

プレースメントのための日本語スピーキングテストSTARの開発

ボイクマン 総子・根本 愛子・松下 達彦(東京大学)

1. はじめに

■プレースメントテスト(PT)の要件

- ・一度に大勢の受験が可能
- ・実施が簡便
- ・判定がすぐに出せる
- ・(判定者間・判定者内で)一貫性の高い結果が出せる



スピーキングテストの開発
“STAR: Speaking Test of Active Reaction”

2. 先行研究と研究課題

■Bachman & Palmar (1996)

- ・テストには「有用性」が大事
有用性: 信頼性, 構成概念妥当性(妥当性), 真正性, 相互性, インパクト, 実用性

■日本語のスピーキングテスト

- ・ACTFL-OPI(牧野他 2001), JF日本語教育スタンダード準拠ロールプレイテスト(国際交流基金 2017)等

課題

- ①対面式テストなので一度に大勢が受験できず時間を要す
 - ②テスト実施者のトレーニングが必要である
- ⇒①②から、実用性の面でPT向きとは言えない
②から信頼性の面で課題が残る

■テスト(STAR)の開発

- ・出題タスクを予め録音しておき, 受験者はそれを聞いて話す⇒実施が簡便な反応型のテストの開発
- ・判定用ルーブリックと音声サンプルも開発



本研究の目的:

筆者ら2名(A, B)の他, 日本語教師2名(C, D)と学部生1名(E)に判定を依頼し, STARがPTとして**有用であるか検証**

研究課題...1)信頼性 2)妥当性 3)実用性 の検証

4. 結果

■5名の判定者間の判定

級内相関係数 $ICC(2,1)=.862$, $ICC(3,1)=.885$
クロンバック $\alpha=.975$, Fleissのカッパ係数 $k=.385$

表3 STARにおける判定者間のピアソンの相関係数(r) n=32

	判定者A	判定者B	判定者C	判定者D	判定者E
判定者A	1	.921**	.901**	.895**	.868**
判定者B	.921**	1	.877**	.889**	.906**
判定者C	.901**	.877**	1	.900**	.918**
判定者D	.895**	.889**	.900**	1	.844**
判定者E	.868**	.906**	.918**	.844**	1

**p<.01

■他のPTとの関係

表4 判定結果とその他のテストの相関 n=32

	総合	SPOT	漢字	語彙	文法	読解	作文
判定者A	.710**	.804**	.402**	.480**	.815**	.503**	.753**
判定者B	.664**	.724**	.429**	.510**	.710**	.503**	.669**
判定者C	.634**	.769**	.347	.411*	.781**	.373*	.718**
判定者D	.726**	.766**	.479**	.542**	.793**	.546**	.681**
判定者E	.584**	.698**	.374*	.451**	.681**	.406*	.638**

*p<.05, **p<.01, ***p<.001

- ⇒漢字・語彙などの書き言葉との相関は相対的に低い
- ・文法とSPOTとの相関が相対的に高い,
- ・受容(読解)よりも産出(作文)の相関が高い

3. 方法

■受験の手順と実験の手順

- ・受験者への指示:「断り」の状況を示し、「この状況で自分ならどう言うか相手の発話を聞き終わったら、各自録音を始めて相手に返事をして下さい。必要だと考え得る十分な発話を1分以内で行って下さい」
- ・相手の発話:「ねえ下北にあるクラブでさっき話してたバンドがこれから演奏するんだけど、授業行くのやめて一緒に聞きに行かない?」

■判定基準 表1 働きかけの発話のSTARにおける判定基準

	レベル1	レベル2	レベル3	レベル4	レベル5	レベル6
課題達成	できるが直接的		簡単に理由や状況を述べたうえで直接的な表現で課題を達成できる	理由や状況を述べたうえで直接的な表現で課題を達成できる	理由や状況を述べたうえで間接的な表現も交えて課題が達成できる	
対人配慮	配慮ができない, または, 「すみません」や「ちょっと」を使った配慮ができる		課題達成のための最低限の配慮ができる	課題達成のためにそれなりの配慮ができる	課題達成のための十分な配慮ができる	
テキストの型	単語レベルだが, いくつか単文が出てくる	文レベルだが, 単文が多い	複文を使って話ができる。一部で段落が見られる	ほぼ段落レベルで話すことができる	段落レベルで話すことができる	
文法・表現の正確さ	基本は単語, 文は定型的なもののみ	定型的な文, 表現のみ	間違いがあり, 聞き手の理解を妨げることがある	間違いがあるものの, 聞き手の理解を妨げることがほとんどない	間違いはあるが, あまり目立たない	ネイティブレベルでほとんど間違いはない
文法・表現の豊かさ	決まったもののみ	日常的に必要な最低限の表現が使える	日常的な表現が十分にできる	さまざまな表現を使おうとする努力がみられる	さまざまな表現を問題なく使うことができる	慣用句や比喩なども交えて話すことができる
流暢さと言い直し	定型的な文が長い休止を伴ったり, 言い直したりしながら言える	言い直しが多い	流暢さがみられるが発音の悪さでわからない部分がある。言い直しも多い	流暢だが, ときどき言い直しがみられる。発音は気になるが, わからないことはない	言い直しを多少交えながら, 流暢に話すことができる。少々発音が気になる	流暢で, 言い直しがあっても気にならず, 発音もよい

■判定者情報

表2 判定者と判定の所要時間 *時間制限を設けなかった

	教師歴	判定の前作業	受験者32人分の判定所要時間
判定者A(筆者)	24年	N/A	20分
判定者B(筆者)	19年	N/A	20分
判定者C(日本語教師)	22年	30分	60分
判定者D(日本語教師)	16年	30分	120分
判定者E(学部生)	0年	15分	30分

5. 考察と今後の課題

1)信頼性

- ・判定者間の相関が高い ⇒**信頼性が高い**
- ・ただし, 中級の差異を判定するのは難しい

2)妥当性

「断り」という1タスクを測ったものであるが, コミュニケーション能力(Celce-Murcia 2007)の5つを満たし, 即応力も課している ⇒**話し言葉の特徴を十分備えている**

3)実用性

短時間で6レベルを弁別でき, 一度に大勢が受験可能 ⇒**実用性がある**

■STARの有用性

妥当なレベル認定のための情報が得られるだけでなく, 書き言葉は苦手でも話し言葉は得意な受験者に, 適当な科目の受講を許可・提案でき, PTとして有用である。

■今後の課題

断りのタスクだけでなく, STARの他のタスク:断り以外の働きかけのタスク(依頼や勧誘)や, タイプの異なるタスク(絵を見て描写, 話を聞いて再話)の有用性も検証する

参考文献

- 国際交流基金(2017)『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(第二版第一刷改訂版)』独立行政法人国際交流基金
- 牧野成一他(2001)『ACTFL-OPI入門』アルク
- Bachman, L. and Palmar, A. S. (1996) *Language testing in practice*. Oxford University Press.
- Celce-Murcia, M. (2007) Rethinking the role of communicative competence in language teaching. In E. Alcón Soler & M. P. Safont Jordà(Eds.), *Intercultural language use and language learning*. pp.41-57. Springer Netherlands.
- Council of Europe (2018) *CEFR Companion Volume with New Descriptors*. <https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989> (2019年2月9日閲覧)