第52回日本言語テスト学会研究例会20210214
（オンライン講演3：13:15～14:30）

教室内技能統合型スピーキングテストにおけるルーブリックと採点：信頼性を高めるために

平井明代
（筑波大学）

1

1

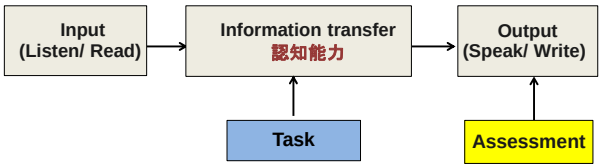
技能統合型評価
（Integrated Skills Assessment）

- １つのテストに複数のスキルを統合させて、実生活で経験するような複数のタスクが要求される状況に対処できるかを見極めるための評価。

2

2

技能統合型評価



ねらい
①受容能力の向上
②文法事項を使えるようになる
③産出能力の向上
④思考力、判断力、表現力を育む

3

3

外部テスト：VERSANT English Speaking Test

妥当性：
リスニング能力とスピーキング能力（自然さ、流暢さ、即時性）を測定

タスク：
A.reading(音読8問)
B.Repeating（復唱16問）
C.Question（質問24問）
D.Sentence built（文の構造10問）
E.Story Retellings（ストーリーリテリング3問）　(L→S)
F.Open questions（自由回答2問）

採点：全体的（わかりやすさ）、4観点（文章構文、語彙、流暢さ、発音）

信頼性：
言語認識システムと自動採点システムを使って採点。客観性・一貫性

(Pearson Education, Inc, 2013).

4

4

外部テスト： TOEIC Speaking Test

妥当性：　スピーキング技能（熟達度）
理解しやすい言葉で話すことができる。/　日常生活や仕事上必要なやりとりをするために適切に言葉を選択し、使うことができる。/一般的な職場において、筋道の通った継続的なやりとりができる。

タスク：6題中2題

4. Respond to questions using information provided (R, L → S)	提示された資料や文書(スケジュール等)に基づいて、設問に答える	発音・イントネーション、アクセント・文法・語彙一貫性・内容の妥当性・内容の完成度	0～3
5. Propose a solution L → S	メッセージまたは会議内容を聞き、その内容を確認した上で、問題の解決策を提案する	同上	0～5

信頼性: 複数者採点 / 評価者訓練

（国際ビジネスコミュニケーション協会, n.d.）

5

5

外部テスト： TOEFL iBT Speaking Test

妥当性
教室内や教室外におけるアカデミックな場面おける、スピーキング技能（熟達度）を測定

タスク
- 【Independent task】 1問
  身近なトピックについて意見を述べる
  準備15秒　解答45秒
- 【Integrated tasks】 3問
  読んだり聞いたりした内容を要約して話す
  （1）Read＋Listen→Speak 2問
  　準備30秒　解答60秒
  （2）Listen→Speak 1問
  　準備20秒　解答60秒

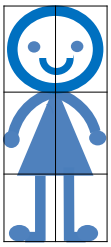
採点：Delivery, Language use, Topic, development, 各0-4点)

信頼性: 複数者採点 / 評価者訓練　（ETS, 2021）

6

6

教室内技能統合型テスト



妥当性

スピーキング技能（到達度）＋α

形成的評価（Formative assessment）

総括的評価（Summative assessment）

タスク

- リテリング (Versant)
- 要約 (TOEFL iBT , TOEIC)
- 質疑応答（ TOEFL iBT, TOEIC ; 英検2次試験）
- ペアによるディスカッション
- ロールプレイ
- プレゼンテーション
- ディベート

→ 年間指導計画に組み込む

→ 「知識・技能」「思考・判断・表現」「主体的に学習に取り組む態度」

7

7

技能統合型テストの利点

1．真正性が高い（現実の一連の行為が出来るかを推測）(Brown, 2007)
2．トピックによる影響が小さい
3．テストが作成しやすい
4．使用テキストとアウトプットタスクでさまざまなレベルに対応できる
5．受容能力と産出能力だけでなく，より思考が活性化される
(Brown, Iwashita, & McNamara 2005)
6．学習項目の挿入
7．フィードバックがしやすい（テキストの参照）

真正性，実用性，波及効果
→ 教室内で行うテストに向いている

8

8

技能統合型テストの短所

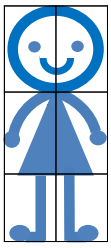
1．入力情報やインプット能力に左右される (e.g., Frost, Elder, & Wigglesworth, 2012)
2．暗記力の影響の可能性

→他の独立型スピーキングタスクと合わせる

9

9

信頼性を上げるために
（実用性を考慮して）



1. 採点者を増やす
2. タスク回数を増やす
3. 評価者訓練をする
4. ピア評価者を匿名にする
5. 生徒と教師間で採点する観点を分ける
6. ルーブリックを工夫する
7. テクノロジーを利用する

10

10

1. 採点者数を増やす

教師評価（1人，または2人）

・学年単位で協力する，AET

・2人で分けて採点する際の注意点

厳しさの異なる評価者の場合

→評価者内信頼性や評価者間信頼性だけでなく、採点得点の一致率(Exact agreement) も大切

1. 一部採点して，相談して調整する
2. できるだけタスク別採点をする。

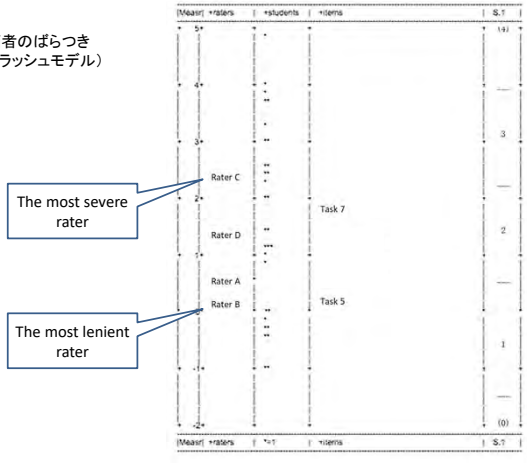
生徒評価 （＝ピア評価）

複数の生徒評価点の平均点にする（e.g., 深澤, 2009）

11

11

評価者のばらつき
（多層ラッシュモデル）



12

12

2. タスク回数を増やす

・スピーキングテスト：SRST（Story Retelling Speaking Test)を使って調査
(Hirai & Koizumi, 2008)

Q1. いくつのタスク（ストーリー）を使うとよいか?

・参加者：52名（大学生、英語初中級レベル）

・材料：４つのタスク（ストーリー）
EBB (Empirically derived, Binary-choice, Boundary-definition) 尺度
4観点：伝達力（Communicative Efficiency：CE), 文法・語彙 (G&V),
内容 (Content), 発音 (Pro)
Analytic尺度
4観点：伝達力（CE), 文法（Grammar）、
流暢さ（Fluency), 発音（Pro)

・採点者：6人の採点者が2種類の尺度を使用する

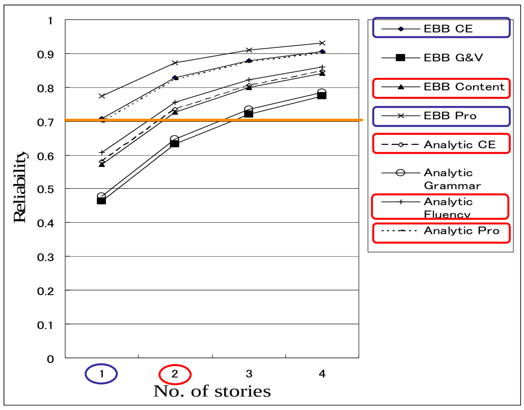
・分析方法：mGENOVA　（多変量一般化可能性理論）

13

13

2. タスク回数を増やす (EBB1を使って)

Changes in phi coefficients in the decision studies (p x s Design).



14

14

3．評価者訓練をする

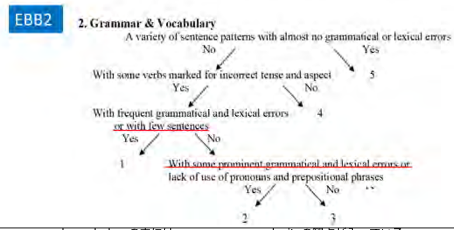
方法
ルーブリックと実際のパフォーマンスサンプルを使った明示的な訓練

効果
1．ピア評価と教員評価の相関が高くなる
（Okuda & Otsu, 2010; Patri, 2002)
2．コメントによる評価力アップ（Saito, 2008）。
3．評価者内信頼性（Saito, 2008, Taylor & Galaczi, 2011;
McNamara, 1996）。
4．短時間で複数回トレーニングによる効果（笠巻, 2021）

効果への疑問
評価者の性格や考え方（Hughes, 1989；Sundqvist, et al. 2020）。

15

15



■ Grammar and vocabularyの中には、accuracy, complexityの観点が入っている。
■ With some verbs marked for incorrect tense and aspect：時制やアスペクト (進行形、完了形など) の間違いはあっても、あまり多くない。1, 2個以内。その間違いがほとんど気にならない。
■ Use of pronouns and prepositional phrases
e.g., 1. They (←Bob and Jean) didn't want to leave the beach. (但し、原文でpronounを使っていないのであれば厳しくつけない)
2. I want to go with him [prepositional phrase].
■ errorの判定は、dysfluency markersを除いた形で考える。
例：She {like} likes English. Correct
■ lexical error: soonerlyなど、英語として存在しない語
■ Tense and aspect: major errorと考えた
Pronouns, prepositional phrases: minor errorと考えた　（また、習得上、正しい時制とアスペクトを使用できるほうが難しいかどうかを確認する）
■Grammarで、発話が少ないことで1になる理由は、発話が少ないということは
learners don't have sufficient grammatical knowledge even to construct short sentencesだから。
■ With few sentences: 文の数については、opinionを含めて数える。and等で続けている場合は、分けて2文にカウントする。4、5文以下の発話はfew sentencesと考える。
■ some prominent grammatical and lexical errorsは、意味が伝わりにくい目立った誤りが2個程度ある。
■ frequent grammatical and lexical errorsは、発話量に比べて、比較的誤りが多い場合。

16

16

4．ピア評価者を匿名にする

Fruhodwlrqv#Ehwz hhq#Shhu#dqg#Whdfkhu#Dvvhvvp hqw#lq#wkh#Wz r#H{shulp hqw

Criteria	Experiment 1 (individual) (*n* = 50)	Exp 2 : Group A Discussion (paired) (*n* = 26)	Exp 2: Group B Anonymous (individual) (*n* = 33)
Communicative Efficiency	.28	.24	.57**
Grammar &Vocabulary	.19	-.10	.42*
Pronunciation	.21	.18	.33
Total	.17	.06	.63**

-s ? #381#--s# #341　　(Hirai, Ito, & O'ki, 2011)

17

17

4．ピア評価者を匿名にする

結果
a. ピアの匿名性で、教師評価との相関が高まる
b. ピア評価者の話し合いで、匿名性効果は失われる

→相手を気にしなくて採点できるため、匿名性は大切
(Hirai, Ito, & O'ki, 2011)

匿名性の確保
→テクノロジーの利用
Zoomで教員にのみ評価を送る
Google Formの利用
PeerEvalアプリの利用
→匿名性と集計機能，フィードバックが早い
(加野 & ゴーベル, 2019)

18

18

5．生徒と教師間で観点別採点をする

参加者
53 名の高校生 (CEFR-J A1.1 to A2.2 levels by TSST)

スピーキングテスト
1. SRST（2つのストーリー）
(a) 流暢さ
(b) ターゲット;ターゲット以外の G & V（言語的正確性）
(c) 内容 (内容の正確性)
(d) 発音
(e) 意見交換（即興性と自分の考えを発信する能力

2．分析的尺度A （高校生によるピア評価と教師評価）
（Hirai & Yokouchi, 2019）

3．分析的尺度B （大学生と教師評価） （Hirai, 2021）

19

19

5．生徒と教師間で観点を分ける

教師評価とピア評価（高校生）の相関とt-test (n =53)

	Max	Teacher		Peer					
		M1	SD	M2	SD	r	M2-M1	t(52)	d
Fluency	6	2.68	1.22	3.28	1.29	.25	0.60	-2.86**	0.48
Target	4	1.59	1.05	3.19	1.04	.13	1.60	-8.48**	1.54
G&V	6	2.23	0.78	2.60	0.88	.19	0.37	-2.59*	0.46
Content	6	3.96	1.48	4.60	1.04	.35*	0.64	-3.15**	0.50
Pronunciation	6	3.32	1.17	3.72	1.46	.36**	0.40	-1.91	0.30
Opinion	6	2.77	1.66	3.55	2.07	.72**	0.78	-3.89**	0.41
Total	34	16.55	5.01	20.94	5.26	.52**	4.39	-6.37**	0.86

*p < .05, **p < .01. （Hirai & Yokouchi, 2019）

20

20

6．ルーブリックを工夫する

1．評価得点を短時間で出したい
全体的（総合的）尺度（holistic scale)でかつ段階が少ない（3段階）
e.g., 全体的評価（4段階） = 観点別評価 (6観点・各4段階)
(Yokoyama, 2016)
マイナス面（診断的(ー), 向上の可視化（ー）
観点によって段階が異なる）
→ ルーブリックの記述を工夫することである程度診断的にできる （例参照）

2．診断的評価にしたい
分析的尺度（analytic scale), 評価
（観点・段階（レベル）が多いほど診断的）
（実用性（ー），信頼性（ー）→評価者訓練）
→ 1回で採点する観点を絞る（3−4観点）
→ 段階を少なくしてリスト形式する（3段階）

21

21

SRST総合的評価例

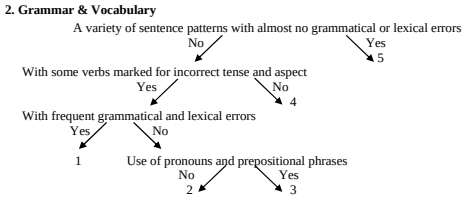
得点	観点	コメント
4	• 発音、文法ともに誤りがほとんどなく内容理解が容易である。 • 言いよどみがなく流暢に話せる。聞きやすい。 • 物語の説明を大枠、詳細ともに述べ、自分の意見を3文以上述べられる。	
3	• 発音の誤りが少ない。 • 文法の間違いはあるが、気にならない。 • 言いよどみや言い直しがあるが、自然なリズムで話せている。 • 物語の詳細の内容にも触れられ、自分の意見も3文以上あるが、4に比べて内容が薄い。	lost lunch "I"
2	• 発音の間違いが多いが、伝わる。 • 文法の間違いが、少し気になる。 • 単語ごとに途切れたり、黙ってしまう箇所があり聞き取りづらい。 • 物語の大枠だけが語られ、自分の意見も1、2文と少ない。	goes 時制
1	• 発音の間違いが多すぎて、伝わらない。 • 文法の間違いが、かなり気になる • 沈黙やいい直しが多く、ほとんど聞き取れない。 • 物語が誤った内容で語られ、自分の意見も述べられていない。	

Yokoyama (2016) 一部改変

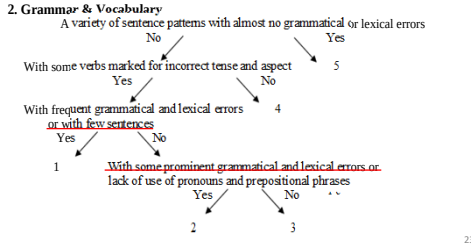
22

22

EBB1



EBB2



23

23

Rating Scale Statistics for Grammar & Vocabulary

	EBB1	EBB2	MT
(a) Person discrimination Person separation ratio:	1.85	2.71	2.01
(b) Rater separation Rater separation ratio:	0.00	0.82	1.91
(c) Rater reliability Rater point biserial: Exact agreement ratio:	 .30 57.5%	 .47 51.2%	 .23 41.1%
(d) Variation in ratings Rater misfit (%): Rater overfit (%):	 0/6 (0.0) 1/6 (16.7)	 0/9 (0.0) 1/9 (11.1)	 2/9 (22.2) 0/9 (0.0)
(e) Scale properties Level 1 Level 2 Level 3 Level 4 Level 5	Count (%) [logit] 31 (11%) [-0.73] 8 (3%) [-0.35] 98 (36%) [-0.24] 112 (41%) [0.32] 23 (8%) [1.75]	Count (%) [logit] 41 (15%) [-2.22] 72 (26%) [-1.33] 92 (34%) [-0.42] 48 (18%) [0.88] 19 (7%) [2.76]	Count (%) [logit] 33 (13%) [-1.96] 55 (22%) [-1.47] 102 (41%) [-0.88] 53 (21%) [0.29] 5 (2%) [2.66]

（Hirai & Koizumi, 2013）

24

24

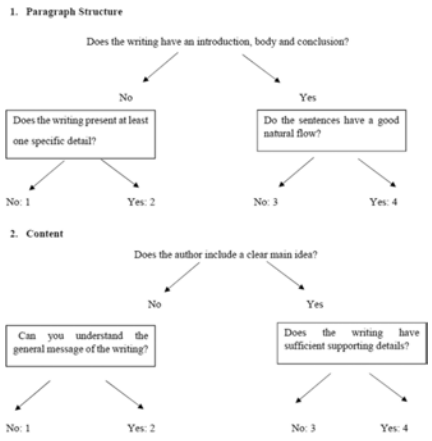
Phi Coefficient (Φ) in Decision Studies of the EBB1, EBB2, and MT Scales for Each Criterion (p x s Design)

EBB1	1 story	2 stories	3 stories	4 stories
Communicative Effi	.741	.851	.896	.920
Content	.642	.782	.843	.878
G & V	.460	.630	.718	.773
Pronunciation	.769	.869	.909	.930
EBB2	**1 story**	**2 stories**	**3 stories**	**4 stories**
Communicative Effi	.621	.766	.831	.867
GG & V	.725	.840	.888	.913
Pronunciation	.821	.902	.932	.948
MT	**1 story**	**2 stories**	**3 stories**	**4 stories**
Communicative Efficiency	.603	.752	.820	.859
G & V	.502	.669	.752	.801
Pronunciation	.608	.757	.823	.861

25

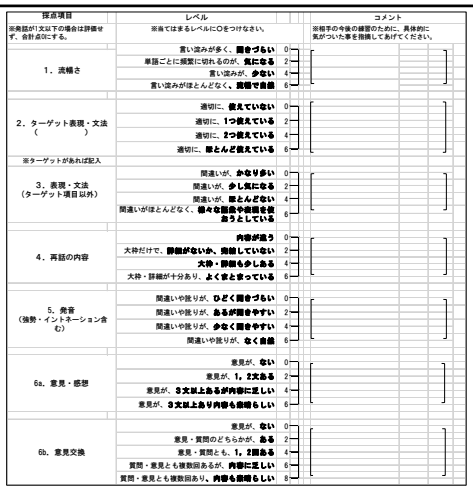
25

Jeong, H. (2019)



26

SRST模型評価シート（2019版）



27

SRSTターゲットありバージョン　　Story 2. Hawaii

Read the story silently within two minutes. Pay attention to underlined grammatical items in order to use them when retelling.（**2分間で次の文章を黙読しなさい。特に、下線部分の表現や文法事項（受身形）をリテリングで使えるように注意して読みなさい。**）

Visiting Hawaii

Hiro's family decided to visit Hawaii for their summer vacation. **Hiro was very excited** because he had always wanted to go abroad. He began to study harder in his English class at school. He also bought a phrase book and learned lots of useful English words and phrases.

When Hiro's family got to Hawaii, **Hiro was surprised** to find that many people spoke to him in Japanese. **He was disappointed** that he could not practice his English. Then one day, his family went to a restaurant in a small town. Nobody in the restaurant knew Japanese. So, Hiro's family had to use English to order their food. After they had ordered, his mother said, "Your English is much better than mine." He was very happy to hear that.

After the signal, read each question aloud and answer it in English.
それぞれの質問の前の合図を待って、1問ずつ質問を読み上げて、英語で答えなさい。
Q1 ～ Q4

Hawaii_K　　Hawaii_U

28

28

Hawaii_Student K_A_R

•Hiro … was going to … go to Hawaii with his family. Hiro was excited … because he had wanted to go abroad. … Hiro was disappointed because many people in Hawaii … spoken to … spoke to Hiro in Japanese. Hiro went to the restaurant with his family. Nobody in the restaurant knew Japanese. So Hiro spoke to … spoke in English. Before Hiro went to Hawaii, Hiro practiced English very hard. So Hiro could speak very well. … After ordered, Hiro's mother said to him your English is much better than mine. … Hiro was very happy to heard … to hear that.

… I think … studying hard is very important … because … I can it someday. …Hiro couldn't use his English at first. But He can…

29

29

SRST採点シート（2021版）

	観点	4	3	2	1	0 採点不可	コメント
1	流暢さ	言い淀みがほとんどなく，流暢で自然	言い淀みが，少ない	頻繁に切れるのが，気になる	言い淀みが多く，ポーズが長すぎる	理解できない	意味の塊まで切らないように，音読やリテリング練習
2a	ターゲット表現・文法	※4つ以上，正しく使えている	3つ以上，正しく使えている	2つ，正しく使えている	1つ，正しく使えている	全て，使えていない	
2b	表現・文法（言い換え表現+1か+2）	間違いが，ほとんどない	間違いはあるが，気にならない	間違いがいくつかあり，気になる	間違いが，多い	理解できない	代名詞（He）も使うとより自然
3	内容	ほぼ全て。内容も正確でまとまっている	8割ほど。ややまとまりにやや欠ける	半分ほど。内容の正確性に欠ける	3割ほど。内容の正確性に欠ける	内容が伝えられていない	
4	発音	間違いや訛りがなく，イントネーションが自然	間違いや訛りが少なく，気にならない	間違いや訛りが，気になる	間違いや訛りが多く，聞き取りにくい箇所がある	聞き取れない	
5a	意見・感想	十分あり，まとまりと深まりがある	十分あるが，まとまりと深まりに欠ける	2，3文あるが，内容に乏しく散文的である	意見・感想が，ほぼない	意見・感想がない	
5b	意見交換	十分なやり取りで，内容の深まりや進展がみられる	十分なやり取りはあるが，内容の深まりに欠ける	やり取りは少しあるが，内容に乏しい	意見交換が，ほぼできていない	意見交換がない	
6	態度			声の大きさ，アイコンタクトが適度で，伝えようとしている	声が小さく，アイコンタクトも少なく，伝わりにくい	聞こえない	

※ターゲット4つ以上設定する場合。2つの場合は2点まで
※迷った場合は中間点（例2.5）でもOK。後で切り上げにするか判断
※観点は目標や指導に合わせて選ぶ。または、担当する観点を分ける

得点

（平井，2021版）

30

30

（SRST　意見交換バージョンb）

Smartphone Addiction

4. Look at keywords and summarize the story in English in 1 minute and 30 seconds. (今読んだ内容を下記のキーワードを見て、英語で **1 分半**で要約しなさい。)

Keywords: Miki, chat, mother, timetable, control

5．After the signal, make a pair and exchange opinionsin English about the following questions for 3 minutes.（合図の後、ペアになって次の質問について、英語で **3 分間**で、お互いの意見を交換しなさい。）

Q 1 . Do you think the treatment suggested in the text is effective? Why or why not?
（テキストにある対処法は効果があると思いますか。理由を挙げて判断しなさい。）
Q2. Raise other treatments with reasons for their effectiveness.
（本文にない対処法について、その効果の理由や具体例なども挙げなさい。）

6．5 に関して評価シートの全ての観点で満点が取れるように書いてまとめなさい。

31

31

（SRST　意見交換バージョンb）**Questions（口頭と筆記）**

Student0002

Oral

Q1

I think the treatment in this sentence is e-, very effective. Because Japanese teenagers, …is, are unconscious for how, how long they use smartphone. So, writing timetable is realized them how much……, 2

Q2

I hear that parents, なんか、control their children's using smartphone. 1

Written

Q1

I think that the treatment which Miki's doctor used is effective. Even though people are told that they use their smartphone too much and they are smartphone addiction, **they cannot realize that**. So, in my opinion, **writing down their daily timetable helps them find how long** a time **they spend to use their smartphone.** 3

Q2

I think that using an application which enables parents to restrict their children's time to use their smartphone is effective, too. 1

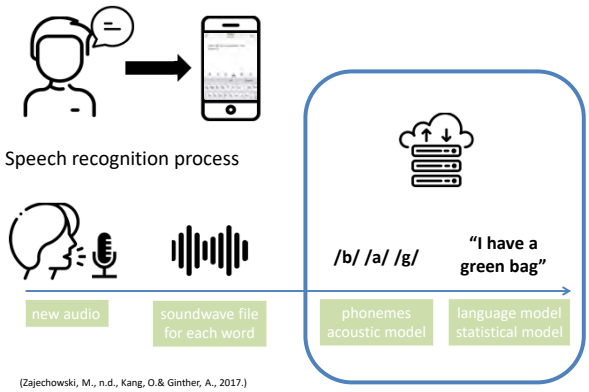
Some video games**, such as Nintendo Switch, also have similar function**. So, parents can watch whether their children play video games a lot and whether they are safe or not.

These applications or functions may allow Japanese teenagers' parents to prevent their children's smartphone addiction. 3

32

32

7. ITを使う:　発音



(Zajechowski, M., n.d., Kang, O.& Ginther, A., 2017.)

33

33

7. ITを使う:　発音

Five Speech-to-Text applications' average transcription accuracy under four tasks

			95%CI	
Tasks	Mean (%)	SE	Lower	Upper
1. Read sentences	59.04	3.12	52.66	65.42
2. Read a passage	62.39	2.97	56.33	68.46
3. Retell the passage	57.60	3.57	50.31	64.90
4. Answer easy questions	64.41	2.41	59.48	69.34

Note. Task 1(Read short): reading two sentences with loan words
Task 2(Read long): reading an easy passage
Task 3 (Retell): retelling the passage without looking at it
Task 4 (QA): answering 3 easy questions

(Kovalyova & Hirai, 2020)

34

34

7. ITを使う　発音

Table．Some consonants not transcribed accurately

Original	Wrong transcriptions (freq)	Issue
lunch	ranch (17)	/r/, /l/
play, explain	pray (10)	/r/, /l/
school	scooter (5)	/r/, /l/
reading	leading (5)	/r/, /l/
three	free (17), tree (10)	/th/
bag	back (100 or more), bank (15)	consonants at the end
shocked	shot (5), choked (5)	consonants at the end
Kenji	can she (6), Kensi (12), candy (17), Kenzie (100), change it(7)	Japanese, loan words

(Kovalyova & Hirai, 2020)

ITの使用：高い信頼性と速やかなフィードバック

35

35

まとめ

1．技能統合的スピーキング評価を行う利点は大きい。
（1）現実の英語使用場面で使えるかを評価することができる
（2）受容能力、学習した文法・語彙の定着、発信能力が鍛えられる
（3）認知能力（思考力、判断力、表現力など）を鍛える

2．学習者のスピーキング能力を正確に把握できるように、年間計画の中で異なるタスクを取り入れる。

3．教室内テストで信頼性を高める工夫をする。
1．採点者を増やす（複数の生徒評価の平均点の取り込みなど）
2．タスク回数を増やす（年間計画の中で）
3．評価者訓練をする（説明程度でも，複数回する）
4．ピア評価を匿名にする（評価を教員が受け取る，テクノロジーを使う）
5．生徒と教師、または生徒間で採点観点を分ける（信頼性の高い部分任せたり、一回の採点の観点を絞る）
6．ルーブリックを工夫する（使いながら改良すれば，信頼性が高いものができる）
7．テクノロジーを利用する（一貫したフィードバックを早く，具体的にできる）

36

36

主な参考文献

Brown, H. D. (2014). Principles of language learning and teaching (6th ed.). NY: Pearson Education.

Brown, N. Iwashita, & T. McNamara (2005). An examination of rater orientations and test-taker performance on English-for-Academic-Purposes speaking tasks. *TOEFL Monograph Report 29*. Princeton, NJ: ETS. https://www.ets.org/Media/Research/pdf/RR-05-05.pdf

Cumming, A. (2013). Assessing integrated skills. In A. Kunan (Ed.), *The companion to language assessment* Vol. 1 (pp. 216–229). New York, NY: John Wiley & Sons, Inc. doi:10.1002/9781118411360.wbcla131

Frost, K., Elder, C., & Wigglesworth, G. (2012). Investigating the validity of an integrated listening-speaking task: A discourse-based analysis of test takers' oral performances. *Language Testing, 29*, 345-369.

Jeong, H. (2019). Writing scale effects on raters: an exploratory study. Language Testing in Asia 9:20 https://doi.org/10.1186/s40468-019-0097-4

Hirai, A., Ito, N. & O'ki, T. (2011). Applicability of peer assessment for classroom oral performance. JLTA Journal 14.

Hirai, A. & Koizumi, R. (2013). Validation of Empirically-Derived Rating Scales for a Story Retelling Speaking Test. *Language Assessment Quarterly*, 10: 398?42 DOI: 10.1080/15434303.2013.824973

Hirai, A., & Yokouchi, Y. (2019). An Investigation of EFL Learners' Diagnostic Assessment Capabilities for a Classroom-based Speaking Test. Annual Review of English Language Education in Japan, pp. 209-224.

Kang, O.& Ginther, A. (Ed). (2017). Assessment in second language pronunciation. Routledge

Saito, H. (2008) EFL Classroom Peer Assessment Training Effects on Rating and Commenting. Language Testing, 25, 553-581.

笠巻知子 (2021) 学生による相互強化よくとプレゼンテーション力に及ぼす要因に関す実証的研究－学生の相互評価と教員による評価との相関と評価者トレーニングに基づいてー（京都外国語大学博士論文発表会資料）

加野まさみ&ゴーベル・ピーター(2019) プレゼンテーション授業における学習者相互評価モバイルアプリ使用とそれに対する学生の意識について.京都産業大学総合学術研究所所報14,47 - 61

平井明代 (2015)「授業を活かすストーリーリテリング・テストの活用」大塚フォーラム(33), pp. 49-69.

平井明代 (2016)「3.3.5 技能統合的スピーキングの評価」20周年記念特別号19巻2号, pp. 116－121. doi. https://doi.org/10.20622/jltajournal.19.2_0

深澤真(2009). スピーチにおける生徒相互評価の妥当性 ―項目応答理論を用いてー. *STEP Bulletin 21*, 31-47.

37

37

ご清聴ありがとうございました。

hirai.akiyo.ft@u.tsukuba.ac.jp

38

38