

日語中級會話課程口試中導入 Rubric 評量法初探 —分析學習者的發話資料—

內田櫻

東吳大學日本語文學系博士生

摘要

Rubric 是近年來外語教學領域常見的評量方式之一。筆者將 Rubric 導入至台灣的大學日文系三年級會話課程中，並根據實踐情況及學習者的會話紀錄資料檢證 Rubric 作為能力指標的適切性。本稿為其檢證過程及結果。

筆者配合課程目標，在會話期中考口試以「能夠各自適切地對長輩、平輩、晚輩提出”請求許可”、”委託”及”道歉”。」為題，與學習者進行教師對學生的一對一角色扮演，使用本課程的 Rubric 表進行口語能力測試。

筆者設計的 Rubric 表參照 JF 日語教育 standard 制定了評量項目、能力指標、評量指標。利用 Rubric 能讓原本採相對性評量的口試成為客觀性的評量。其客觀性在於「測量項目」與「達成度的等第」清楚，能判斷出學習者能達到何種階段，發言中哪個項目有問題。但 Rubric 也存在著問題，設定的評量指標若太過抽象會造成評價上的困難，或所記述的評量指標無法契合能力內涵。

因此，筆者事後檢視學習者口試發言內容，並對學生訪談，檢證 Rubric 評量妥當性，並設計出對學習者日語會話有所助益的能力指標描述方法與射程內容。

關鍵字：中級程度、口頭考試、評價、Rubric、JF 日語教育 standard

受理日期：2016.03.18

通過日期：2016.05.13

Verification of Evaluation Method by Rubric in Oral Test : Utterance Data Obtained from Learners

Uchida Sakura

Doctoral candidate, Department of Japanese Language and Culture,
Soochow University

Abstract

Rubric is a way that use on evaluating the proficiency of foreign learners recently. The study took the utterance data from learners in the conversation class of Japanese department in Taiwanese university to verify the adequacy of rubric.

The author conducted a one-by-one roleplay between teacher and learner and set “I can request, ask a favor, or apology to elders, peer, or junior properly.” as the object in the midterm exam of conversation class. To develop the rubric, it referenced the JF standard for Japanese-Language Education and designed the evaluated items, proficiency and the object of evaluating. As a result, the rubric can evaluate more objectively what level the learners achieved or what problems learners have in the oral exam. On the other hand, the problem of the rubric is that the evaluating goal is too abstractive. Therefore, making the evaluating goal be more specific is indispensable. In the future, how learners take this evaluating activity is an issue. In addition, I would like to conduct a follow up interview, and design an evaluating activity that is helpful to the Japanese conversation of learners.

Keyword: intermediate level, oral exam, evaluating, rubric, JF Standard for Japanese-Language Education

中級レベルの口頭テストにおけるルーブリックを利用した 評価法の一試み —学習者の発話データをもとに—

内田さくら

東呉大学日本語文学系博士課程

要旨

本稿は、台湾の大学の日本語学科の会話授業において、近年外国語能力を評価する手段として用いられているルーブリックを導入し、実践及び学習者の発話データをもとにルーブリックの指標としての適切さを検証するものである。

筆者は、会話授業の中間テストにおいて、「目上の相手、対等/目下の相手に適切に許可を求める・依頼をする・謝罪をすることができる」という目標のもと、教師と学生の1対1のロールプレイを行った。ルーブリックを開発する際、JF日本語教育スタンダードを参照し、評価項目、能力レベル、評価の指標をデザインした。その結果、相対的になりがちだった口頭テストが、ルーブリックを利用することで、学習者の発話がどの段階まで達成できているのか、どの部分が問題なのかがわかりより客観的に評価を行うことができた。問題点として、評価指標の抽象度が高すぎると評価に困ることがあるため、より具体的な評価指標の記述が必要であることがわかった。今後は、このような評価活動を学生たちはどのようにとらえているのか、フォローアップインタビューを行い、学習者の日本語会話に役立つ評価活動をデザインしていきたい。

キーワード：中級レベル、口頭テスト、評価、ルーブリック、JF日本語教育スタンダード

中級レベルの口頭テストにおけるルーブリックを利用した 評価法の一試み —学習者の発話データをもとに—

内田さくら

東呉大学日本語文学系博士課程

1. はじめに

近年 CEFR や JF 日本語教育スタンダードを利用し、日本語を使って何ができるかという能力記述文の形で目標を掲げ実践を行っている研究が日本に限らず台湾でも発表されるようになった(羅 2012、陳 2014、内田 2015、黄 2015、周 2015)。これらは、目標—学習—評価の一貫性をなし、コミュニケーション能力の可視化を目指している。これまでほとんどがコースデザインに関する報告であり、評価法についての検証がまだ少ないようである。

本稿は台湾の大学の日本語学科の会話クラスで、CEFR、JF 日本語教育スタンダード(以下、JF スタンダード)を利用し、「～ができる」という形式の目標を挙げ、それに基づき学習活動をデザインしている。しかしその会話テストにおいて、特に中級レベルの学習者の口頭能力は、応答する力が伸び、発話量が多くなり、教師が本来考えていた評価基準がゆれてしまい、客観的に評価するのが難しいことがよくある。

本稿は、より信頼できる会話のテストを求めるため、近年外国語能力を評価する手段として用いられるルーブリックを開発し、さらに CEFR、JF スタンダードを利用してデザインした評価のあり方について、実践及び学習者の発話データを材料にルーブリックの指標としての適切さを検証するものである。

2. 先行研究

2.1 会話の評価方法

口頭能力を測るテストとしては、OPI(Oral Proficiency Interview)が開発され、インタビューを通して外国語能力が判定される方法として広く利用されている。レベルは、初級、中級、上級、超級の4つの主なレベルに分かれている。評価基準は「機能・タスク」、「場面・内容」、「正確さ」「テキスト・談話の型」という4つの観点から構成され、話題や語彙、文法項目は基本的に無限定である。テストは資格を持ったテストターが30分間のインタビューを行い、OPI終了時に録音をもう一度聞いて、評価が下される(牧野1991、中野2003、大西2006、近藤2012)。OPIは学習者の日本語口頭能力のレベルを判定する熟達度テストとしてよく利用されているが、OPIでは訓練を受けたテストターがレベル判定を行うこと、インタビューの時間が30分間かかってしまうことを考えると、そのまま導入することは難しいと考えられる。それに、それぞれのコースの目標や、活動に合った評価基準は担当者が決めなければならない部分がある。学習者たちと教師で作った教室文化があり、それに合っているテストティングは何かという課題が残る。

OPI以外の方法について、近年口頭テストの評価に注目した研究が見られるようになった(片岡1998、井之川・大場2002、関崎他2011、新井他2012、工藤2014)。これらの研究も、テスト前に評価基準を決め作成された評価表を使用している。その評価表の利用の仕方は様々であり、教師が使うためのもの(片岡1998、井之川・大場2002、新井他2012)、教師だけでなくTAにも使わせるもの(工藤2014)、学習者に自己評価や他者評価などに使わせるもの(関崎他2011)などがある。ここから評価は従来の教師が行うという一方的なものから、自己評価、他者評価のように多角的視点に開放されるようになっている傾向が見られた。この変化は学習者の内省、気づきに重きを置き、学習者主体に関心が寄せられるようになってきたとも言えるのではないだろうか。

2.2 評価表のデザイン

より到達度の高い教育が追求される中、評価表のデザインも工夫されている。テストの内容や方法によって、または使用者やテストの役割によって、多種多様なものが開発される。近年注目されてきているのが、ルーブリックである。

ルーブリックとは数段階に向けてレベルのめやすを記述し、達成度を判断する評価基準表である。「目標に準拠した評価」のための「基準」作りの方法論であり、学習者が何を学習するのかを示す評価基準と学習者が学習到達しているレベルを示す具体的な評価基準をマトリクス形式で示す評価指標である。これは、高次の認知過程や実践を可視化したものを評価するツールとして、パフォーマンスを評価する基準を示すものとして活用されている（當作・中野 2013、濱名 2012、山田他 2015）

ルーブリックをデザインする上で、近年 CEFR や JF 日本語教育スタンダードを利用したものが出てきた（関崎他 2011、新井他 2012、工藤 2014）。

CEFR とは、Common European Framework of Reference for Languages: Learning, teaching, assessment の略で、ヨーロッパの言語教育・学習の場で共有される枠組みである。その目的は、シラバス、カリキュラムのガイドライン、試験、教科書等々の向上のために一般的基盤を与えるものである。また言語学習者が言語をコミュニケーションのために使用するためには何を学ぶ必要があるか、効果的に行動できるようになるためには、どんな知識と技能を身につければよいかを総合的に記述するものである（吉島・大島 2004）。この CEFR を日本語教育の共通のスタンダードに作りかえられたものが、国際交流基金が開発した JF スタンダードである。

CEFR や JF スタンダードのような外的基準を用いてルーブリックを利用する効果としては、新井他 (2012) は、「外的基準を参照したことで、学習者間の相対評価になりがちだった評価方法が改善され、言語能力をより客観的に評価できるようになった」と述べ、さらに

関崎他（2011）は、「自己評価や他者評価のシートへの記述やディスカッションでの発言の中に文法や語彙などについての指摘ではない、発表全体の構成や態度などに関する様々な指摘が見られた」と述べている。つまり CEFR や JF スタandard を利用し、ルーブリックをデザインすると、相対的ではなく、より何を評価するのかがはっきりし、文法や語彙の正確さ以外の評価項目が増え、評価の幅が広がることがわかった。本稿でも、日本語の正確さは重要ではあるが、コミュニケーション能力を測る必要があるため、CEFR、JF Standard を利用する意義があるように考えられる。

2.3 ルーブリックを利用した評価の成果と問題点

実際にルーブリックを利用し、口頭能力を評価するとどのような効果があり、どのような問題点が議論されるのだろうか。以下、ルーブリックによる評価を検証したものを見ていく（井之川・大場 2002、関崎他 2011、新井他 2012、工藤 2014）。

井之川・大場（2002）では、群馬大学 2001 年度日本語予備教育の初級会話クラスで実施された会話テストにおいて、テストが学習者に有益な影響を及ぼすよう評価表をデザインしている。評価方法について、教師の実践及び学習者のアンケートにより検証を行っている。その結果、評価表に関する問題点として二つあげられた。

一つは、テストをその場で評価する際、評価項目の全ては評価できなかった。焦点をしばって評価をする必要がある。二つ目は、評価の基準が「満足」「不満足」という表記であり、これに対して学習者から点数化を望む意見があったが、評価とテストの実施方法への共通理解ができるよう、十分な話し合いを持つ必要があるということが指摘された。

関崎他（2011）では、国際交流基金で行っている教師研修での「総合日本語」授業における口頭発表の評価を JF Standard を参照し

評価シート¹をデザインしている。またルーブリックを学習者たちの自己評価、他者評価に利用している。この評価活動をアンケート、学生が実際に書いた評価シート、教師の目線から検証をしている。その結果、評価基準を参照し他者評価を複数回行う中でそれぞれの課題を自覚し、自ら次のパフォーマンスでの目標設定をしているようすが教師の目から見られた。またルーブリックを利用する際の留意点も述べている。

- (1) 研修参加者の評価活動のために必要な日本語への配慮
- (2) パフォーマンス全体とそれを支える観点の両方に注目させる工夫
- (3) 発表後のディスカッション
- (4) 参照できる客観的な指標の利用

新井他（2012）では、筑波大学留学生センターで、学習者の言語能力を内外に明確に示すことを目的とし、CEFR、JF スタンドアードの Can-do Statements を用いてルーブリック評価をデザインし、検討を行っている。その結果、「外的基準を参照することで相対評価からより客観的な評価になった」と述べている。しかし外的基準を用いる際、「記述文をそのまま引用しては、学習者の能力や『何ができるのか』を詳細に反映させることが難しかったため、項目の編成、配点比重の調整が重要なポイントである」とも指摘している。

工藤（2014）では、台湾の日本語学科の2年次の学習者が教師や日本人留学生と日本語を使う接触場面を想定した会話の授業を行っており、その活動の一つに、日本籍の TA を相手にロールプレイを行い、TA が達成度をルーブリックを使い評価するものがある。この研究では、学習者へのインタビューを行い、自己評価してもらうことで、到達度評価のシステムについて検証している。その結果、学習者たちは、一人を除き5人が研究者の評価とほぼ一致しており、学習者でもある程度の確に問題の分析ができることがわかった。ただ

¹関崎他（2011）では、ルーブリックによる評価表を「評価シート」としている。

し、評価表の基準の理解については、スラスラ言えないことやいいよどみを必要以上に低く、自己評価する傾向があり、どんなパフォーマンスがいいパフォーマンスなのか具体的に示しながら評価項目を再度詳しく説明する必要があることがわかった。

以上をまとめると、ルーブリックの成果としては以下のようにまとめられる。

〈成果〉

- ①ルーブリックは学習者自身の自己評価、他のクラスメート进行评估する他者評価においても利用可能である。
- ②自己評価や他者評価の際、学習者が評価基準を見ることで、自身のパフォーマンスを客観的にとらえる機会となっている。
- ③テストをその場で評価する際、評価項目の全てを評価することができなかった。焦点をしばって評価項目を決めるべきである。
- ④CEFR や JF スタンドアードの記述文をそのまま、学習者の能力や「何ができるか」を詳細に反映させることが難しいため、項目の編成、配点比重の調整が重要である。

〈留意点〉

- ⑤ルーブリックは、目標到達度を見せるものでもあり、学習者が理解できる記述文にする。
- ⑥教師が求めるよいパフォーマンスを具体的に示す必要がある。
- ⑦テスト後にルーブリックを学習者に返す際、点数の「0点」をなくしたり、点数のみでなく文章を追加など、フィードバックの仕方に配慮が必要である。

以上より、ルーブリックにより口頭能力を評価することは、学習者の自己評価にも利用できることから、教師が評価する際にも効果的であると考えられる。また関崎他（2011）、新井他（2012）でも述べられていたように、CEFR、JF スタンドアードなどの外的基準を参照してルーブリックをデザインすることも、より評価が客観的になる

ことがわかった。しかし、初級レベルの実践に集中しており、より進んだ段階へ向けるクラスの実践が少ないようである。筆者が担当した中級レベルの学習者は、一学期中、発話量が多く、内容も抽象度が高くなる。そのため、評価する内容が多くなり、評価項目も具体的に改めなければならない部分がある。また日本で行われた少人数の実践とは異なり、台湾の大学の日本語学科の会話クラスは、20人を越えるのが一般的である。筆者と同じく台湾の日本語学科の会話クラスで評価について検証しているのは、工藤（2014）のみであるが、工藤（2014）は、教師ではなく TA による評価活動を取り上げ、少し評価のスタイルが異なる。そのため、台湾においてさらに、ルーブリックによる評価活動を検討していく必要がある。

2.4 本稿の立場

本稿では、中級レベル²対象の口頭テストにルーブリックによる評価方法を採用する。評価基準は学習目標に準拠した内容、項目をデザインする。そしてルーブリックを使用する際、適切に学習者の発話を評価できるものであったのかをテストのとき録音した学習者の発話データを材料に適切さを検証する。なお本稿の評価活動で扱うテストは、日々の授業の達成度を測る到達度テストである。

3. 授業デザインと評価表作成の流れ

3.1 コースデザイン

この授業は、台湾の某大学の日本語学科3年次の20名の学習者³を対象にした「日語会話（三）」のクラスである。このクラスは、1週

² JF スタンドでは、CEFR の共通参照レベルを採用し「～ができる」(Can-do)形式で、学習者の日本語の熟達度の尺度を示している。その言語熟達度は、「A：基礎段階の言語使用者 (Basic User)」、「B：自立した言語使用者 (Independent User)」、「C：熟達した言語使用者 (Proficient User)」の3つの大きな段階に分かれている。そして各段階がさらに2つに分かれ、全部で6つのレベル (A1、A2、B1、B2、C1、C2) となっている。本稿で言う中級レベルとは、基礎段階から一歩進んだ「B 自立した言語使用者」のレベルに相当すると考えられる。

³ 4年制大学で、卒業時に日本語能力検定試験のN1をとることが卒業基準と要求される。3年生のレベルは、N3、N2、N1のいずれの合格者もいる程度である。

間に1回100分の授業であり、中間テスト、期末テストを含めて1学期で18週間ある。必修の授業であり、学習者たちはこの科目以外に聴解、作文、読解の授業を履修している。

1学期の授業のコースをデザインをする際は、国際交流基金(2010a:21)を参考にし、図1のような流れで行った。

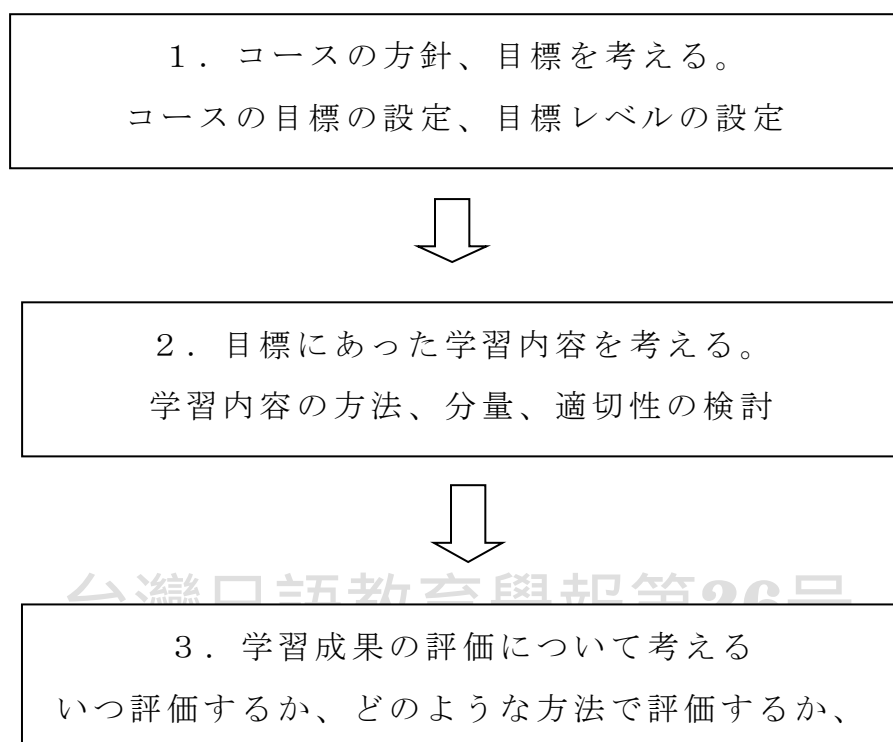


図1 コースデザインの流れ

本稿で扱う実践は、第2学期(2015年2月～6月)のものであり、学習目標は以下の通りである。

母語話者を相手に、お互いにストレスを感じさせることなく、普通の対話や関係が維持できる程度に、流暢に自然に対話できる・問題を解決することができる。(場面や相手、内容に応じて話すことができる。)

まず中間テストまでは、場面や相手(目上、対等)に合わせて、許可、依頼、謝罪などをできるようにする。中間テスト以降は、就

職面接のための日本語会話を取り扱った。

次に、これらの目標に沿った学習内容・実施の手順を考えた。以下の(1)～(5)が学習内容である。なお本稿では、中間テストのみを扱うため、学習内容の(4)(5)は本稿では扱わない。

- (1) 目上・対等の相手に許可を求めることができる。
- (2) 目上・対等の相手に依頼することができる。
- (3) 目上・対等の相手に謝罪することができる。
- (4) 説得力のある志望動機、自己PRを言うことができる/書くことができる。
- (5) 面接官からの質問に答えることができる。

以上の学習内容を振り分けた結果が、表1の通りである。

表1「日本語会話(三)」シラバス

1	2/26	オリエンテーション
2	3/5	許可求め
3	3/12	依頼
4	3/19	謝罪①
5	3/26	謝罪②
6	4/2	祝日のため休講
7	4/9	テスト前の総復習
8	4/16	中間テスト ⁴
9	4/23	中間テスト

3.2 評価のデザイン

3.2.1 口頭テストの概要

口頭テストは、2015年4月に、2週間にわたって行った。テスト形式は、教師と学習者の1対1のロールプレイである。その内容は、「許可」「依頼」「謝罪」に関して目上の相手、対等/目下の相手のそれぞれの場面で課題を与え、それらが、達成できるかを見た。学習

⁴ 中間テストを2回に振り分けたのは、1対1の口頭試験で1回では時間内に終わらないためである。

者には、まず初めに 3 枚のルールカードから 1 枚を引いてもらい、3 分間準備し、その後、教師を相手にロールプレイを行った。そのルールカードの一部の例が、以下の示す通りである。1 枚のルールカードに、目上の相手に対するものと対等/目下の相手に対するロールプレイのタスクが両方書かれている。なお口頭能力を測るためテスト前まで、ルールカードの内容は学習者には知らせなかった。

ルールカード例

依頼のタスク①

役割：あなたは学生です。相手は担任の先生です。

状況：あなたは日本留学のための申請書類を準備しています。その中で、志望動機の作文がありますが、日本語に自信がありません。その締め切りは 1 週間後です。

(1)先生に志望動機の日本語チェックを頼んでください。

(2)相手が許可しなかったり、はっきり答えなかったときは、相手が許可してくれるように説得してください。

依頼のタスク②

役割：あなたはレストランでアルバイトをしている学生です。小林もあなたと同じ年齢で、アルバイトの仲間です。

状況：あなたは急に明日仕事の面接が入りました。ですが、明日はアルバイトがあります。小林にアルバイトを代わってもらよう依頼してください。

(1)相手にアルバイトを代わってもらよう頼んでください。

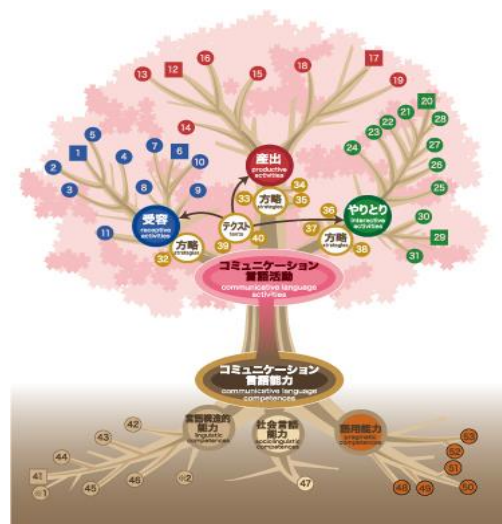
(2)相手が引き受けなかったり、はっきり答えなかったときは、相手が引き受けてくれるよう説得してください。

3.2.2 ルーブリックのデザイン

ルーブリックのデザインを行う上で、評価尺度には、包括的尺度

(holistic scale)⁵ではなく、分析的尺度(analytic scale)を利用する。それは、分析的尺度の場合、学習者がどのパフォーマンスレベルに到達したのかをいくつかの評価項目に分けて評定することで、被験者の長所や短所を分析でき、学習者に意味のあるフィードバックを与えることができるためである(近藤 2012)。

次に、国際交流基金(2010a)の「JF スタンドールの木」⁶を参考にした(図 2)。ここにはコミュニケーションを行うための言語能力として必要な能力が示されているとともに、言語活動も示されている。この中から、評価項目を決めるため、本稿と関わりがあると考えられる「語彙の使いこなし」「文法的正確さ」「社会言語的な適切さ」



「話題の展開」「話しことばの流暢さ」を選択した。これらをそのまま使うのではなく、「談話の展開」「語彙・文法・発音の正確さ」「流暢さ」とした。「社会言語的な適切さ」は、目上、対等/目下とそれぞれの相手に対して適切に対応できるかを見る必要があったため、「目上の相手に対する会話の適切さ」「対等/目下の相手に対する会話の適切さ」と2つに分けた。

図 2 JF スタンドールの木⁷

⁵ 包括的尺度は、学習者のパフォーマンスを包括的に評価するもので、分析的尺度は、評定項目ごとに評価するものである(近藤 2012)。

⁶ JF スタンドールは、「総理解のための日本語」を理念としている。相互理解のためには日本語を使って何がどのようにできるかという「課題遂行能力」と、様々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「異文化理解能力」が必要であると考えられる。この考え方に基づいて、各教育現場において学習者の「課題遂行能力」の向上を目指した教育実践を行いやすくするために、CEFR の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の関係を整理し、一本の木で表現した(国際交流基金 2010: 3)

⁷ <http://jfstandard.jp/summary/ja/render.do>

表 2 「JF スタンダードの木」と評価項目

JF スタンダードの木	本実践における評価項目
「語彙の使いこなし」「文法的正確さ」	「語彙・文法・発音の正確さ」
「話題の展開」	「談話の展開」
「話しことばの流暢さ」	「流暢さ」
「社会言語学的な適切さ」	「目上の相手との会話の適切さ」 「対等/目下の相手との会話の適切さ」

次に、能力レベル数を決めた。本実践では4段階に設定し、良い方から「4目標以上達成」「3目標達成」「2目標まであと少し」「1がんばりましょう」とした。

最後に、各レベルの評価指標を記述した。

「目上の相手との会話の適切さ」「対等/目下の相手との会話の適切さ」については、3種類のタスク（「謝罪」「依頼」「許可」）すべてを網羅させる必要があり、学習目標の「母語話者を相手に、お互いにストレスを感じさせることなく、普通の対話や関係が維持できる程度に、流暢に自然に対話できる・問題を解決することができる。」を考慮に入れ記述した。その際、「謝罪」「依頼」「許可」を行う場合、「迷惑の度合い」や「その場に適した理由」は非常に重要であると考えられるため、これらも記述に加えた。

「談話の展開」については、「謝罪」「依頼」「許可」を行うとき、一方的にそれらの形式や表現だけを言えばいいのではなく、会話の始まりから終わらせ方にも注意を払う必要があると考えたため、「会話の始まりから終わりまで自然な流れで会話を進めることができる。」（4目標以上達成）という記述文にした。

以上より、作成されたルーブリックが表3である。なお、ルーブリックは、テスト前に学習者に配布し、ルーブリックの見方及びどのようなパフォーマンスが、高得点がとれるのか説明した。

表 3 ルーブリック評価表

評価項目	4 目標以上達成	3 目標達成	2 目標まですと少し	1 がらまらぼう
手の相話の適切さ 目の適切さ	手迷いに由相レさに決が。相、合場理、トじ手解とる。のし度のたしス感上をこる。目上のそつ話をせず題るき(25) ⁸	手迷いに由問すで相、合場理、決が(20)のし度のたし解と。目上のそつ話をこる。目に惑やあを題るき	手な題き部合場ふな説で相、問で一度のにくや方(15)のしか決、のそ況し由仕上対と解が惑や状わ理のる。目にんをた迷いのさい得あ	手惑や状わ理の体な、決い。相迷いのさい得全礼と解い。のし、合場ふな説で、失ををな目上のそ況し由仕的印問で(5)
下の切目と適切 等相話の「 対の会さ	手迷いに由相レさに決が。相、合場理、トじ手解とる。のし度のたしス感上をこる。等対のそつ話をせず題るき(25)対に惑やあを手スせ問すで	手迷いに由問すで相、合場理、決が(20)のし度のたし解と。等対のそつ話をこる。対に惑やあを題るき	手、問で一度のにくや方。相てか決、のそ況し由仕のしと解が惑や状わ理のる。等対んをた迷いのさい得あ(15)対にな題き部合場ふな説で	手惑や状わ理の礼と解い。相迷いのさい得失ををい。のし、合場ふな説で、象題ない。等対のそ況し由仕なえ、決(5)対にのそ況し由仕なえ、決(5)
展開の談話	まわ然会るき始終自でめでのらでれ進が(20)話かま流をと。会りりな話こる。	まり少なあ難進が。始わてちが無をとる。の終いこ分、話こる。話やおぎ部が会るにめ(15)会りにしいるにめ(15)	まりぎが話こ。始わてさ会るる。の終いな、まあ話やおちり止が(10)会りにこあがと(10)	要て話か。にっ談なる。的言り、きたりす。一方をたでたり(5)一件いがった(5)
文の音発・彙語の正確さ	法誤な文のく(20)と音全(20)彙発は語と用い。	彙誤に意の。語の常、へし。と音非く通なと音非く通な法発はな疎影響(15)文と用少思影(15)	多く思影な多意の少(10)は、へは少(10)は、へは少(10)誤な疎響い。	彙多通語が疎あり。と語用思影(5)文法誤、影響に(5)
流暢さ	話な一て者どとピし話(10)んとス関語(10)とる、に母み。ほまくドは並	たさこがはつ暢るるに(8)つ流ちあ体的。ま流ちあ体的。つりがと全流暢。	は々、は度時が通(5)速、る疎(5)すくま思能。話遅つ意可	と黙う。こ沈まる。こ沈まる。つがしと(3)つがしと(3)

⁸ ルーブリックの各欄の記述文の後についている数字(25)は、点数の意味である。

3.3 授業の流れ

授業では、まず「許可求め」「依頼」「謝罪」の会話を目上の相手の場合、対等の相手の場合と両方見ながら、言語形式や談話の組み立てを確認した。その際、言語形式の様々なバリエーションを提示し練習した。その後、ルールカードを配布し、ロールプレイを行った。授業の最後に、ペアごとに発表するという流れであった。ロールプレイに対しては、状況に対し言語形式が適切であるかどうか、談話の流れがふさわしいかどうかをチェックし、フィードバックを与えた。

4. 結果

4.1 テスト結果

テストの結果は、最高点 93 点、最低点 45 点の平均 68 点だった。項目ごとの人数は表 4 の通りである。

表 4 評価項目ごとの人数

	4 目標以上達成	3 目標達成	2 目標まであと少し	1 がんばりましょう
目上の相手との会話の適切さ	1	9	9	1
対等/目下の相手との会話の適切さ	2	9	9	0
談話の展開	2	9	9	0
語彙・文法・発音の正確さ	0	10	9	1
流暢さ	0	14	6	0

表 4 より、学習目標と関係のある「目上の相手との会話の適切さ」と「対等/目下の相手との会話の適切さ」の両方とも、クラスの半分が目標達成できており、残りの半分が達成できていなかった。また談話の流れ、正確さ、流暢さについても、クラスの半分が目標達成していた。「正確さ」と「流暢さ」だけ、「4 目標以上達成」がいなかったが、これは「正確さ」では、「語彙と文法と発音の誤用は全く

ない。」としたこと、「流暢さ」については、「ほとんど詰まることなく、スピードに関しては母語話者並み。」と誤用が全くない、スピードが「母語話者並み」としたため、「4 目標以上達成」に判定できる人がいなかったと考えられる。

4.2 評価の行い方

実際、どのようにループリックを利用し評価を行ったのか。データをもとに説明していく。以下は、「依頼」のタスクを行った学習者 A とのロールプレイである。なお依頼をするタスクのロールカードは、3.2.1 を参照。

タスク 1 先生へ留学資料の日本語チェックの依頼をする。

S : 学習者 T : 先生

- 1⁹ S 先生、今お時間よろしいでしょうか。
- 2 T はい、どうしました。
- 3 S おそれはあります。お願いがあるんです。
- 4 T はい。
- 5 S 実はね、来年留学にする、行くつもりです。最近、留学の申請、しょ、書類を準備しています。その中で志望動機の作文が、日本語でしよるいがあります。しかし自分の日本語が上手ではなく、ちょっと自信をもってないですが、先生よろしかったら、ちょっと見てくれていただけませんか。
- 6 T そうですね、その作文の締め切りはいつなんですか
- 7 S 来週、来週の金曜日なんです。
- 8 T ぎりぎりですね。
- 9 S 先生はお忙しいのは承知の上で、気にいたします。
- 10 T いつもなら、担任なので、やってあげたいのですが、ちょうどその 1 週間は忙しくて、他の先生は頼めないですか。

⁹ 「1」とは、データ番号を指す。

- 11 S でも担任の先生、自分の状況をよく知りますと思っています。
まず日本語が上手な友達に、見てもらって、次先生に、チェックしてさせていただきませんか。
- 12 T それならいいかもしれませんね。でも3日間は忙しいので、四日後でもいいですか。
- 13 S はい、大丈夫です。ご迷惑をかけて本当にごめんなさいです、申し訳ございません。
- 14 T こちらこそすみませんね。
- 15 S 先生ありがとうございます。

まず「目上の相手への会話の適切さ」について、「今お時間よろしいでしょうか(1S)」や「おそれはいります。お願いがあるんです(3S)」などの前置き、「実はね、来年留学にする、行くつもりです。最近、留学の申請、しよ、書類を準備しています。その中で志望動機の作文が、日本語でしよるいがあります。しかし自分の日本語が上手ではなく、ちょっと自信をもってないですが、先生よろしかったら、ちょっと見てくれていただけませんか(5S)」というように、理由を言ってから依頼をしている。さらに、教師が忙しいと言った後に、「でも担任の先生、自分の状況をよく知りますと思っています。(11S)」と担任だから自分のことをよくわかっているという説得をし、「まず日本語が上手な友達に、見てもらって、次先生に、チェックしてさせていただきませんか。(11S)」と、相手の負担を考慮し、新たな提案を述べている。依頼が成功したあとに、再度、「ご迷惑をかけて本当にごめんなさいです、申し訳ございません。(13S)」と感謝の言葉を述べている。日本語の正確さが気になる所はあるが、全体的な流れとしては、迷惑の度合いを考慮し、その場の理由を話しており、何とか問題解決ができているため、「3目標達成」と判定した。

タスク2 アルバイトの仲間に、アルバイトを代わりに入ってもら

うよう頼む。

- 16 S 小林さん、明日、暇ですか。
- 17 T 明日、どうしたの
- 18 S アルバイトの時間をチェンジしてもいいですか
- 19 T 明日
- 20 S 4時間だけ。
- 21 T 私もあれなんだよね、
- 22 S 約束がありますか
- 23 T 3連勤目で、ちょっときついんだよね。
- 24 S 忙しい、疲れすぎて。
- 25 T うん
- 26 S 明日は、ずっと入りたいの会社の面接があります。そして行かなければならないです。小林さんは今回のアルバイトは手伝ってくれて、今回、今回はご飯をおごる、おごるよ。何でもいい。好きものだったら。
- 27 T お手伝いしてあげたいんだけどさ、私もあさってテストがあつて
- 28 S そっか。
- 29 T うん、勉強したいなと思って。
- 30 S じゃあ、2時間だけでいい。
- 31 T 2時間だけでいいの。
- 32 S 他の2時間は、他の友達を聞いて。
- 33 T 他の人に4時間やってもらえば。
- 34 S そうですね。じゃあ、じゃあ先の2時間やってもらって、面接終わったらすぐアルバイトの場所に帰って、えっと、チェンジして
- 35 T 楊さんが戻って来る
- 36 S そう、戻って来る。うん、お願い、大切な仕事です。
- 37 T うん、そっか。うんうん。2時間ならいいか
- 38 S あ、じゃああさってのテストのノートをかりますよ、かり、

- かり。テストのノートのコピーを書いて、かります
- 39 T かします。あ、そっか。でもあさってのテスト、楊さん受けるって
- 40 S 私、大丈夫よ。
- 41 T わかった。明日2時間だけならいいよ。
- 42 S 今回一緒にご飯を食べて。おごりますよ。
- 43 T 何でもいいのね。
- 44 S はい、どうもありがとうございます。

次に、「対等/目下の相手との会話の適切さ」について、「小林さん、明日、暇ですか。(16S)」と前置きで始まり、依頼をしている。その後、「明日は、ずっと入りたいの会社の面接があります。そして行かなければならないです。(26S)」と理由を話し、「今回、今回はご飯をおごる、おごるよ。何でもいい。好きものだったら。(26S)」と提案をしている。その後、「2時間だけ」という条件を提案し、本人が面接が終わったら戻って来るという形で終結させている。唐突な印象があるが、教師が「3連勤目で、ちょっときついなよね。

(23T)」と言った後に、「じゃあ、2時間だけでいい。(30S)」と時間を減らしている。さらに、教師が本来テスト勉強したかったという点を考慮し、「あ、じゃああさってのテストのノートをかりますよ、かり、かり。テストのノートのコピーを書いて、かります(38S)」とノートを貸すという提案をし、相手への迷惑の度合いを考慮しながら説得する姿が見られた。だが、前置きからすぐに、「アルバイトの時間をチェンジしてもいいですか(18S)」と依頼をし、「明日は、ずっと入りたいの会社の面接があります。そして行かなければならないです。小林さんは今回のアルバイトは手伝ってくれて、今回、今回はご飯をおごる、おごるよ。何でもいい。好きものだったら。」とここまで会話を見ると、教師側のことを考えずに一方的に要望を伝えているような印象がある。そのためここでは、「2目標まであと少し」と判定した。

「談話の展開」については、少しぎこちない部分があったが、無難に会話を進めることができていたので、「3 目標達成」と判定した。

「語彙・文法・発音の正確さ」については、以下の7つが挙げられた。ここでは、誤用を言って、再度言い直している場合は誤用と数えなかった。

3S 「おそれはいります (=いります)」

5S 「ちょっと見てくれていただけませんか (=見て)」

9S 「先生はお忙しいのは承知の上で、気にいたします」

11S 「チェックしてさせていただけませんか (=していただけませんか)」

26S 「ずっと入りたいの会社の面接があります (=入りたい会社)」

32S 「他の友達を聞いて (=他の友達に)」

38S 「ノートをかりますよ (=かしますよ)」

台湾日語教育學報第26号

以上の誤用は、誤用は非常に少ないわけではないが、多くもない。また全体的には前後の文脈により、誤用でも意味が通じ、意志疎通には影響がないと考えられるので、「2 目標まであと少し」と判定した。

「流暢さ」については、文字化資料からは判断することは難しいが、一部つまったりすることがあったが、全体的には話す速度は遅くなく流暢であったため「3 目標達成」と判定した。

以上より、学習者 A は、「目上の相手との会話の適切さ」は「3 目標達成 (20 点)」、「対等/目下との相手との会話の適切さ」は「2 目標まであと少し (15 点)」、「談話の展開」は「3 目標達成 (15 点)」、「語彙・文法・発音の正確さ」は「2 目標まであと少し (10 点)」、「流暢さ」は、「3 目標達成 (8 点)」と判定し、68 点となった。

4.3 ルーブリックの問題点

ルーブリックの評価指標の記述は、デザインの段階である程度、

想像して記述したが、実際に学習者の発話からはこちらが想像していなかった誤用が見られ、判断に困る点があった。それらを以下に見ていく。

4.3.1 「目上の相手、対等の相手との会話の適切さ」の記述について

(1) 言語形式の誤用

「許可」を求めるタスクで、上司に対し会議を欠席してもいいかと許可を求める際、「欠席していただきたいのですがよろしいでしょうか」、「欠席していただけますか」などの誤用が見られた。また「依頼」を求めるタスクで、教師に対し、研究計画書の日本語チェックを依頼する際「チェックさせてもらえませんか」「チェックしてもいいですか」などの誤用が見られた。本来ならば「欠席させていただきたいのですがよろしいでしょうか」、「欠席させていただけますか」、「チェックしてもらえませんか」、「チェックしてくれませんか」と言うべきところである。これらの誤用は、許可を求めるときの言語形式「～せていただけていただきたいのですがよろしいでしょうか」と依頼の言語形式の「～してもらえませんか」「～してくれませんか」が混同して使用したため、意味が全く変わってしまい、相手に対して失礼な印象を与えてしまった。ルーブリックにおいては、「一部迷惑の度合いやその場の状況にふさわしくない理由や説得の仕方である（3目標達成）」「迷惑の度合いやその場の状況にふさわしくない理由や説得の仕方ですべて失礼な印象を与え、問題を解決できない。（2目標まであと少し）」など、「理由」や「説得の仕方」について述べられているが、言語形式については述べられていなかった。そうすると、「語彙・文法・発音の正確さ」で評価することになるが、このような言語形式形式の誤用は、その他の誤用と少し種類が違い、タスク未達成につながる重大な誤用である。そのため「適切に伝えることができる」ということを目上、同等のそれぞれの記述に追加する必要がある。

(2) 相手を説得する内容

今回のテストでは、ただ許可、依頼、謝罪をするのではなく、相手は簡単に納得しないことにし、ロールカードには、「(2)相手が許可しなかったり、はっきり答えなかったときは、相手が許可してくれるように説得してください。」という指示を与えた。その結果、最初の部分はよかったのだが、こちらがすぐに納得しなかったので、説得の内容がその場にふさわしくない内容・現実的ではない内容が出てしまい、点数が低下してしまう人が多かった。

例〈T：教師 S：学習者 依頼のタスク①〉

S：あの先生

T：はい

S：今、ちょっとお時間よろしいでしょうか

T：はい、大丈夫ですよ。どうしましたか。

S：今、日本留学のためのしん、申請書類を準備しています。でも、志望動機の作文がありますから、日本語に自信がありません。志望動機の作文をチェックしていただけませんか。

T：あ、そうですか、いつまでですか。

S：1週間後で、この作文を完成します。

T：1週間後ですか。

S：ほんとに、お忙しいのは承知の上で参りました。

T：ちょっと、最近忙しくて、3日後に学会があって大変なんですよ。

S：あの、先生の日本語はあの、私は日本語に自信があります。私は日本に留学したいです。お願いします。お願いします。

例では、最初に教師へ適切に依頼をしており、さらに「ほんとに、お忙しいのは承知の上で参りました。」というように、迷惑の度合いも考慮して配慮の言葉も述べられていた。だが、こちらが納得しなかった後、「あの、先生の日本語はあの、私は日本語に自信がありません。私は日本に留学したいです。お願いします。お願いします。」と

「私は留学したいです。お願いします」と説得はしているが、その前の「あの、先生の日本語はあの、私は日本語に自信があります。」という部分が、よくわからない内容になってしまったため、点数を落とした。

説得する部分は、教師の反応に臨機応変に答えなければならないため、非常に難しい部分であると考えられる。そのため、評価項目は、言語形式と説得の部分で評価項目を分けるべきではないかと考えられる。

4.3.2 「談話の展開」の記述について

「談話の展開」については、2種類のタスクから総合的に判断することにしてはいたが、目上の相手と同等の相手によって、展開の仕方が一方がうまくいって、一方がうまくいかないということがあった。この場合は、やはりそれぞれの目上の相手の場合と同等の相手の場合それぞれの「談話の展開」を評価する必要がある。

4.3.3 「語彙・文法・発音の正確さ」の記述について

評価レベルの「3目標達成」と「2目標まであと少し」の「正確さ」の記述について、

〈3目標達成〉「誤用は非常に少なく、意志疎通への影響なし」

〈2目標まであと少し〉「誤用は多くなく、意志疎通への影響は少ない」

となっており、この2つの差が、「意志疎通への影響がないか少しあるか」という点であった。ただ実際に評価を行うと、どちらに評価した方がいいのかわからず、あまり差が見られないと考えられ、もう少し具体的な記述にする必要があると考えられる。

5. ルーブリックによる口頭能力の評価から得られた知見

ルーブリックを台湾の大学の中級レベルを対象とした会話テストに導入した結果、事前に評価項目・評価指標を決めておき、それら

を元に評価を行うことで、「なぜこの点数になるのか」と4.2で示したような学習者の発話について説明ができるようになった。学習者たちは、筆記テストの場合は、自分の日本語の問題点がはっきりわかるが口頭テストの場合は、それが難しい。またルーブリックと言っても、評価指標を記述せず5段階評価したものも見られるが、学習者は、ルーブリックを見た際に自分はどのレベルであるかはわかったとしても、何が問題だったのかはわからない。そのため、学習者に自分の発話がどの段階で、何が問題なのかを提示するためには、評価指標の記述が必要であり、ルーブリックは口頭テストに導入する意義があるということがわかった。

次に、ルーブリックをデザインする際、JFスタンダードを参照した。今回は、「JFスタンダードの木」を参照し、評価項目を選定した。テストの内容に合わせた結果、従来用いられてきた評価項目とあまり変化がなかったが、外的指標を参照することで、よりコミュニケーションに必要な評価の観点を知ることができた。また学習者のレベルが上がり発話量が増えれば増えるほど、彼らの発話もカバーしやすくなると考えられる。

4.3で示したように、今回は初の試みであり、実際にテストを行うと、筆者が予想しなかった発話が見られ、評価指標の記述とのずれがいくつか見られた。それらは、評価指標の抽象度が高すぎるものが原因であったため、より具体的な評価指標の記述は必要である。また実践者である教師は学生の運用能力向上を考えるため、学習者の口頭データを大切に対应的に修正していく必要がある。

6. ルーブリックの改善

ここでは、4.3で示したルーブリックの問題点をもとに、ルーブリックの修正を行った。

表5は、目上の相手との会話の評価表である。元々は、目上の相手との会話、対等の相手との会話を別に評価項目を立て、談話の展開、正確さ、流暢さは2つのタスクを総合して判定していたが、4.

3で述べたように、2つのタスクで、学習者の日本語の表現にも違いがでてしまったため、目上の相手、同等/目下の相手それぞれで、「談話の展開」「正確さ」「流暢さ」を判定することにした。また目上の相手に対する会話50点、対等/目下の相手に対する会話50点という点数の振り分けにした。

次に、「目上の相手との会話の適切さ」と「説得」は、元々「目上相手との会話の適切さ」と1つの評価項目であったが、許可、依頼を適切に言うことができず、さらにロールプレイの内容もテスト当日に知るため、説得することは、学習者にとって難しいため、評価項目を2つに分けた。

今後、この新しいルーブリックをもとに、再度実践を行っていきたい。

表5 目上の相手との会話の評価表¹⁰

評価項目	4 目標以上達成	3 目標達成	2 ほぼ目標達成	1 がんばりましょう
目上の相手との会話の適切さ	目上の相手に対し、迷いなく、適切な言い方を話し、問題を解決することができる。	目上の相手に対し、迷いなく、適切な言い方を話し、問題を解決することができる。	目上の相手に対し、なにか問題を解決できなかったが、一部迷惑やその場の状況にふさわしい理由や説得の仕方がある。	目上の相手に対し、迷いなく、適切な言い方を話し、問題を解決できない。失礼な印象を与え、問題を解決できない。
説得力	相手への迷いなく、適切な言い方を話し、説得力がある。	相手への迷いなく、適切な言い方を話し、説得力がある。	相手への迷いなく、適切な言い方を話し、説得力がある。	相手への迷いなく、適切な言い方を話し、説得力がない。
談話の流れ	会話の始まりから終わりまで自然	会話の始まりや終わりに少	会話の始まりや終わりに少	一方的に要件を言っており、相手

¹⁰ 対等/目下の会話の評価表は、表5の「目上の相手との会話の適切さ」の部分の「目上」を「同等/目下」に変えるのみであり、その他はすべて同じである。そのため、ここでは紙幅の関係で省略する。

	な流れで会 話を進め る。	しぎこちな いる部が、 に会話を める。	こちな あが止が、 とが、か と。	が話さ こるあ を何め る。	のこを考 え、て、な 滞す、ま す、た す、す と、が 多。
正確さ	語彙・文 法・発音 誤用は全 くない。	文法・語 彙・発音 誤用は非 常に少な く、疎通 の影へな し。	文法・語 彙・発音 誤用はあ る疎通の 影へ少な い。	誤用が多 く、意旨 通にも影 響がある。	
流暢さ	ほとんどの まるとな く、スピー ドに、母語 は母語話者 並み。	つまつた り、流暢さ がおちる とはある 全体的には 流暢。	話す速度は 遅く、時々 つまるとは 意疎通可能。	つまると沈 まるともあ る。	

7. 今後の課題

ルーブリックは、学習者自身が自分の口頭能力のレベルの特徴を把握することができるものであり、工藤(2014)のように、学習者にこの評価活動がどのように捉えられているのか、このような評価の方法は学習者の日本語会話に役立つものであるのか、フォローアップインタビューで明らかにしていくことを今後の課題としたい。

<付記>

本稿は2015年8月20日から21日にかけてサイモンフレーザー大学ハーバーセンター大学で開催された「CAJLE 2015 年次大会」において口頭発表を行い、その内容に加筆修正したものである。

参考文献

新井優子・和氣圭子・石上綾子・石田麻美・小浦方理恵・関崎博紀(2012)「タスク型オーラルテストにおける Can-do Statements を用いた評価方法の一試案」『筑波大学留学生センター日本語教育論集』27 p 81-103 筑波大学留学生センター

- 井之川睦美・大場美和子 (2002) 「コミュニケーション能力テストの実施と評価方法の開発—学習を促進させるテストをめざして—」『日本語教育方法研究会誌』9 p4-p5 日本語教育方法研究会
- 内田さくら (2015) 「A1 における Can-do 概念を導入した日本語授業の実践研究—生徒からのフィードバックを分析して—」『台湾日語教育學報』23 p324-352 台湾日語教育學會
- 大西博子 (2006) 「言語能力はいかにして評価すべきか:ACTFL-OPI における言語能力観の分析と考察をとおして」『言語文化教育研究』4 p17-30 早稲田大学大学院日本語教育研究科言語文化教育研究室
- 片岡朋子 (1998) 「海外で学ぶ初級日本語学習者のための口頭テストの作成」『東京家政学院大学紀要. 人文・社会科学系』38 p191-p203 東京家政学院大学
- 工藤節子 (2014) 「ロールプレイの到達度評価の検証—授業と評価の改善に向けて—」『台湾日語教育學報』23 p28-57 台湾日語教育學會
- 黄鈺涵 (2015) 「Can-do を取り入れた第二外国語授業のクラス活動と評価方法—学習者の自己評価に関する考察—」『台湾日語教育學報』25 p58-85 台湾日語教育學會
- 国際交流基金 (2010a) 『JF 日本語教育スタンダード 2010』[第二版]
- 国際交流基金 (2010b) 『JF 日本語教育スタンダード 2010 利用者ガイドブック』[第二版]
- 近藤ブラウン妃美 (2012) 『日本語教師のための評価入門』くろしお出版
- 周欣佳 (2015) 「JF 日本語教育スタンダードに基づく台湾の実業高校の部活動としての日本語授業の事例研究—PAC 分析を通して—」『東吳日語教育學報』44 p125-154 東吳大學日本語文學系
- 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 p119-p133 独立行政法人国際交流基金日本語教育方法研究会

- 陳姿菁 (2014) 「『Can-do』を利用した道案内の試み—会話分析を通して—」 Proceedings of the 18th Conference of The Japanese Studies Association of Australia: Peer-reviewed full papers (online), available from
 <<http://japaninstitute.anu.edu.au/projects/jsaa2013/papers>
 >2014 月 8 月 2 日アクセス
- 當作靖彦・中野佳代子 (2013) 『外国語学習のめやす—高等学校の中国語と韓国語教育からの提言—』公益財団法人国際文化フォーラム
- 中野はるみ (2003) 「ACTFL-OPI による一考察(1)」『長崎国際大学論叢』3 p73-81 長崎国際大学
- 仁科浩美 (2006) 「口頭テストにおける日本語教師の評価と発話データとの関係—研究留学生を対象とした形成的評価を例に—」『日本語教育方法研究会誌』13(1)p4-p5 日本語教育方法研究会
- 濱名篤 (2012) 「ループリックを活用したアセスメント」中央教育審議会高等学校教育部会
- 牧野成一 (1991) 「ACTFL の外国語能力基準およびそれに基づく会話能力テストの理念と問題」『世界の日本語教育』1 p15-32 国際交流基金、
- 山田嘉徳・森朋子・毛利美穂・岩崎千晶・田中俊也 (2015) 「学びに活用するループリックの評価に関する方法論の検討」『関西大学高等教育研究』6 p21-30 関西大学教育開発支援センター
- 吉島茂・大島理枝他 (訳編) (2004) 『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』朝日出版社 [Council of Europe(2002). Common European framework of reference for languages: Learning, teaching, assessment. 3rd printing. Cambridge: Cambridge University Press.]
- 羅素娟 (2012) 「『大学外文 (日本語)』の新しい試み—JF スタンドードに基づく初級日本語の授業—」『東吳外語学報』34 p137-152 東吳大學