JF 日本語スタンダードによる日本語会話能力測定（OPI）の課題と可能性について

笹　川　洋　子　　三　井　絢　子

Issues and Possibility of JF Standard for Japanese Oral Proficiency Test

Yoko SASAGAWA　　Junko MITSUI

要　旨

JF 日本語教育スタンダードは，CEFL に基づき，外国人日本語学習者の日本語会話能力を測るために開発された。従来の日本語能力試験では，語彙や文法など言語構造的能力の測定を主としてきたが，JF スタンダードは CAN-do リスト（言語運用能力）を評価基準に取り入れ，多文化理解のためのコミュニケーション能力をつけることを目的としている。本稿では，JF スタンダードと，基盤となった CEFL を比較し，JF スタンダードの変更点を確認する。さらに，実際に JF スタンダードを用いて会話評価を行った研究例を取り上げ，(1)評価の安定性，(2)評価のしやすさ，(3)タスク達成能力の評価について考える。最後に，JF スタンダードの利点と課題について述べたい。

キーワード：JF スタンダードの課題，日本語会話能力の測定

1　はじめに

本稿では，国際交流基金によって開発された JF 日本語教育スタンダード（JF スタンダード）を実際の会話テストに応用する利点，JF スタンダードの可能性，そして課題について考えてみたい。

外国語学習者，そして日本語学習者は，外国語で母語とは違う言語を話す人々と自由に話せるという状況に憧れ，言語学習の目標とする。言うまでもなく，外国語能力は読む，聞く，書く，話す能力を統合したものであるが，特に話す能力，コミュニケーション能力の育成は外国語教育にとって欠かすことのできないものである。しかし，外国語の会話能力の評価は必要とされながら，読解，文法，作文，聴解などの評価の中でも，研究や実践が進んでいない分野の一つである。しかも，日本語を実際に教える現場の教師は学習者の会話能力を，学習の最も重要な技能と考えており，その経験から多少直感的に評価を行ってきた。つまり，いちばん重要だと考えられる課題が，最後尾に置かれるという現象が起こっているのである。もちろん，経験豊かな教師は，短い時間内の面接で，様々な角度から言語能力を観察し，的確に学習者の能力を評価することができる。日本語教育の会話能力評価は長い間，この経験値による観察に依

拠して行われてきたと言えるであろう。

この経験による観察を越えて，できるだけ客観的な指標を立て，インタビューに活かそうという試みの一つが ACTFL（The American Council on the Teaching of Foreign Languages）である。詳しくは後述するが，ACTFL は英語圏で英語会話の能力テストとして開発され，日本語の会話能力測定に応用された。しかし，会話評価の指標は主に語彙・文法に拠るものであり，会話で重視される流暢さは測りにくい。その点，CEFR（Common European Framework of Reference for Languages: Learning, teaching, assessment：「外国語の学習，教授，評価のためのヨーロッパ共通参照枠」）に依拠して開発された JF スタンダードによる会話テストは，より積極的に会話の流暢さを測る可能性を備えている。ここでは，先行研究を振り返り，評価することで，JF スタンダードによる会話テストの可能性と課題を探ることにする。（国際交流基金，2007他参照）

2　日本語スタンダードによる日本語能力の測定

日本語教育スタンダード（以下 JF スタンダード）はコースデザイン，授業設計，評価を考えるために国際交流基金によって開発された日本語能力の評価枠組みである。JF スタンダードは，CEFR（外国語の学習，教授，評価のためのヨーロッパ共通参照枠）を基盤としている。CEFL は，英語，フランス語，ドイツ語など EU 圏で用いられている言語を同じ基準で計ろうとするものである。そして，語彙や文法規則に違いのある言語の上達度を評価するために，その言語で何ができるかという CAN-DO リストが開発された。例えば，「産出：経験や物語を語る」は，初級のＡ１レベルでは「自分が作った料理の味や材料について，“魚の料理です”，“からいです”など，簡単な言葉で友人に紹介することができる」という具体的な場面が想定されているが，中級のＢ１レベルでは「子供時代の習い事や学校生活などについて，その当時の夢などと関連付けながら，友人に話すことができる。」と会話内容の抽象度が高くなっている。このように設定された課題を達成できたかどうかで測れば，理論的には，ヨーロッパの言語と，これとは異なる文法構造をもった日本語や中国語などの言語能力測定の基準を共有することができる。なお，CEFR は，2001年にヨーロッパの言語教育・学習・評価の場で共有される枠組みとして発表され，以後，広く外国語教育や評価の場で利用されている。CEFL に依拠した JF スタンダードの基準を用いることにより，日本語の熟達度を CEFR の基準に照らし合わせることができる。（岡崎，古川，三原，2011他参照）

JF スタンダードの「日本語で何がどれだけできるか」という課題遂行能力はＡ１，Ａ２，Ｂ１，Ｂ２，Ｃ１，Ｃ２の６レベル指標になっている。カテゴリー（言語能力や言語活動の例）ごとに分けられ，課題遂行能力は Can-do（「～できる」という文）で表されている。

産出には「経験や物語を語る」や「論述する」などの技能が含まれるが，そのうち「講演やプレゼンテーションをする」のＡ１～Ｂ２の段階別技能には次のような例が含まれる。

Ｂ２「学会発表などで，あらかじめ準備してあれば，自分の研究内容についてデータなどを示しながら，明確に詳しく説明し，多少発表の趣旨とずれた質問に対しても柔軟に対応することができる。」

Ｂ１「勉強会や授業などで，あらかじめ準備してあれば，"留学生の悩み"など身近な話題についてのアンケート調査の結果や感想などを含んだ，まとまりのある簡単な発表をし，想定した質問に答えることができる。」

Ａ２「弁論大会などで，メモをときどき見ることができれば，異文化体験などについて，短い簡単なスピーチをすることができる。」

Ａ１「転勤や退職などの機会に開いてもらった送別会で，出席者の前で，“お世話になりました”“ありがとうございました”など簡単な定型表現でお礼の挨拶をすることができる」

複雑な表現や論理構成を必要とする発表から，簡単な挨拶までの幅広いプレゼンテーション場面が想定されている。

なお，JF スタンダードでは，「JF スタンダードの木」という概念図が示されている（図１）。木の根にあたる能力には「言語構造的能力」，「社会言語能力」，「語用能力」がある。そして，このような基礎能力はコミュニケーション活動を支える，木の枝である「コミュニケーション言語活動」として表される。ここでは，「受容」，「産出」，「やりとり」，「テクスト」，「方略」が重要な支点となる。

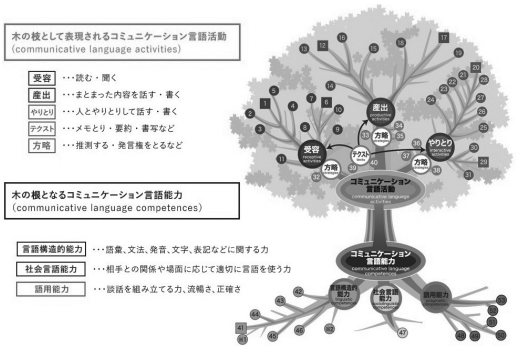


図１「JF スタンダードの木」

JF スタンダードの利点は次のように説明されている。

- 国や機関が違う教師同士でも，同じ基準で話し合いや情報交換ができる。
- 教師と学習者が学習の目標を共有することができる。
- 学習者の日本語力を，他の言語との共通基準にもとづいて説明できる。

なお，Can-do で測る課題遂行能力を評価基準とするということは，従来行われてきた日本語能力試験（Ｎ１，Ｎ２など）とは別の，達成すべき言語能力を設定する必要があるということになる[注1)]。先に紹介した「JF スタンダードの木」では，JF スタンダードが測ろうとするコミュニケーション能力が明示されているが，注目すべきは，これまで日本語の言語能力で優先的に測定されてきた文法や語彙力が，言語構造的能力として，基盤となる能力の一部として捉えられていることである。これは，CEFL が多文化共存を理念として作成された経緯に起因する。（国際交流基金，ibid.，吉島，2013他参照）

> 言葉を通した相互理解のためには，その言語を使って何がどのようにできるかという課題遂行の能力と，さまざまな文化に触れることでいかに視野を広げ他者の文化を理解し尊重するかという異文化理解の能力が必要です。（国際交流基金 2010a：1）

相互理解のための課題遂行能力，異文化理解能力は CEFL の母体，EU 諸国の理念である。吉島茂（2013）は CEFL が言語と文化の両方を学習対象としており，Can-do statement が COE（Council of Europe：欧州評議会）の目指す多種民族の平和的共存，そのための Plurilingualism・Pluriculturalism（複数言語・複数文化状態）の実現のために，言語能力測定の理念を共有できることを前提としていると記している。つまり，そうした理念を念頭に置き自己評価していくことで，個々人の多文化共存のための知識が養われ，多文化共存を支える市民が育つと考えるのである。（吉島茂，2013：11参照）

このような理念が JF スタンダードによる会話テストの設定にも反映されているが，特にＢ２からＣ１のレベルではそれが顕著に感じられる。次にこのような多文化理解の要素が JF スタンダードにどのように体現されているかを確認してみよう。

3　JF スタンダードによる言語評価テストでは，多文化理解がどのように測られるか。

JF スタンダードには「聞くこと」「読むこと」「書くこと」と「話すこと（やり取り・表現）」の能力基準が設けられている。ここでは，そのうち会話能力評価に関わる項目のうち，多文化への姿

勢が問われる「Ｂ２」「Ｃ１」レベルについてみていこう。

まず，「話すこと」について，どのようなことが話せるかという CAN-DO リストは，会話の相手との「やりとり」の技術と自分自身の会話「表現」に分けられる。ここでは，コミュニケーション・ストラテジー，語彙・文法力，流量さなどが評価され，併せて対象となる言語の社会文化に対する知識，多文化社会に対する個人の姿勢が測られていることがわかる。

「Ｂ２：話すこと」

［やり取り］流暢に自然に会話をすることができ，母語話者と普通にやり取りができる。身近なコンテクストの議論に積極的に参加し，自分の意見を説明し，弁明できる。

［表現］自分の興味関心のある分野に関連する限り，幅広い話題について，明瞭で詳細な説明をすることができる。時事問題について，いろいろな可能性の 長所，短所を示して自己の見方を説明できる。

「Ｃ１：話すこと」

［やり取り］ことばをことさら探さずに流暢に自然に自己表現ができる。 社会上，仕事上の目的に合ったことば遣いが，意のままに効果的にできる。自分の考えや意見を正確に表現でき，自分の発言を上手に他の話し手の発言に合わせることができる。

［表現］複雑な話題を，派生的話題にも立ち入って，詳しく論ずることができ，一定の観点を展開しながら，適切な結論でまとめ上げることができる。

次にテスター側の「Ｂ２・Ｃ１」で測るべき評価基準を確認してみよう。もちろん，これは CAN-DO リストの評価と連動しており，内容は重なる部分が多い。

「Ｂ２」レベルでは「意見とその根拠を言える。社会的な話題を扱ったフォーマルなインタビューで流暢にやりとりができる」が求められる。「Ｂ２」の会話テストはテスターとのディベートを行う課題で，JF スタンダードのモデルでは「新しい日本語教科書に切り替えるか」「フレックス制を取り入れるか」「社会の連絡はメールが良いか，直接話すのが良いか」という課題が示されている。また，「Ｃ１」は「環境問題」「教育問題」など社会問題について話す課題である。「Ｂ１」までは趣味や読んだ本，料理など自分の身近な話題について話すものだが，「Ｂ２」以上のレベルでは言語能力に加え，被検者の素養が関わってくることがわかる。つまり，CEFR のめざす，多文化社会の市民としての資質が備わっているかどうかが重視されるものになっている。

会話テストでは，さらにテスター側の判定の詳細な指標が設けられているが，これを見ると，このような多文化理解の資質が求められることが明確になる。例えば，「Ｂ２」のディベートでは，ただ反論するだけではなく，相手の反論に理解を示しながら，新しい論点を示していくことと，明記されている。

［Ｂ２の判定の指標］

◎意見や提案と，その根拠を，はっきりと複数述のべることができる。相手の反論に対たいして理解を示しながらも自分の解釈を述べたり，積極的に新しい論点を示し筋道立てて意見を述べたり，仮説を立てて意見を補強したりして説得力ある主張を維持できる。

○意見や提案と，その根拠をはっきりと述べることができる。相手の反論に対して自分から積極的に新たな論点を示し何とか 意見を補強できる。

同様に，社会問題について語る「Ｃ１」の課題では，問題の背景と関連付け，複数の視点から論点を深めることが求められる。言語能力だけではなく，被験者が市民としてどのように社会に対峙しているかという姿勢が併せて測られるわけである。

［C 1の判定の指標］

◎社会的な問題の具体的内容を社会的な影響や背景と関連づけて，常に自然で流暢に筋道立てて詳細に説明できる。相手からの質問やコメントに，面接に合った言葉遣いで積極的に対応し，議論を複数の視点から自在に深めたり広げたりすることができる。

○社会的な問題の具体的内容を社会的な影響や背景と関連づけて，常に自然で流暢に筋道立てて詳細に説明できる。相手からの質問やコメントに，面接に合った言葉遣いで積極的に対応し，議論を何とか深めたり広げたりすることができる。

なお，JF スタンダードでは，自律学習教材として，ポートフォリオを作成することを推奨している。これは ELP（ヨーロッパ言語ポートフォリオ）の考え方を基にしており，言語学習者および利用者が言語能力を管理運用する手段とされている。ELP には記録・レポートとしての機能に加え，教育的機能として「言語学習者の動機を高める」「学習者を刺激して，自らの目的を考えさせ，学習を計画させる」「自己評価のためのグリッドを提供する」「学習者に多言語的・異文化間経験の獲得を促す」が設定されている。JF スタンダードではポートフォリオの効果として次のような点があげられている。ほぼ，ELP の理念を受け継いだものと言えよう。(国際交流基金，2007，吉島，ibid. 他参照)

・教師と学習者が学習目標と学習の過程を共有できます。
・学習者が他の教育機関に移動したときにそれまでの学習成果を正確に伝えることができます。
・学習者が自己評価や体験を記録することで，課題遂行能力や異文化理解能力だけでなく，自律的学習能力や学習の動機づけを高めることができます。
・日本語能力だけでなく，教室の中や外で学んださまざまな知識や技能の学習成果の評価も行うことができます。

ヨーロッパ言語ポートフォリオは，言語パスポート，言語バイオグラフィー（経歴 記録），ドシエ（関連書類）の 3 つの部分から成るが，JF スタンダードでは，(1)評価表，(2)言語的・文化的体験の記録，(3)学習の成果が設定されている。「(1)評価表」は「①自己評価チェックリスト」「②学習活動の評価基準や評価シート」，「(2)言語的・文化的体験の記録」は「①言語的・文化的学びと体験」「②自己目標や学習計画とふり返り」に分けられる。ここでも，自国とは異なる文化に対する意識を高めることが目標とされている。「(2)言語的・文化的体験の記録」に含まれる「①言語的・文化的学びと体験」は次のように説明されている。

言語的・文化的体験と学び　言語的・文化的体験を記録することによって，学習者は自分と異なる言語や文化に対する意識を高めることができます。それが，複合的な視野を持ったり，自文化について新しい視点や態度を得たりする，異文化理解能力の育成につながります。日本語を使う機会が少ない海外の場合でも，日本語で書かれたマンガを読む，インターネットで日本語の記事を読むなど，間接的な体験について記録することができます。(国際交流基金，2010a)

このように，JF スタンダードの評価基準は，CEFL 同様，多文化に対する理解能力，コミュニケーションへの対処能力を保持することを前提として作られていることがわかる。次に，JF スタンダードのモデルとなった CEFL と JF スタンダードの違いを確認しておきたい。次節で述べる JF スタンダードの課題は，CEFL との違いに遡る必要があると考えるためである。

4　CEFL と JF スタンダードの違い

国際交流基金では，CEFL と JF スタンダードの違いを，次のように表にまとめている(表 1 参照)。

表１には，JF スタンダードにはなく，CEFL では明記された能力として，「方略」，「テクスト」「コミュニケーション能力」があげられる。しかし，これは JF スタンダードが，場面想定型の統合的な「活動」に焦点をあてているためで，例えば，CEFL では「テクスト」というカテゴリーを作り，ノートをとるや要約などの能力をチェックさせるが，JF スタンダードではＢ２の「仕事と職業」「自由時間と娯楽」のテーマに「活動―作文を書く」が含まれているという具合である。

国際交流基金では，CEFR Can-do は汎言語的な記述で抽象性や包括性が高いが，日本語での言語活動が想定しにくい面があるため，JF スタンダードでは，日本語の使用場面を想定し，日本語での具体的な言語活動を例示した Can-do としたと説明している。Ａ１からＢ２までの活動 Can-do にトピックが付与し，言語活動の場面が具体的で，言語活動がよりイメージしやすく，抽象的な CEFR Can-do に比べ JF Can-do は日本語教育の現場で使いやすくなっているとされている。15のトピックは，「自分の家族」「住まいと住環境」「自由時間と娯楽」「生活と人生」「仕事と職業」「旅行と交通」「健康」「買い物」「食生活」「自然と環境」　「人との関係」「学校と教育」「言語と文化」「社会」「科学技術」に分けられる。

CEFL の CAN-DO リストもタスク（行動課題）と結びつけられているが，JF Can-do はこれを一歩進めたものとして評価されよう。ただし，CEFL は「Ｃ１」「Ｃ２」のレベルにも対応し，「Ｂ２」「Ｂ１」「Ａ２」はさらに二つのレベルに細分化されている。

CEFL の示す「言語能力と言語活動のカテゴリー」を簡単にまとめると次のようになる。これは先に図１で示した国際交流基金の「日本語スタンダードの木」にも継承され，わかりやすく表現されている。

「コミュニケーション言語活動」
　「活動－受容（理解する）産出（表現する）やりとり（相互行為）」話す，聞く等
　「方略（受容・産出・やりとり）」自分の発話をモニターしたり，議論に協力する能力等
　「テクスト」ノートを取る，要約等をする能力
「コミュニケーション言語能力」
　「言語構造的能力」語彙・文法・発音の知識等
　「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・帰納的能力）」場面に応じ，調整する能力

国際交流基金 CAN-DO 資料参照

表１　CEFL と JF スタンダードのレベル設定の違い

（国際交流基金 HP より）

	Can-do の種類				合計	レベル	備　　考
	活動	方略	テクスト	能力			
CEFR	319	46	15	113	493	Ｃ２ Ｃ１ Ｂ２ Ｂ2.1 Ｂ2.2 Ｂ１ Ｂ1.1 Ｂ1.2 Ａ２ Ａ2.1 Ａ2.2 Ａ１	CEFR Can-do は、全てのカテゴリーにあります。
JF	552*	0	0	0	552	Ｂ２ Ｂ１ Ａ２ Ａ１	*JF Can-do は次のカテゴリーにはありません。 ・聞くこと全般　・母語話者同士の会話を聞く ・読むこと全般　・話すこと全般 ・きくこと全般　・口頭でのやりとり全般 ・母語話者とやりとりする ・文書でのやりとり全般

CEFL と JF スタンダードの違いは，まず CEFL が 9 レベルを設定しているのに対し，JF スタンダードでは CEFL では細分化されていたＡ２，Ｂ１，Ｂ２のレベルをまとめ，Ａ１～Ｃ２の 6 レベル対応になっているという被検者の言語能力レベルの設定の違いがある。また，CAN-DO リストが，CEFL では言語能力別で構成されているが，JF スタンダードでは具体的な場面設定での CAN-DO に変更されていることがあげられる。しかし，多文化に対する知識や姿勢を測るという点については，共通している。

各国語版に先行して作られた英語版の CEFL に対し，ドイツ語，フランス語などでも CEFL に準拠した，TestDaF（Test Deutsch als Frem Dsprache），CECRL（Cadre Europeen Commun de Reference pour les Langues）が作成された。

なお，国際交流基金では特に触れていないが，CEFL には用意され，JF スタンダードに取り入れられていないのが，教員養成に関わる指導者，大学関係者を対象とした EPLTE（European Prolfile for Language Teacher Education －A Frame of Reference）である。教師支援のための EPOSTL（European Portfoiio for Teacher Student for Languages）は，日本語版として日本人英語教師用の JPOSTL もある。これと比べて，国際交流基金では JF スタンダード理解のための，資料，論文，会話ロールプレイのための動画などを用意しているが，これはいずれも現役教師のためのものである。（吉島，2013，参照）

本節では，CEFL と JF スタンダードの違いについて簡単に確認したが，次に実際に JF スタンダードを用い，日本語能力評価を行った先行研究では，どのような問題が指摘されたか，紹介したいと思う。

5　日本語スタンダードによる日本語能力測定法の評価と課題

JF スタンダードによる言語能力の測定は，これまでの文法・語彙能力中心の評価指標とは一線を画したものである。以前，私自身，ACTFUL の評価基準を参考に会話テストを行ったことがあったが，発話の流暢さが十分に測れないと感じた。ACTFUL は，学習者の発話の語彙，文法，文型に注目して評価を行うために，学習項目を丁寧に積み上げ，勉強しているかどうかを観察することができる。しかし，その反面語彙が少なく，文法が多少曖昧でも，聴解能力があり，流暢さに優れている学生を適切に評価することができないという印象を持った。一方，JF スタンダードは，タスク達成型であるため，流暢さ，聴解能力を観察することができる。その意味で，JF スタンダードは日本語教師が感じる「会話が上手な学生」の能力により近づいた基準を持つと推測される。

ここでは，JF スタンダードの評価指標を導入し，会話テストを実施した先行研究の結果と，この JF スタンダードで示された言語能力が，この会話テストを通してどの程度測れたかについて考えたいと思う。

（１）評価の安定性

評価は正確性、そしてどの状況でも同じような評価結果を出せる安定性を持つことが望ましい。テストに安定性を持たせるためには、様々な点から検討する必要があるが、ここでは評価段階の細分化について考えたい。長い期間日本に在住すれば、会話力が伸び、評価結果も当然上がる。一般的に、初級では比較的目に見える形で会話力を伸ばすことができるが、中級以降は会話力が伸びるのに時間がかかるという状況が、第二言語習得の発達段階には見られると言われる。CEFL では、Ａ２，Ｂ１，Ｂ２はそれぞれ「１・２」も分けられ（「表１」参照）、外国語能力がより細かく設定されている。初級では、伸びる速度が速いため、発達段階を細分化し、評価することが必要だが、現在の JF スタンダードの評価基準は中級以降、特に「Ｂ１」と「Ｂ２」レベルの開きが大きいように感じる。実際に JF スタンダードを会話テストに応用した、真島・山本（2017）は評価段階の設定の問題は、初級の「簡単なタスク」のレベルで起こっていると記している。また、真島・山本（ibid.）は JF スタンダードを用いた日本語会話

表 2　ロールプレイテスト判定結果（真島・山本，2017，による）

番号	ロールプレイ記録						判定
	1 回目	2 回目	3 回目	4 回目	5 回目	6 回目	
1	A 2 ◎	B 1 ○	B 1 ◎	B 2 △	B 2 ○	B 2 ×	B 1
2	A 2 ◎	B 1 ◎	B 2 ×				B 1
3	A 2 ◎	B 1 △	B 1 ×				A 2
4	A 2 ◎	B 1 ◎	B 2 ○	B 2 ○			B 2
5	A 2 ◎	B 1	B 1 ○	B 2 △	B 2 ○	B 2 △	B 1
6	A 2 ◎	B 1 ◎	B 2 ○	B 2 ○			B 2
7	A 2 ○	A 2 ○	B 1 ×				A 2
8	A 2 ◎	B 1 △	B 1 ○	B 1 △			A 2
9	A 2 ◎	B 1 ○	B 1 ◎	B 2 ×			B 1
10	A 2 ◎	B 1 ○	B 1 ◎	B 2 ×			B 1
11	A 2 ◎	B 1 ○	B 1 ○	B 2 ×			B 1
12	A 2 ◎	A 2 ◎	B 1 ○	B 1 ○	確定		B 1
13	A 2 ○	B 1 ◎	B 2 △	B 2 ○	B 2 △		B 1
14	A 2 ◎	B 1 ◎	B 2 △	B 2 △			B 1

能力の判定結果をまとめている（表 2，参照）。JF スタンダードのテストでは、「◎」が難なく達成、とされているが、何とか達成の「○」が二つの場合も、そのレベル相当と判定されている（データ番号 4・5・6・7・11・12）。つまり、下のレベルではないが、対象レベルにも達していないという状況が中級レベルでも起こっていると言えよう。CEFL と比べ、JF スタンダードでは、この中級レベル以降のレベルも簡略化されているが、CEFL ではさらに細分化されている B 2 以降の JF スタンダードの日本語能力段階の設定、及び評価方法には再考の余地があると考えられる。

（2）判定のしやすさ

周知のように言語力の正確さを測る ACTFUL はテスターになるのに，時間も費用もかかる。現実には取得の難しい資格である。これに対して，JF スタンダードは国際交流基金 HP で，テスター・マニュアルを含むすべての資料が公開されている。テスター・マニュアル自体は，詳細な説明が添えられており，かなりの厚さがある。誰もが，手に取りやすいという印象ではない。しかし，実際のテスト場面では，日本語教師の経験者であれば，授業のロールプレイタスクの延長として，A 1 ～ B 1 レベルでは，比較的容易に会話テストを行うことができる。表 3 のように，評価基準も極めてシンプルである。有益な評価テストであると実感できる。ただし，A 1 に達しないレベルはこの指標では測ることができないという問題がある。

表 3　被験者別 JF スタンダード評価の変化

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たたない

一方，B 2 以降のレベルでは，かなりマニュアルを読み込み，また，準備をしなければならない。例えば，B 2 はディベート課題である。テスターは被験者と反対の立場に立ち，反論しながら，ディベートを進めていく必要がある。国際交流基金がサンプルとしてあげているのは，「日本語の古い教科書を使うか」「メールか，直接話すか」「フレックス制への賛否」という 3 つの課題だが，被験者が複数いる場合は，多くの課題を用意する必要が

ある。併せて必要なのが，賛否両方の立場に立った反論の具体例である。さらに，Ｃ１は社会問題の現状と背景，解決策を多くの視点から述べるという課題で，課題数は多いが，具体的な議論展開のマニュアルがない。Ｂ２レベルでは，課題数を多くすること，Ｂ２，Ｃ１レベルともに，具体的な議論の進め方の例が多く記載されていれば，テスターも容易に会話テストに臨むことができよう。

（３）タスク達成能力の評価について

これまで，会話テストは主観で行うため，日本語能力テストに比して二次的な，参考資料程度の扱いである場合が多かった。しかし，JF スタンダードの登場により，会話テストの位置づけは大きく改善されたと言えよう。

ただし，いくつかの課題がある。課題を達成する時に，最小限の語彙で効果的に表現する学習者と，多くの語彙を用いて語れるが，論理的に話が進められず，話題が拡散してしまう学生，またＢ２では反論が苦手な学生など学生の個人差の影響があるだろう。もちろん，JF スタンダードが基本能力とする，社会言語能力，語用能力の問題が関わってくるが，ACTFUL で測られた語彙の豊かさや文法的正確さを，JF スタンダードテストに反映させるテスター側の工夫が必要であると感じる[注2)]。

なお，諸外国のスタンダード・テストについても，同じような問題点が指摘されている。パウル・ルッシュ（2007：145）は，Profile Deutsch（ドイツ語版 CEFL）の事例をあげている。

・Ｃレベルのための言語手段リストに関する質問－要求とは言わないが－がひんぱんに寄せられる。しかし，Ｃレベルに見合うだけの，恣意的ではない基準がほとんどみあたらない。合意できる共通の核というべきものがない。外国語のＣレベルでの使用は，かなりの程度まで，個別的である。例えば，ビジネスの分野でのコミュニケーションのニーズは，ドイツ語の教師になるためにドイツ語を学ぶ，といったときのニーズとは大幅に異なる。

・最端レベル（CEFR のＡ１レベルおよびＣ２レベルの一部）について評価が行なわれていなかったことがわかった。

Ｃ１レベルでは，社会問題について語ることが求められる。Ｃ１についてはある程度語れるが，Ｂ２のディベートでは説得力のある議論ができないということも起こりうる。議論が上手くできないということは日本人の外国語学習でも想定できる。Ｃ１の課題に，個人差によるバリエーションを反映できる工夫が必要であろう。例えば，JF スタンダードではディベートが求められるが，アジア圏の学生は初対面の，多くの場合，年上のテスターに反論することが，文化的なポライトネスの問題，つまり礼儀への違反行為と接触してしまう場合が起こる。このような学習者には，Ｃ１レベルで用意されているような，論述・プレゼンテーション式の課題が選択肢として用意されるべきであろう。もちろん，多文化共存のためには議論をする能力が必要だが，このような能力はある程度対象言語文化圏に慣れた上で，自然に身に着けていくことができる。対象言語の文化圏以外の外国に在住する学習者の言語能力を測ることも含めるなら，会話テスト自体の多文化性，すなわち多文化に対応する選択肢を準備することが必要であろう。

ここまで，日本語スタンダードが会話テストとして機能するか，またテスターとしての使いやすさはどうか，タスク達成能力を評価できるかという点をみてきた。Ａ１～Ｂ２，またＢ２以降のレベル，そしてＡ１以前のレベルについては課題が残るが，JF スタンダードは，留学生の日本語力を測るための有用なテストであると評価できよう。

最後に，テスター側の問題に触れたい。まず，被検者の自然な言語力を測るために，JF スタンダードを測る様々な会話場面を用意することが求められる。また，評価の揺れの問題に触れたが，課題を達成できる流暢さに加え，正確さを評価するためには，まずテスター側の「CAN-DO リスト」の読み込みが求められる。さらに，テスター

による誤差を減じるためには，インタビューアーに加え，複数の評価者で評価を行い，安定した評価に近づけることが必要であろう。（長坂・押尾，2013，鈴木，2005，他参照）

なお，リチャード・D・ブレクト（2007：140）は，Crozet & Liddycoat（1999）を引用し，各国語版スタンダードについて，次のような点を指摘している。

・文化は外からの浸透作用によって獲得されるのではない。それは教えられなければならない。異文化対応力を獲得するためには概念的および経験的な学習が要求される。そのため，文化についての深い概念的教授がカリキュラムとスタンダードの構成要素でなければならない。

さらにブレクト（ibid.：141）は，こうした文化体験は言語によってのみ可能となることを前提とし，「ACTFL（アメリカ外国語教育協会）のK-12アプローチ（５つのＣの有用性）」[注3)]，「文化に触れるための現地での学習の必要性」「異文化理解を進めるための研究，政策的展望と分析の必要性」に触れている。つまり，外国語学習における多文化理解は，外国語学習の場を，実際の文化体験へと広げていくことで可能になると考えるのである。こうした視点は，近年の「コミュニティとの協働学習をめざす日本語教育」の考え方に継承されることになる。（文化庁，2018，分科会資料参照）

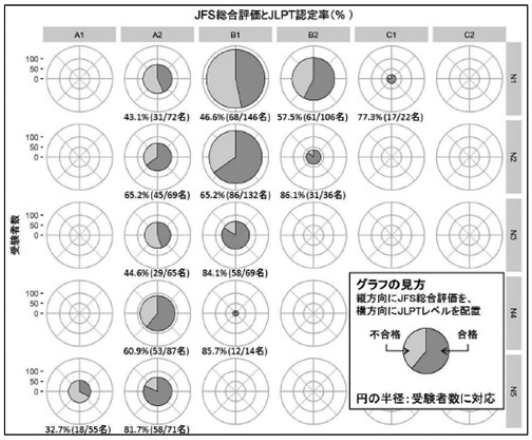
6　おわりに

本稿では，JF スタンダードによる会話テストの実践から，この会話テストの利点と課題について述べてきた。いくつかの課題を指摘したが，JF スタンダードの基準を用いることで，留学生の会話能力測定は実務的な意味で飛躍的に向上したと言えよう。まず，ある程度一定の基準による評価が可能になった。また，会話能力を日本語のクラス分けの参考にすることができる。また，個々人のレベルでは，会話能力評価を比較することで，会話力の伸びを確認することができる。CEFL 自体も多くの課題を指摘されているが，最新の会話能力測定基準を取り込んだ JF スタンダードの意義は大きい。

なお，会話能力の質的な評価課題については，本稿ではほとんど扱えなかった。別稿に譲りたい。

注

1）言語構造的能力を基盤とする日本語能力試験と，JF スタンダードの対応について国際交流基金は次のような資料を示している。



【図 1】　JFS 総合評価と JLPT 認定率（%）

2）長坂・押尾（2013）も，テスターによる評価の揺れを，また鈴木（2005）は，短時間のテストレーニングがマイナスに評価したことを報告している。なお，真島・山元（2017）は質的なレベル評価について考察しているが，調査で行われたレベル判定は，国際交流基金の開発者らが想定したレベルよりも甘かったと記している。

3）５つのＣとは，１ Communication コミュニケーション，２ Cultures 文化，３ Connections コネクション，関連づけ，４ Comparisons 比較，５ Communities コミュニティである。

参考文献

新井優子，和氣圭子，石上綾子，石田麻実，関崎博紀（2011）「Can-do statement を用いたタスク型オーラルテスト評価の一試案」『日本語教育方法研究会誌 18(2)』，日本語方法研究会，Pp.42-43.

押尾和美・長坂水晶（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)』お茶の水女子大学日本言語文化学研究会，Pp.70-73.

嘉数勝美（2011）「『JF 日本語教育スタンダード』がめざす日本語能力とは何か」『早稲田日本語教育学 8・9』Pp.107-113.

関崎友愛，古川嘉子，三原 龍志（2011）「評価基準と評価シートによる口頭発表の評価：JF 日本語教育スタンダードを利用して」『国際交流基金日本語教育紀要(7)』国際交流基金，Pp.119-133.

国際交流基金（2007）『日本語教育スタンダードの構築をめざす国際ラウンドテーブル会議録』

塩澤真季・石司えり・島田徳子（2010）「言語能力の熟達度を表す Can-do 記述の分析 －JF Can-do 作成のためのガイドライン策定に向けて－」『国際交流基金 日本語教育紀要(6)』Pp.23-39.

篠原亜紀（2017）「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」『言語学論叢(36)』Pp.28-46.

柴原智代（2007）「各国のスタンダード作成の意義と日本の課題―ヨーロッパ，米国，オーストラリア及び中国，韓国の比較・分析－」『日本語教育紀要(3)』国際交流基金，Pp.113-122.

島田めぐみ，野口裕之，谷部弘子，斎藤純男（2009）「Can-do statements を利用した教育機関相互の日本語科目の対応づけ」『日本語教育(141)』日本語教育学会，Pp.90-100.

島田めぐみ（2006）「Can-do-statements を利用した言語行動記述の試み－日本語能力試験受験者を対象として－」『世界の日本語教育(16)』，Pp.75-88.

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(11)』神田外語大学大学院紀要，Pp.77-94.

パウル・ルッシュ（2007）「"Profile deutsch"－多目的ツールを開発する」『日本語教育スタンダードの構築をめざす国際ラウンドテーブル会議録』国際交流基金，Pp.142-145.

松尾馨他（2006）「外国語の学習，教授，評価のためのヨーロッパ共通参照枠（CEF）の日本語教育における活用－ドイツ・ベルリンの中等教育ガイドラインの例－」『世界の日本語教育』国際交流基金，Pp.155-168.

牧野成一他（2001）『ACTFL-OPI 入門：日本語学習者の「話す力」を客観的に測る』アルク

真島知秀，山本淑乃（2017）「課題遂行能力の向上を重視した初級日本語学習－JF 日本語教育スタンダード準拠ロールプレイテストによる評価結果－」『琉球大学国際教育センター紀要(1)』琉球大学グローバル教育支援機構国際教育センター，Pp.39-52.

真嶋潤子（2010）「大学の外国語教育における CEFR を参照した到達度評価制度の実践－大阪大学外国語学部の事例を中心に」『外国語教育フォーラム(4)』金沢大学，Pp.3-12.

森本由佳子・塩澤真季・小松知子・石司えり・島田徳子（2011）「コミュニケーション言語活動の熟達度を表す JF Can-do の作成と評価－CEFR のＡ２・Ｂ１レベルに基づいて－」『国際交流基金 日本語教育紀要(7)』Pp.25-42.

山本弘子（2008）「日本語学校から見た評価の観点の見直し－ヨーロッパ共通参照枠の視点から－」『日本語教育(136)』Pp.38-48.

吉島茂（2014）「CEFR の日本の外国語教育への応用」『JACET 関東支部紀要(1)』大学英語教育学会関東支部，Pp.4-19.

Council of Europe，吉島茂・大橋理枝訳・編（2008）『外国語教育Ⅱ－外国語の学習，教授，評価のためのヨーロッパ共通参照枠－』初版第 2 刷，朝日出版社 =Common European Framework for Reference of Languages: earning, teaching, assessment. 3rd, John Trim, Brian North, Daniel Coste, 2002, Cambridge University Press.

来嶋洋美，柴原智代，八田直美（2012）「JF 日本語教育スタンダード準拠コースブックの開発」『国際交流基金日本語教育紀要(8)』国際交流基金，Pp.103-117.

リチャード，Ｄ．ブレクト（2007）「海外における日本語教育のためのスタンダードとテストに関する政策的展望」『日本語教育スタンダードの構築をめざす国際ラウンドテーブル会議録』国際交流基金，Pp.139-141.

参考 WEB サイト

JF 日本語教育スタンダード　http://jfstandard.jp

みんなの Can-do サイト　http://jfstandard.jp/cando

国際交流基金日本語国際センター20周年記念シンポジウム「JF 日本語教育スタンダード－その活用と可能性－」　http://www.jpf.go.jp/j/urawa/j_rsorcs/standard/index.html

『ヨーロッパの日本語教育の現状－CEFR に基づいた

日本語教育実践と JF 日本語教育ス タンダード活用の可能性－』2010年度 CEFR-JF 日本語教育スタンダード研修」論集 http://www.mcjp.fr/francais/langue-japonaise/seminaire-cecrl-jf-standard-for-2 55/seminaire-cecrl-jf-standard-for