断する要因—一般日本人と母語話者日本語教師の比較を通して—」『人間文化創成科学論叢』11: 165-174.
- 野原ゆかり (2011)「日本語学習者の口頭説明に対する母語話者の評価—評価コメントの質的分析による事例研究—」『比較文学研究』99: 69-80.
- 原田明子 (1998)「一般の日本人は日本語学習者の日本語をどのように評価するか」『北海道大学留学生センター紀要』2: 157-168.
- 牧野成一 (監修), The American Council on the Teaching of Foreign Languages (著), 日本語 OPI 研究会翻訳プロジェクトチーム (訳) (1999) 『ACTFL-OPI 試験官養成マニュアル』, アルク.
- 牧野成一・鎌田修・山内博之・齊藤真理子・荻原稚佳子・伊藤とく美・池崎美代子・中島和子 (2001) 『ACTFL-OPI 入門』, アルク.
- 松崎寛 (2002)「リピートのとき学習者は何を考えて発音しているか」『広島大学日本語教育研究』12: 33-41.
- 松崎寛 (2004)「発音評価研究の展望」『日本人は何に注目して外国人の日本語運用を評価するか 平成 12 年度～平成 15 年度 科学研究費補助金 基盤研究 (B) (2)研究成果報告書』: 37-51.
- 松崎寛 (2007)「発音評価研究に関する覚え書き」藤原雅憲・堀恵子・西村よしみ・才田いずみ・内山潤 (編)『大学における日本語教育の構築と展開』: 297-309, ひつじ書房.
- 三木理 (2008)「韓国語を母語とする日本語学習者の「これじゃない」の発話意図とイントネーション—音響分析と聴取に基づく考察—」『言葉と文化』9: 321-341.
- 水谷信子 (2015)『感じのよい英語・感じのよい日本語—日英比較コミュニケーションの文法—』, くろしお出版.
- 関光準 (1989)「韓国語話者の日本語音声における韻律的特徴とその日本語話者による評価」『日本語教育』68: 175-190.
- 山内博之 (2004)「OPI 評価基準に関する一考察—日本人評価の視点から—」平成 12～15 年度 科学研究費補助金研究成果報告書『日本人は何に注目して外国人の日本語運用を評価するか』: 67-75.
- 渡部倫子 (2002)「日本人評価研究の発展と課題」『教育学研究紀要』48 第 2 部: 341-346.
- 渡部倫子 (2003)「日本語口頭運用能力の評価基準—評価者による相違—」『日本語教科教育学会誌』25 (4): 11-17.
- 渡部倫子 (2004)「日本語学習者の発話に対する日本語母語話者の評価—評価尺度開発の試み—」『広島大学大学院教育学研究科紀要 第二部 文化教育開発関連領域』52: 175-183.
- 渡部倫子 (2005)「日本語学習者の発話に対する日本語母語話者の評価—共分散構造分析による

評価基準の解明—」『日本語教育』125: 67-75.

渡辺裕美・松崎寛 (2014)「発音評価の相違—日本人教師・ロシア人教師・一般日本人の比較—」
『日本語教育』159: 61-75.

Council of Europe (2004)『外国語の学習、教授、評価のためのヨーロッパ共通参照枠』, 吉島茂、
大橋理枝ほか (訳・編), 朝日出版社.

OJAE2010 チーム (2010)『OJAE (Oral Japanese Assessment Europe) CEFR 準拠日本語口頭産出能力
評価表』, 自費出版.

Canale, M., and M. Swain (1980) Theoretical bases of communicative approaches to second language
teaching and testing. *Applied Linguistics 1*, 1-47.

Canale, M. (1983) From communicative competence to language pedagogy. IN J. Richards and Schmidt
(Eds.), *Language and communication*. London, Longman.

Hymes, D.H. (1972) On Communicative Competence. In: J.B. Pride and J. Holmes (eds),
Sociolinguistics: Selected Readings. 269-293. Harmondsworth: Penguin.

Okamura, A. (1995) Teachers' and nonteachers' perception of elementary learners' spoken Japanese. *The
Modern Language Journal*, 79, 29-40.

(篠原 亜紀 筑波大学大学院生)

Trend and future of evaluation research of oral communication

SHINOHARA Aki

This paper reviews previous research on evaluation of oral communication, and discusses the evaluation research required in the future. First, the paper describes how communication has been captured in language education so far. Then, it reviews how oral communication is evaluated in the Japanese language education currently. In addition, it overviews the previous research from the point of "who", "what kind of speech", "what", and "how" evaluated. Based on the review, this paper proposes the evaluation research we need in the future, in order to evaluate oral communication.