

目次

はじめに.....	7
第 1 章	
言語テストの作成と使用に関する基本事項.....	12
Fundamental issues in developing and using language tests	
1.1. 「評価」をどう捉えるか.....	13
What is assessment?	
1.2. 診断的評価、形成的評価、そして総括的評価.....	16
Diagnostic, formative, and summative assessments	
1.3. 目標基準準拠テストと集団基準準拠テスト.....	18
Criterion-referenced and norm-referenced tests	
1.4. テストの妥当性と信頼性.....	21
Validity and reliability of tests	
1.5. 役に立つ言語テストとは.....	28
What is a useful language test?	
1.6. 構成概念としてのコミュニケーション能力.....	31
Communicative competence as a construct	
1.7. クラス内評価におけるシラバスの役割.....	36
The role of syllabus in classroom assessment	
1.8. 成績判定と報告.....	39
Determining and reporting grades	
第 2 章	
日本語クラスにおけるテストと自己評価の実践.....	44
Practical applications of tests and self-assessments in Japanese language classes	
2.1. 到達度テストの妥当性を高めるには.....	45
For enhancing the validity of achievement tests	
2.2. テストの出題方法と解答形式.....	48
Test question methods and response formats	
2.3. 口頭のパフォーマンス評価：手順と評定方法.....	52
Oral performance assessment: Procedures and rating methods	
2.4. 採点の一貫性と評定者訓練について.....	57
About score consistency and rater training	
2.5. 目標基準準拠テストの項目品質分析.....	60
Item quality analysis of criterion-referenced tests	
2.6. 自己評価の方法と役割.....	64
Approaches to and roles of self-assessments	

第3章

日本語プログラムの学習成果アセスメント.....	70
--------------------------	----

Assessing learning outcomes of Japanese language programs

3.1. プログラム対象の学習成果アセスメント：現状と動向.....	71
Assessing program-level learning outcomes: Current situations and directions	
3.2. プログラム対象の学習成果アセスメント：目的と多面的役割.....	75
Assessing program-level learning outcomes: Purpose and multiple roles	
3.3. プログラム対象の学習成果アセスメント：計画と具現化.....	78
Assessing program-level learning outcomes: Planning and implementation procedures	
3.4. プログラム対象の学習成果アセスメントのためのアンケート調査：事例研究.....	81
A survey for assessing program-level learning outcomes: A case study	
3.5. 直接的な学習成果アセスメント：日本語作文テストと採点に関する問題点.....	85
Direct learning outcomes assessment: Japanese composition tests and scoring issues	
3.6. 間接的な学習成果アセスメント：プログラム修了アンケートの方法と用途.....	90
Indirect learning outcomes assessment: Procedures and uses of exit surveys	

第4章

日本語の能力テストとプレースメント・テストを利用した評価.....	94
-----------------------------------	----

Assessing with Japanese proficiency and placement tests

4.1. ACTFL 言語能力基準と能力テスト.....	95
ACTFL proficiency guidelines and proficiency tests	
4.2. 能力(熟達度)テストとしての ACTFL OPI.....	98
The ACTFL OPI as a proficiency test	
4.3. ACTFL OPI の妥当性と信頼性に関する先行研究.....	100
Previous research on the validity and reliability of the ACTFL OPI	
4.4. 日本語 ACTFL OPI を利用した口頭能力の評価.....	104
Assessing Japanese oral proficiency using the ACTFL OPI	
4.5. ACTFL 言語能力基準に基づく日本語読解力の評価.....	108
Assessing Japanese reading proficiency based on the ACTFL proficiency guidelines	
4.6. 日本語のプレースメント・テスト.....	114
Japanese language placement tests	
4.7. 項目応答理論を用いたプレースメント・テスト項目の分析.....	118
Item analysis of placement tests using Item Response Theory	

第5章

年少者対象の日本語プログラムの評価.....	124
------------------------	-----

Assessing Japanese language programs for young learners

5.1. 年少者対象の日本語プログラムの評価方法.....	125
Approaches to assessing Japanese language programs for young learners	
5.2. 小学校における学習成果アセスメントの計画と実施.....	129
Planning and implementing learning outcomes assessments at an elementary school	
5.3. 年少者用の日本語口頭テストと採点方法の開発.....	132
Developing Japanese oral tests and scoring methods for young learners	

5.4.	口頭テスト実施中の試験官と被験者との相互作用	135
	Tester-candidate interactions during oral tests	
5.5.	口頭テスト結果の追跡(縦断的)分析の仕方	139
	Longitudinal analysis of oral test results	
5.6.	口頭テスト、聴解力テスト、及び自己評価結果の比較	142
	Comparing oral test, listening test, and self-assessment results	
第 6 章		
	日本語教育実習における評価	146
	Assessing in a Japanese language teaching practicum	
6.1.	教育実習の目的、内容、及び評価.....	147
	Purpose, content, and assessment of a teaching practicum	
6.2.	ティーチング・ポートフォリオとは	151
	What is a teaching portfolio?	
6.3.	ティーチング・ポートフォリオの評価方法	153
	Approaches to assessing teaching portfolios	
6.4.	ティーチング・ポートフォリオ評価の有用性について	155
	About useful portfolio assessment	
6.5.	授業観察及びフィードバックの目的と手順.....	157
	Purposes and procedures for class observations and feedback	
6.6.	実習中のフィードバックの与え方と受け止め方.....	159
	How to give/receive feedback during a teaching practicum	
巻末資料.....		
	163	
資料 1-1A.	日本語使用及び口頭能力に関する自己評価	164
資料 1-1B.	日本語使用及び口頭能力に関する自己評価(和訳)	165
資料 1-2.	日本語解説文の読解力テストの見本.....	166
資料 1-3A.	ニーズ分析アンケート(学生用)の見本	168
資料 1-3B.	ニーズ分析アンケート(学生用)の見本(和訳)	169
資料 1-4.	Project work in Japanese(日本語のプロジェクト・ワーク)のシラバス	170
資料 1-5.	中間・期末成績レポート用紙の見本.....	173
資料 2-1.	テスト項目細目一覧の見本(単元テスト用)	174
資料 2-2.	単元テスト用口頭タスクカードと評価用紙(教師用)の見本	178
資料 2-3.	単元テスト用筆記問題の見本(聴解力・読解力問題を含む)	179
資料 2-4.	資料 2-3 の聴解力テスト問題に使われた録音会話	181
資料 2-5.	単元テスト用筆記問題(資料 2-3)の採点基準	182
資料 2-6.	初級日本語クラス用書きタスクの見本	183
資料 2-7A.	初級日本語クラス用書きタスクの採点ルーブリックの見本(分析的尺度)	184
資料 2-7B.	初級日本語クラス用書きタスクの採点ルーブリックの見本(分析的尺度)(和訳)	185
資料 2-8.	上級日本語クラス用読解力テストの見本	186
資料 2-9.	初級日本語クラス用口頭テスト(ロールプレイ・カードとサンプル・パフォーマンス)	188
資料 2-10A.	初級日本語クラス用口頭テストの評定尺度の見本(包括的尺度)	190
資料 2-10B.	初級日本語クラス用口頭テストの評定尺度の見本(包括的尺度)(和訳)	191
資料 2-11A.	初級日本語クラス用口頭テストの評定尺度の見本(分析的尺度)	192
資料 2-11B.	初級日本語クラス用口頭テストの評定尺度の見本(分析的尺度)(和訳)	193

目次

資料 2-12.	日本語二年課程のプログラム評価用読解力テストの見本	194
資料 2-13A.	上級日本語クラス用口頭発表の自己評価シートの見本	196
資料 2-13B.	上級日本語クラス用口頭発表の自己評価シートの見本(和訳)	198
資料 2-14A.	初級日本語クラス用ウィークリー・ジャーナルの自己評価シート	200
資料 2-14B.	初級日本語クラス用ウィークリー・ジャーナルの自己評価シート(和訳)	201
資料 2-15.	インタビュー・タスクとインタビュー協力者用評価シート	202
資料 3-1.	必修科目としての第二言語教育の評価計画アンケート(教師用)	204
資料 3-2.	レポート提出チェックシートの見本	207
資料 3-3A.	日本語小論文の評価シートの見本(クリティカル思考の評価を考慮)	208
資料 3-3B.	日本語小論文の評価シートの見本(クリティカル思考の評価を考慮)(和訳)	209
資料 3-4A.	日本語作文評価シート(Japanese Composition Scoring Sheet)	210
資料 3-4B.	日本語作文評価シート(Japanese Composition Scoring Sheet)(和訳)	211
資料 3-5.	ハワイ大学言語・言語学・文学部のプログラム修了アンケート(選択式設問のサンプル)	212
資料 3-6.	ハワイ大学言語・言語学・文学部のプログラム修了アンケート(記述式設問のサンプル)	213
資料 3-7.	日本語プログラム修了アンケートの設問の見本: プログラムに対する意見	214
資料 3-8.	日本語プログラム修了アンケートの設問の見本: ACTFL OPI を使用した評価に関する質問	215
資料 4-1A.	基礎レベルを超えた日本語カリキュラムの枠組み [草案]	216
資料 4-1B.	基礎レベルを超えた日本語カリキュラムの枠組み [草案](和訳)	217
資料 4-2.	OPI 初級から超級までの基準の概略	218
資料 5-1.	年少者用の日本語口頭テストの見本(タスク別)	220
資料 5-2A.	タスク共通尺度(Task-independent scale)	222
資料 5-2B.	タスク共通尺度(和訳)	222
資料 5-3A.	タスク別尺度(Task-dependent scale)の見本(挨拶と自己)	223
資料 5-3B.	タスク別尺度の見本(挨拶と自己)(和訳)	223
資料 5-4.	SPSS(version17)を使用した反復測定一元配置分散分析(One-way repeated measures ANOVA)	224
資料 5-5.	SPSS(version17)を使用した対応のある二標本の <i>t</i> 検定(paired <i>t</i> -test)	228
資料 5-6.	年少者用の日本語聴解力テストの見本	230
資料 6-1.	日本語教育実習のシラバスの見本	232
資料 6-2A.	教科書分析基準の見本	234
資料 6-2B.	教科書分析基準の見本(和訳)	235
資料 6-3.	日本語教育実習用の教案用紙の見本	236
資料 6-4.	日本語授業の観察用紙の見本	237
資料 6-5A.	日本語教育実習の学期末授業評価用紙	238
資料 6-5B.	日本語教育実習の学期末授業評価用紙(和訳)	240
資料 6-6A.	日本語学習者対象の教育実習生評価のためのアンケート用紙	242
資料 6-6B.	日本語学習者対象の教育実習生評価のためのアンケート用紙(和訳)	244
資料 6-7.	日本語教育実習ポートフォリオの評価基準	246
引用文献リスト(Works cited)		248
主要用語リスト(Keyword list)		267

第1章

言語テストの作成と使用に関する基本事項

Fundamental issues in developing and using language tests

● 2章[2.6]⇒64 ページ

日本語クラスにおける評価で、教師が一番時間をかけているのは、到達度テストを使った評価でしょう。小テスト、単元テスト、中間・期末試験など、教師は、数多くのテストを作成し採点しています。教師なら誰もが、「テストの採点に明け暮れた休日」の経験を持っているのではないのでしょうか。最近注目されている、テストを使用しない「代替的アセスメント」(2章 [2.6])をクラス内評価に導入している教師でも、到達度テストを全く使わないという人は、あまりいないのではないかと思います。

クラスで定期的テストを行う第一の目的は、学習者の目標達成度を適切に把握し、その情報に基づいて、成績判定や学習指導をすることです。また、適切なテストを定期的に行うことにより、学習者の学習意欲を高め、教師自身も、日本語指導法や学習教材の適切さについて省察できます。逆に、適切とはいえないテストの作成や採点を行っていると、学習者の成績や、日本語指導の効果が正確に把握できません。それだけでなく、「学生の評価をいい加減にしている先生」として、学習者の学習動機の低下を招き、さらには、教師としての信頼を失いかねないでしょう。

1章では、基礎的な評価理論や言語テスト理論を踏まえながら、テストを使った適切な評価とは、どういう評価なのかについて、理論と実践の両面から考えます。まず、評価の定義に始まり、続いて、評価の目的を診断的評価、形成的評価、そして、総括的評価の観点から述べます。次に、目標基準準拠テストと集団基準準拠テストの基本的な違いについて説明します。さらに、テストの妥当性、信頼性、そして、有用性を取り上げ、言語テストの作成、実施、そして、テスト結果の解釈に関する数々の基本事項を取り上げます。テストの妥当性の関連事項として、構成概念である「コミュニケーション能力」に対する理解がどのように発達し、それが口頭言語能力の評価にどのように影響してきたのかについても考察します。最後に、クラス内評価におけるシラバスの役割と成績判定の方法を取り上げます。

第1章 の目標

1. 評価の定義と目的について把握する。
2. 診断的評価、形成的評価、そして、総括的評価についての理解を深める。
3. 目標基準準拠テストと集団基準準拠テストの違いが、さまざまな角度から

説明できる。

4. 言語テスト、特に到達度テストの妥当性、信頼性、そして、有用性についての知識を深め、現場で応用できる。
5. 構成概念としての「コミュニケーション能力」に対する理解を高め、妥当性を考慮しながらコミュニケーション能力の評価ができる。
6. シラバスの機能を評価の観点から理解し、授業に活用できる。
7. クラスで行った絶対評価のデータ処理の仕方を学び、適切な成績判定ができる。

1.1. 「評価」をどう捉えるか

What is assessment?

日本語でいう「評価」という言葉には、**アセスメント**(assessment)と**エバリュエーション**(evaluation)の二つの意味があります(新版日本語教育事典 2005, p. 778)。「アセスメント」は、主に、学習過程と学習成果に関する評価のことを意味します(Norris, 2006; Wright, 2006)。テスト、ポートフォリオ、自己評価、観察など、さまざまな評価ツールを使って、学習目標に照らし合わせながら、学習に関する評価を行い、そして、その結果を指導強化や学習向上のために使用していく過程を「アセスメント」と呼びます。このような意味でのアセスメントは、カリキュラム開発や指導法改善に欠くことのできないものであり、現場で常に行われるべき教育活動の一つです(Bailey, 1998; Graves, 2000; Leskes & Wright, 2005)。「**学習成果アセスメント**(learning outcomes assessment)」という言葉に象徴されるように、教育現場における学習成果を重視したアセスメントの積極的な導入は、近年になって特に強調されています(3章 [3.1])。

一方の「エバリュエーション」は、アセスメントより広域に捉えられることが多く、教育の質や価値を維持もしくは向上するのに必要な評価のすべてが、その対象となりえます(図1-1参照)(Brown, 1995)。学習に関する評価(=アセスメント)はもとより、新入生や卒業生に関する調査から、受講登録状況、学生アドバイス、学生への経済的援助、教育環境設備に関する調査、そして教員評価に至るまで、教育に関する情報のすべてが「エバリュエーション」の対象となりえるのです。また、アセスメントでは、学習に関する評価の結果をカリキュラム改善や学習向上に生かすことに重きが置かれるのに対し、エバリュエーションは、評価結果を基に何らかの判定や意思決定を行うという点が強調されます(例. 教育の質保証認定[accreditation]、学位課程新設の認定、雇用契約更新、助成金の更新など)(Bachman, 1990; Brown, 1995)。もちろん、アセスメントも、評価結果に基づいて、授業やカリキュラムに関する事項について、何らかの判定や意思決定を行いますが、それよりも評価結果をカリキュラム改善や学習向上

● 3章[3.1] 71 ページ



話し合おう

1. アセスメントとエバリュエーションの違いは何ですか。本書においては、「評価」という言葉は、主にどのような意味で使用されていますか。
2. 診断的評価、形成的評価、そして、総括的評価とは何ですか。
3. 学習者を支援するための評価という観点からは、どの評価が最も大切ですか。その理由は何ですか。
4. 目標基準準拠テスト(CRT)と集団基準準拠テスト(NRT)の違いを挙げてください。
5. 言語テストの妥当性にはどのような種類がありますか。
6. どのようにすれば、到達度テストの妥当性を高めることができますか。
7. テストの信頼性とは何ですか。信頼性はなぜ大切なのですか。テスト結果の信頼性を高めるために、教師はどのようなことに気をつけなくてはなりませんか。
8. コミュニケーション能力の理論的枠組みとして、代表的なものに何がありますか。その共通点及び違いは何ですか。
9. テストの有用性を高めるには、どのような配慮が必要ですか。
10. シラバスの役割を、評価という観点から述べてください。
11. 日本語クラスで、テスト結果を相対的に解釈する方法で成績判定を行うことの問題点を話し合ってください。
12. 成績判定の報告をする時に、どのようなことに気をつけなくてはなりませんか。



プロジェクト・アイデア

1. 現役の日本語教師や研修生が、「アセスメント」をどう捉え、また、どのような方法で行っているのかを調査してみましょう。そして、その共通点や問題点を分析してください。
2. ある特定の日本語クラスでどのような診断的評価、形成的評価、そして、総括的評価が行われているのかを調べてみましょう。
3. 日本語の到達度テストを一つ選んで、学習目標や指導項目と照らし合わせながら、その構成概念的妥当性、内容的妥当性、結果的妥当性を検討してみましょう。
4. 日本語能力テストを一つ選び、NRTとしての妥当性、信頼性、有用性に関する分析がどのように報告されているかを調べてみましょう。
5. 日本語クラスのシラバスを入手し、評価に関する説明も含めて、内容の適切さを検討してみましょう。また、シラバスの内容に沿ったクラスの成績表を作成してみましょう。
6. エクセルを使って実際の(もしくは架空の)成績データの処理をしてみましょう。