

学習者オートノミーのためのルーブリック： 入門期の中国語の授業を例に

山崎直樹

(関西大学外国語学部／外国語教育学研究科)

ymzknk@kansai-u.ac.jp

第32回広島大学 外国語教育研究センター 外国語教育研究集会

シンポジウム「パフォーマンス課題の作成と実践・評価：

自ら学び、考え、表現する力を育成する新しい外国語教育のために」

2024年3月8日、オンライン

今日のポイント

参照したもののなど【↓】

[https://scrapbox.io/
2024-03-08/INDEX](https://scrapbox.io/2024-03-08/INDEX)

1. ルーブリックの松竹梅
2. 評価のブラックボックス化
3. パフォーマンス課題とルーブリックの実例
4. 「できばえ」を評価する評価方法の放棄



ルーブリックは何をするためのものか

教師が重視したい点を反映でき、細かく記述すれば、学習者に「何を」「どのように」学ぶべきか、「どのくらい」できたら目標達成といえるのか、を明確に示すものとなる。

つまり学習者を測るもののさしとしてだけでなく、学習の指針を示すものとなる。

(植村麻紀子先生談)

FAQ:

ループリックは「**評価の信頼性**」を増すためのものではないのですか？

山崎の実験によると、「**評価の信頼性**」に関しては、特筆するほど効果があるかどうかは疑問

中国語話者の大学院生数人に、複数の学習者のパフォーマンスの動画を見せて、「1~10点の間の点数で評価せよ」という課題を与えたところ、評者間の食い違いはあまりなかった。ループリックを使ったところで、評者間のばらつきは大差ないのでは？

「**評価の妥当性**」を増すためのものとして考えるのがよいかも

ループブリックの松竹梅

梅

竹

松

学習者が、何を評価されるかを理解できる

+

+

+

学習者が、どこまでできればよいかを納得できる

+

+

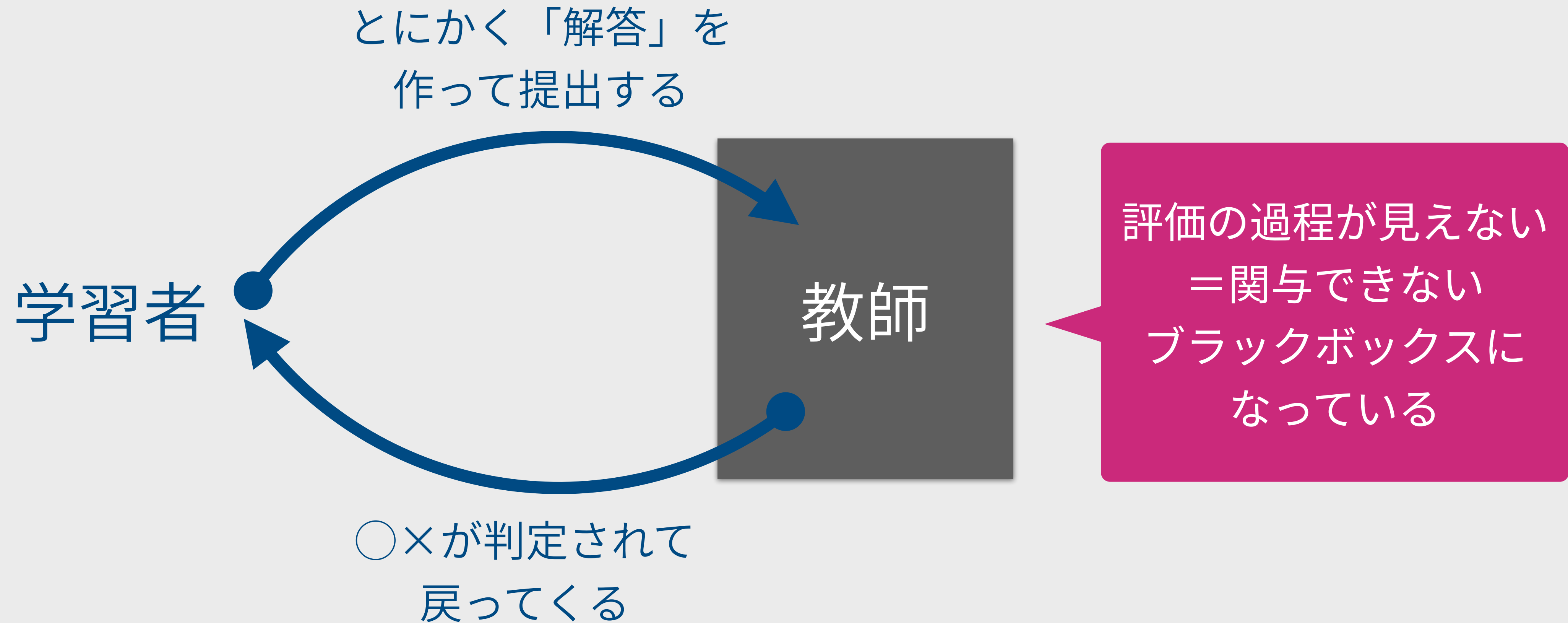
学習者自身が、評価をおこなえる

+

なぜ、学習者自身が評価をおこなえることが重要なのか？

- 入門～初級レベルの外国語の学習では、「評価」がブラックボックスになっているから

評価のブラックボックスモデル



- 評価のブラックボックスモデル：自分が知らないことを知っている（らしい）というだけで、ある者の**権威**を認め、それを「学習者のアウトプットを受け取る→検証する→学習者にフィードバックを与える」というサイクルを実行してくれるブラックボックスとして利用する学習モデル
- このブラックボックスはなくならないかもしれない
- 言語を教える教師の多くにとっては、教師としてのprofessionalityの主要な部分かもしれないので（特に母語話者教師）
- 📖 山崎（2018; 2023）

なぜこのようなことにこだわるのか？

- 学習者のオートノミー autonomy
- 何が「正しい」情報であるかを判断する権威は誰が持つべきなのか（青木, 2018)

パフォーマンス評価で心がけていること

- コンテンツ（については、今日は省略させていただきます）
- 言語形式
 - intelligibility, comprehensibility を評価する
 - accuracy, well-formedness は、形成的にフィードバックをおこなう機会があるときのみ、評価する
- メタ認知（学習者自身が自分のパフォーマンスをどうとらえているか）

タスク(1) 「～での理想の休日」

- 留学に行きたい大学を決めてください。
- その大学がある街を確認してください。
- 留学してその街に住んでいる自分を想像してください。
- あなたは、その街で、休日をどう過ごしますか？何時に起きる？起きてから何をする？近所の／近郊のどこかに出かける？昼ごはんは何を食べる？どんな休日の過ごし方が理想的だと思いますか？
- その街でのあなたの理想の休日の日課を、中国語で語ってください。


提出物（録音した音声）の構成

1. 中国語でパフォーマンスをする（氏名を言う、テーマを言う、日課を紹介する）
2. うまくできたところ、できなかったところを、日本語でコメントする

	目標以上(4pt)	目標達成(3pt)	あと一歩(2pt)	さらなる努力を(1pt)
音声 (時刻)	ストレスなく正確な時刻を聞き取れる。	多少のストレスはあるが正確な時刻を聞き取れる。	一部、時刻を聞き取りにくいところがある。	時刻を聞き取りにくい、あるいは聞き取れないところがある。
音声 (アクティビティ)	かなり正確で、わかりやすい発音である。ポーズの位置なども適切で、 どんな活動をするのか聞いてすぐわかる。	わかりやすい発音で、またポーズの位置なども適切である。どんな活動をするのか理解するのに問題はない。	多少、わかりにくい音声があっても、事前にスクリプトを見ていれば、どのような活動をするのかすぐわかる。	わかりにくい音声がところどころにあり、事前にスクリプトを見ていても、どの活動をするのか理解できないところがある。

(4つの評価の観点のうちの2つです)

言語形式の評価

- 「正確さ」を評価するのは、「正確さ」に関するフィードバックを得られるプロセスがあるときだけ
- 基本的には、言語形式が正確かどうかではなく、正確さがコミュニケーションにどんな影響をもたらしているかを評価する（客観的に観察可能な事象のみを評価するというルーブリックの原則）
- 「正しい／正しくない」は客観的に観察可能な事象ではなく、**教師の主観**
- **形式の正確さや適格性の判断は、どうしても教師に依存しやすい**→評価のブラックボックス化

不十分な点

- 理解可能度 *comprehensibility* を判断しているのは教師
- 入門段階の学習者は、他の学習者の（クリエイティブな）パフォーマンスを理解できるほど、リソースをもっていない
- ルーブリックで理解可能度を強調してもオートノミーにつながらない

	目標以上(4pt)	目標達成(3pt)	あと一步(2pt)	さらなる努力を(1pt)
アクティビティの内容	その人の休日のイメージがよくわかり、なおかつ、留学先の街の雰囲気や伝わる日課である。	その人の休日のイメージがよくわかる日課である。	その日をどんなふうに過ごすかということは伝わる。	日課としての情報が不足しているところがある。
振り返り (×2)	うまくできたところ、そうでもないところが具体的に把握できており、今後、どうしたらよいかという展望も具体的である。	うまくできたところ、そうでもないところを具体的に把握しようとしており、今後、どうしたらよいかという展望も具体的に考えようとしている。	うまくできたところ、そうでもないところを把握しようとしているが、具体性に欠ける。	うまくできたところ、そうでもないところを把握しようとしてしていない。

(4つの評価の観点のうちの残りの2つです)

学生の振り返り(1)

意味の区切りまで
ひといきに発音で
きなかった

変なところに
ポーズを置いて
しまった

この時のテーマは「意味の切れ目
にきちんとポーズを置く」

文を自然につ
なげることが
できなかった

このあたりは、やはり認識しやす
いのかも

もう少しハキハキ
読めばわかりやす
くなると思った

学生の振り返り(2)

第2声をすつ
と上げられ
なかった

第3声を低
くできな
かった

超分節音素（音の高さ）に
関しては意識できている

息を出す／出
さないを意識
できていない

子音／母音を意識で
きている者もいる

モデル音声と自分の音
声を比べる活動の経験
のある学習者たちです
が、この課題ではモデ
ル音声はありません

学生の振り返り(3)

アプリになかなか認識してもらえなかった

一発で認識してもらえるように練習したい

スマホのアプリに音声認識してもらえるかどうか、1つの基準になっている

学生の振り返り(4)

「慣れている」こと
ばと「慣れていな
い」ことばは違う

使ってみることで
××を覚えられた

ふだんの学習方法へのフィードバック

具体的な振り返りが無い学習者

- 「できない」のか「重要だと思っていない」のか……
- ポイントを×2にして加重化したことは、おそらく動機づけにはなっていない（こちらが何を重視しているかは伝わったと思いますが）
- 期限を過ぎて提出した学生は（申し訳ないと思っただらしく）かなり詳細な振り返りをしている

学習者が自己評価／相互評価に
使える理想的なルーブリックの例



ふつうなら、お弁当のおかずについて、「子供が好き嫌いをしないように、工夫してください」とか、「栄養が片寄らないようにお願いします」とか、いうところだけど、校長先生はひとこと、「**海のもの**と、**山のもの**を持たせてください」と、子供たちの家の人に、頼んだ、というわけだった。

山は……例えば、お野菜とか、お肉とか（お肉は山でとれるってわけじゃないけど、大きく分けると、牛とか豚とかニワトリとかは、陸に住んでいるのだから、山のほうに入るって考え）、海は、お魚とか、佃煮とか。この二種類を、必ずお弁当のおかずに入れてほしい、というのだった。

お弁当の時間に、校長先生が、自分たちのお弁当箱の中をのぞいて、

「**海のもの**と、**山**のものは、あるかい？」と、ひとりずつ確かめてくださるのが、うれしかったし、それから、自分たちも、**どれが海で、どれが山か**を発見するのも、ものすごいスリルだった。

タスク(2) ポーズの指示

1. 4人 (a, b, c, d) が1組になってください。aとbが並んで立ちます。
2. cは、aとbに中国語で指示を出してポーズを取らせます。
3. dは、その様子をcの肩越しに動画で撮影します。
4. 自分が指示を出している動画を、自分でチェックします (👉 「評価基準」)。
5. チェックした結果を提出します。

	目標以上 (4pt)	目標達成 (3pt)	あと一歩 (2pt)	さらなる努力を (1pt)
(1) 学習効果	指示を出す側も聞く側も楽しくかつ有益な学習ができるよう、個性的かつ多彩なポーズを取らせていた。	指示を出す側も聞く側も有益な学習ができるよう、多彩なポーズを取らせていた。	できるだけ多彩なポーズを取らせようという努力はしていた。	多彩なポーズを取らせようという努力が見えなかった。

「みんなのお勉強のために、できるだけ予測不可能な多彩なポーズを取らせる指示を出してね」という課題

(2) 確認のやりとり

目標以上 (4pt)

動作ごとにきちんと確認をし、相手が安心してポーズを取ることができたことがよくわかった。

目標達成 (3pt)

相手が安心していただけどうかはわからないが、動作ごとに確認のやりとりはしていた。

あと一步 (2pt)

ときどき確認を忘れて、相手がとまどったり、不安に思ったりしている様子が少しあった。

さらなる努力を (1pt)

確認のやりとりが少なく、相手が終始不安を感じていることがはっきりわかった。

インタラクションの重要性をわかってもらいたかった……

	目標以上 (4pt)	目標達成 (3pt)	あと一歩 (2pt)	さらなる努力を(1pt)
(3) 成功したかどうか	確認のやりとりも順調で、相手がスムーズの目標のポーズを取ることができた。	誤解やまちがいはあったが、やりとりの中で修正でき、最終的には、目標のポーズを取ることができた。	誤解やまちがいをことばでは修正できず、大幅に身振りに頼ったり、日本語に頼ったりして、目標のポーズを取ることができた。	確認のやりとりの中では修正できず、相手が目標のポーズを取ることができなかった。
(4) 音声	指示がすぐわかる、ストレスを感じない、はっきりした発音である。	指示がわかりやすい、はっきりした発音である。	ところどころ指示が聞き取りにくい発音であるが、意思の伝達はできる。	聞き取りにくいところが多くあり、意思の伝達に支障がある発音である。

「理解可能度」と「音声のクオリティ」を結びつけて考えてほしい

目標以上 (4pt)

目標達成 (3pt)

あと一歩 (2pt)

さらなる努力を(1pt)

該当項目なし

1～2項目該当

3～4項目該当

5～6項目該当

**(5) 文の
組み立て**

- [把 0] 放在 [ばしょ] という語順が狂っていた。 “把”を言い忘れた。
- “放在”を言い忘れた。 場所化の接辞を言い忘れた。 誰への指示かを言い忘れた。
- 「誰の～」かを言い忘れた。

文構造の適格性のチェックポイントをリスト化

学習者が自信をもてるところは？

学習者に「**得点を2倍**にしてほしい項目を2つ選んでください」という指示をしました。どれを選んだと思いますか？

1. （学習効果を高めるため）多彩なポーズ
2. 確認のためのインタラクション
3. 音声のクオリティ→理解可能度
4. 文構造の適格性

この自己評価をどう利用したか

- 当初は、同じルーズブリックを使って、**学習者の自己評価と、教員の評価を併せた評価を成績に算入しよう**と**思っていた**。が……
- 学習者の自己評価をほとんどそのまま採用した。なぜか……？
- 「そのまま」でない部分について話すと長くなるので、別の機会に……

今後やってみたいこと

- 「パフォーマンスのできばえ」を評価しないパフォーマンス評価
- 「うまくできたかどうか」を評価しないということです

15の話題領域、4段階のレベルに分けられた〈コミュニケーション能力指標〉

例えば【↓】のような指標がありますが……

- 日本と相手の国それぞれの年中行事で何を食べるかについて、写真などを見せながら口頭で紹介しあうことができる。（「食」Level 2）



『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』

学習者は、コンテンツベースの学習のどの過程を
自分でおこなう必要があるのか（あったのか）？

コンテンツを生成する過程

言語表現を生成する過程

何もなかったころ



機械翻訳の精度が
向上すると…



生成系AIが実用化
されると…

aidを使ったコミュニケーション

- 現実には、aid（補助手段）を使ったコミュニケーションが多用される
- aidの使用を禁じてコミュニケーションのスキルを測るのは、裸眼視力を測るようなもの
- 最終的な完成度には、いろいろな補助手段（機械翻訳、AI、熟達した使用者）が寄与している。もしかしたら、本人がほとんど寄与していないかもしれない……それを評価してどうする？

オートノミーのための発想の転換

- 自分自身のパフォーマンスのできばえをよくする努力をさせることより、他者のパフォーマンスをきちんと評価できる能力を身につけさせることのほうが重要ではないのか？
- 我々はそもそも、1回限りのパフォーマンスで見せた学習者のスキルの保持と転移を検証したことがあったか？（とくに、各種のスピーチコンテストのようなパフォーマンスの場合……）

「できばえ」を競う 評価方法をやめる

- × 学習者Aのパフォーマンスを教師が評価する（**できばえ**のブラックボックス的評価）
- × 学習者Bが学習者Aのパフォーマンスを評価した結果が、Aの評価となる（**できばえ**の相互評価）
- 学習者Aが学習者Bのパフォーマンスをどのように評価したかが、Aの評価となる

ねらい

1. 「よいパフォーマンスとは何か」という**メタ認知**
 2. パフォーマンスを評価する過程での**気づき**
- ……**学習者オートノミー**につながるはず