
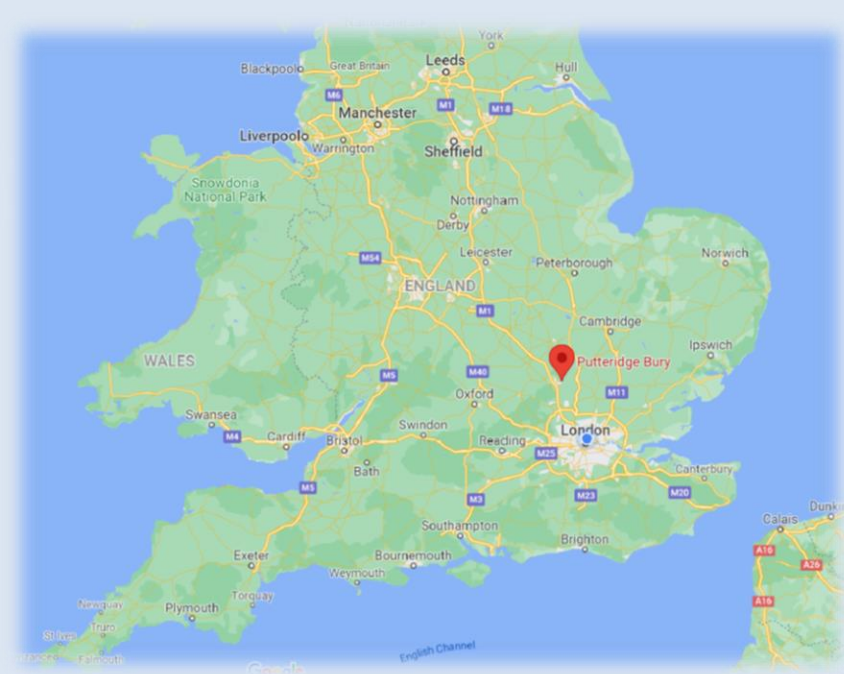


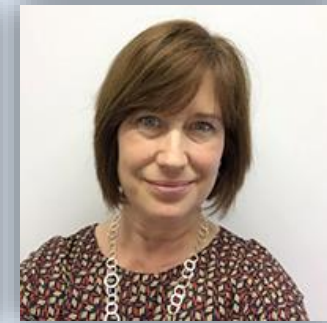
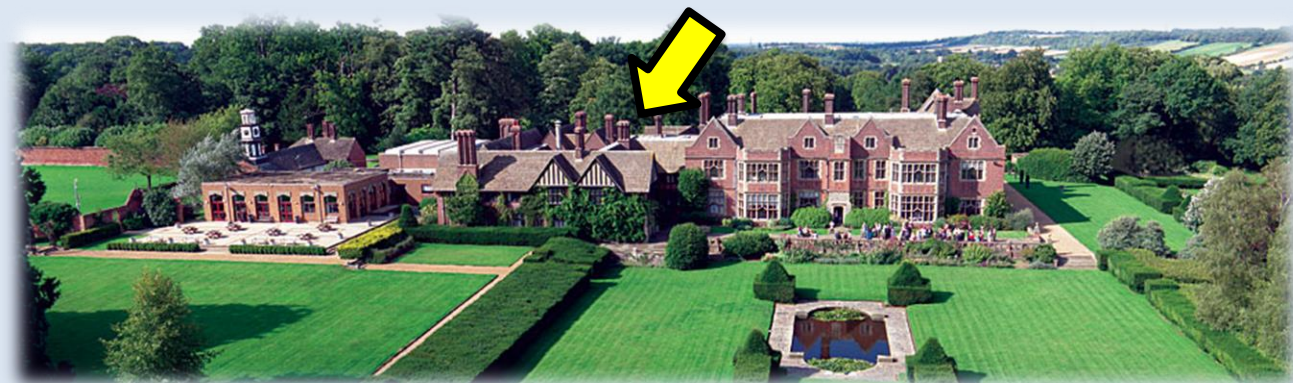
英語スピーキングテストって何？



英国ベッドフォードシャー大学
言語学習評価研究センター
中津原 文代



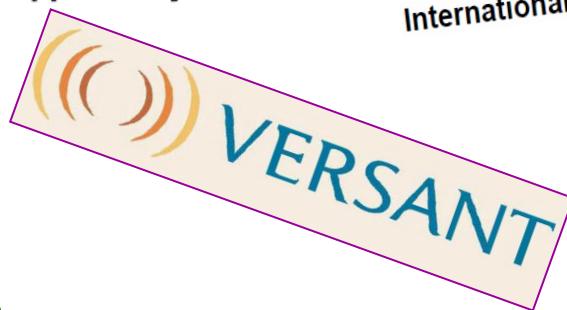
CRELLA Centre for **R**esearch in **E**nglish **L**anguage **L**earning and **A**ssessment University of Bedfordshire



www.beds.ac.uk/crella



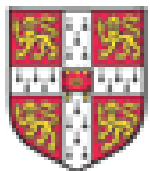
テスト開発・妥当性検証プロジェクト(一部)



Plan Ceibal



TEMPUS



UNIVERSITY of CAMBRIDGE
ESOL Examinations



Pearson



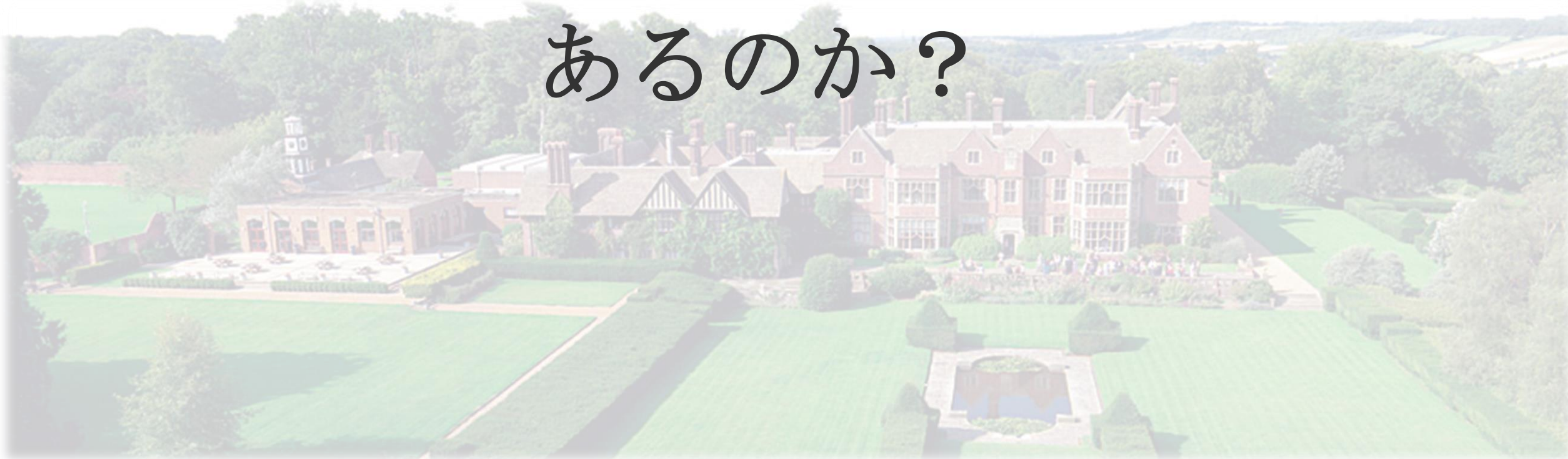
EnglishScore

本日の話の流れ



- どんなスピーキングテストがあるのか？
- スピーキングテストは何を測っているのか？
- スピーキングテストはどのように作られているのか？
- スピーキング能力向上につながるスピーキングテストとは？

どんなスピーキングテストが
あるのか？



Q & A (Password Speaking test)

There are five simple questions to answer in section 2. Test-takers have 20 seconds to answer each question.

EXAMPLE 1

Where do you live?

Answer: *I live in London next to Hyde Park.*

EXAMPLE 2

What is the weather like today?

Answer: *It's warm and sunny.*

Individual long-turn task (IELTS)

Describe a memorable event in your life.

You should say:

**when the event took place
where the event took place
what happened exactly**

and explain why this event was memorable to you.

You will have to talk about the topic for 1 to 2 minutes.

You have one minute to think about what you're going to say.

You can make some notes to help you if you wish.

Prepared monologue with Q&A (Trinity ISE Speaking & Listening)

TRINITY
COLLEGE LONDON

Integrated Skills in English Topic Form – ISE Foundation

What we did and saw

The activity I enjoyed the most

Title of topic:
My school trip to Paris

What we're going to do next year

Some differences between Paris and my city

The information on this form must be presented to the examiner during the exam.
Please note, candidates do not have to use the Trinity topic form, they can make their own with four points for discussion.

Voicemail task (Oxford Test of English)

OXFORD
UNIVERSITY PRESS

Back to main menu

Speaking Part 2



Speaking Part 2

2

Voicemail 2

You are going to reply to a voicemail message. First read and listen to the task and decide what you want to say.

You have 40 seconds to leave your voicemail. Start speaking when you hear the tone.

Listen to the message from your friend about a free course she has won as a prize. Then, leave a voicemail message for your friend. In your message, you should:

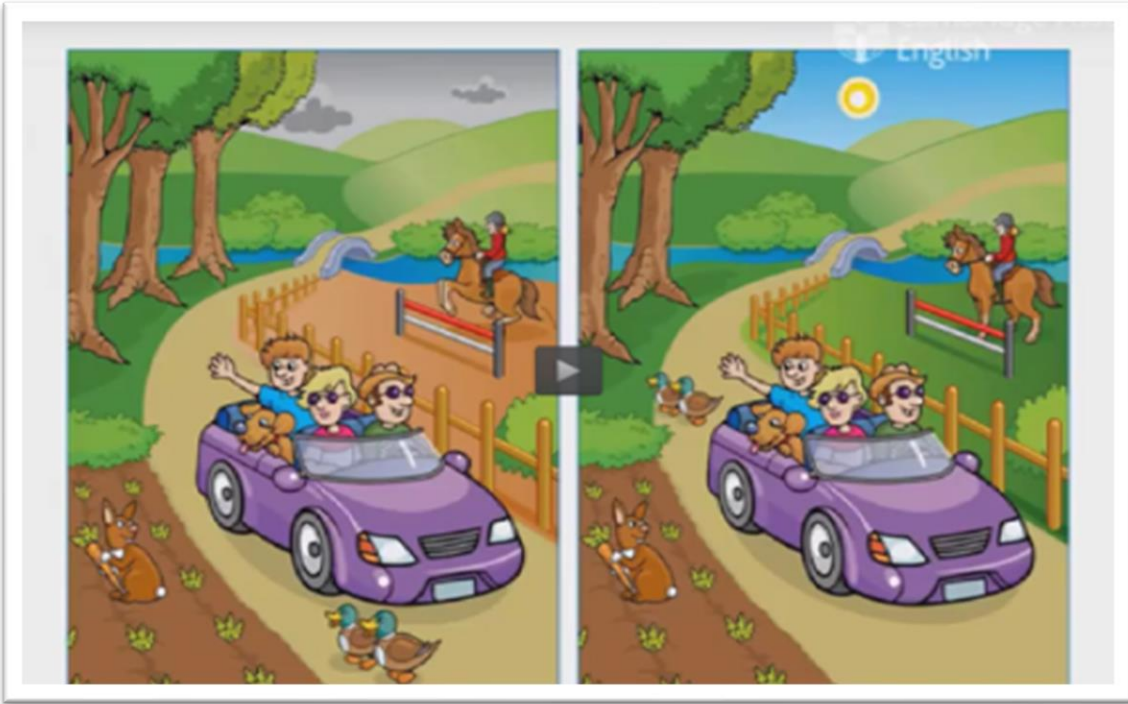
- congratulate your friend
- ask some questions about the course
- say what you think your friend should do.

Now listen to the message.

You now have 20 seconds to think about what you want to say.

Picture comparison task (Cambridge YLE Movers)

Look at these pictures. They look the same, but some things are different. Here, there are three trees, but here there are two trees. What are the different things can you see?



Multiple picture narration task (Cambridge YLE Movers)

These pictures show a story. It's called 'a clever dog'. Look at the pictures first. Holly and John are playing table tennis in the garden. The baby and the dog are there too. Mum is saying, 'Please watch the baby for me'. Now you tell the story.



Listen & Repeat task (PTE Academic)

After listening to a recording of a sentence, repeat the sentence.

Repeat sentence

You will hear a sentence. Please repeat the sentence exactly as you hear it. You will hear the sentence only once.

Audio Player

Playing

0:00 / 0:04

Most teaching staff make their lectu

Retelling task (PTE Academic)

After listening to or watching a lecture, re-tell the lecture in your own words.



PTE Academic - Speaking: Retell Lecture

Re-tell lecture

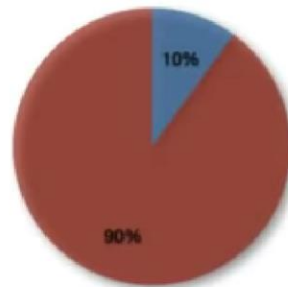
1 of 3

00:10 / 09:00

Watch

You will hear an interview. After listening to the interview, in 10 seconds, please speak into the microphone and retell what you have just heard from the lecture in your own words. You will have 40 seconds to give your response.

Distribution of large fish 2005 - 06



● Remaining in the oceans
● Caught and consumed by humans

Audio Player

Playing

0:07 / 1:05

Info gap task (Cambridge YLE Flyers)

George and Holly have new pets. I don't know anything about Gorge's pet, but you do. So I'm going to ask you some questions.

What animal does George have? / What's Gorge's pets' name? / When did George get his pet? / Is it quiet or noisy? / What colour is it?

Now you don't know anything about Holly's pet. So, you ask me some questions.



George's pet

Name	Socks
What animal	puppy
Colour	black
When / get	last weekend
Quiet / noisy	noisy

Holly's pet

Name	?
What animal	?
Colour	?
When / get	?
Quiet / noisy	?

Paired discussion task (Cambridge B2 First)

Happiness task (5 minutes)

Here are some of the things in life which can affect our happiness.

1. Talk to each other about how **important** these things are for a happy life.
2. Decide which two are the most **important**.



タスク・フォーマットの例

Read-aloud (e.g. TOEFL Jr)	Multiple picture narration/description (e.g. TOEIC Bridge)	Prepared monologue with Q&A (e.g. Trinity Integrated Skills in English)
Listen & repeat (e.g. PTE Academic)	Information gap (e.g. Michigan Young Learners English Test)	Role-play with examiner (e.g. ACTFL Oral Proficiency Interview)
Retelling with written or audio input (e.g. Versant)	Q&A - short / long (e.g. IELTS)	Discussion with peer(s) (e.g. General English Proficiency Test, Taiwan)
Single picture description / Name pictures (e.g. Uruguay Plan Ceibal Speaking Test)	Monologue (e.g. Aptis Speaking Test)	Discussion with peer(s) and examiner (e.g. Cambridge General English Test)
Picture comparison/ spot the differences (e.g. Cambridge Young Learners English Test)	Voicemail (e.g. Oxford Test of English)	Integrated Listening (& Reading) into Speaking (e.g. TOEFL)

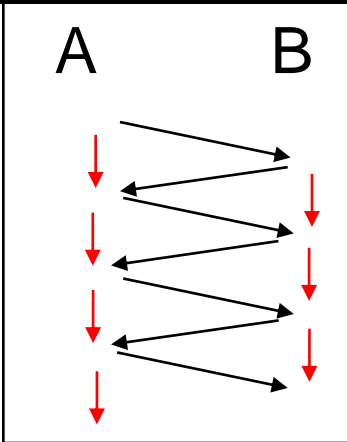
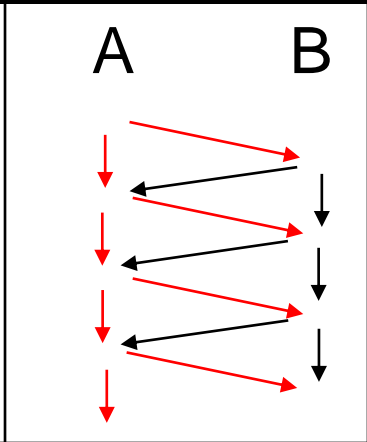
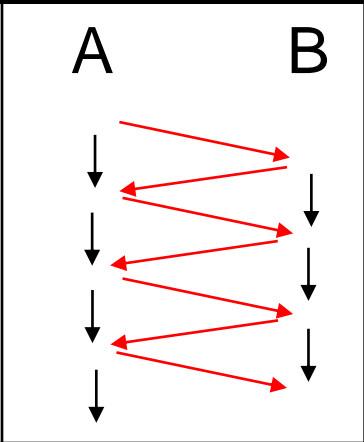
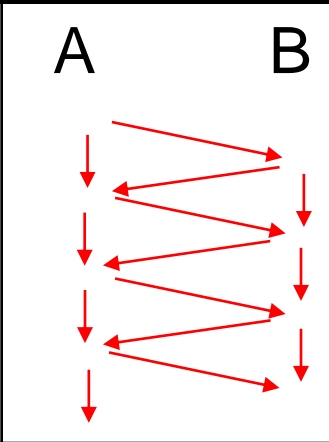
Interaction (やり取り)

Production
(一方通行の発話)

Mechanical
Reproduction
(機械的な再生)

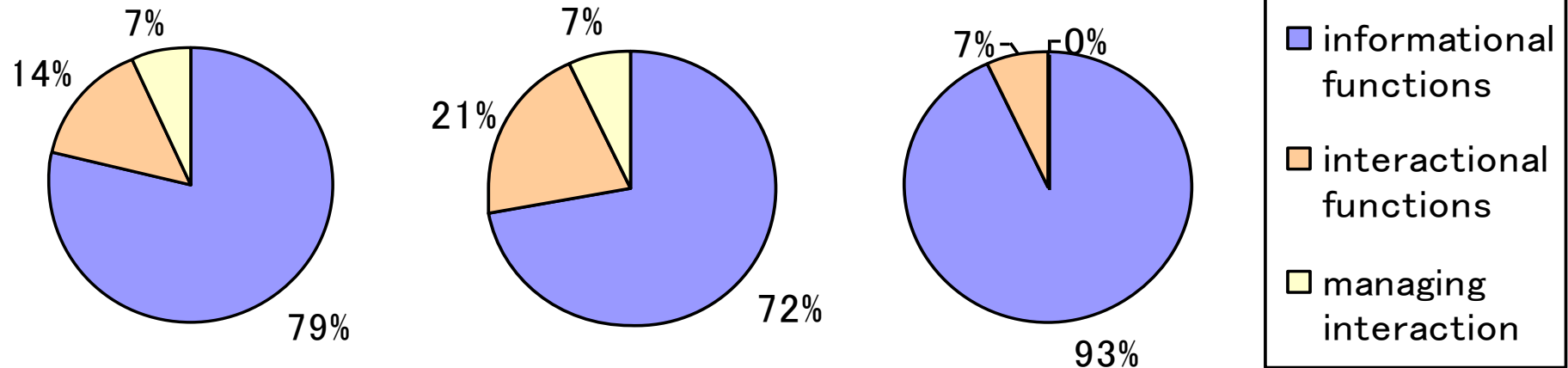
インタビュー形式と会話形式のディスコースの違い

(van Leir,1989;Young & Milanovic,1992; Young,1995; Kormos, 1999)

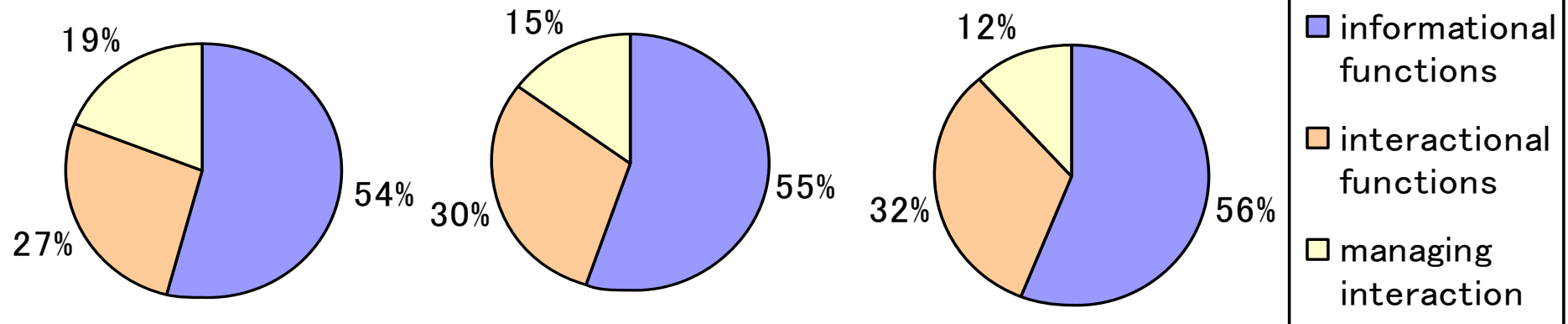
時間 ↓	A B 	A B 	A B 	A B 
	Pseudo-Contingency	Asymmetrical Contingency	Reactive Contingency	Mutual Contingency
インタラクションの例	<ul style="list-style-type: none"> ・ 劇 ・ 挨拶などの決まり文句 	<ul style="list-style-type: none"> ・ インタビュー ・ (伝統的な形式の) 授業 	<ul style="list-style-type: none"> ・ 特別な目的のない会話 	<ul style="list-style-type: none"> ・ 交渉 ・ ディスカッション・ディベート

インタビュー形式とペア形式のスピーキングテストで 引き出される言語機能の比較 (ffrench, 2003, p.413)

インタビュー形式で引き出された言語機能



ペア形式で引き出された言語機能



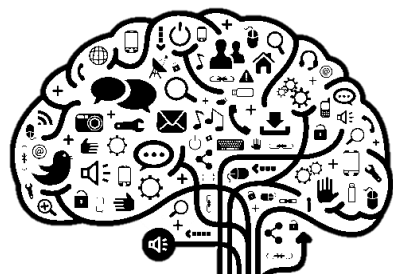
スピーキングテストは
何を測っているのか？



第二言語スピーキング能力

認知的な側面

言語知識と
言語プロセス



測りたいスピーキング・コンストラクトをテストの目的によって決める。

← テストのスコアは何に使われるのか？

社会的な側面

社会的なコンテキスト
での言語使用



スピーキング能力の認知的な側面

Speaker's general goals
World knowledge
Knowledge of listener/situation
Recall of discourse so far



Conceptualisation (何を話そう?)

Syntactical & lexical knowledge
Pragmatic knowledge
Knowledge of formulaic chunks



Grammatical Encoding (どういう語彙・文法?)

Lexical knowledge
Phonological knowledge



Phonological Encoding (音声体系・構造?)

Knowledge of articulatory settings



Phonetic Encoding (どうやって音を出す?)

[Articulation: 発話]

Speaker's general goals
Recall of discourse so far



Self-Monitoring (発話は合ってる?)

(Levelt 1989; Field 2011)

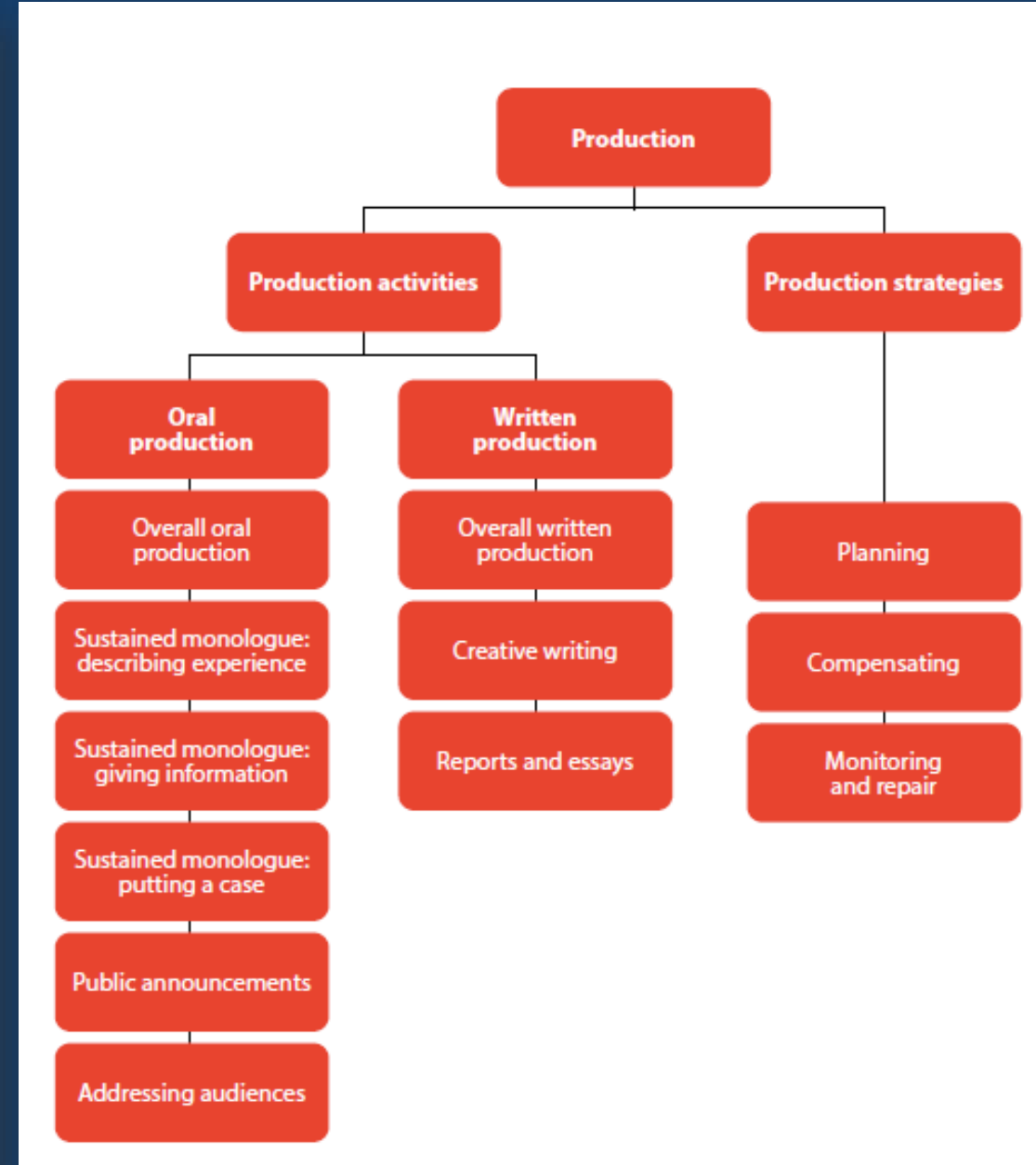
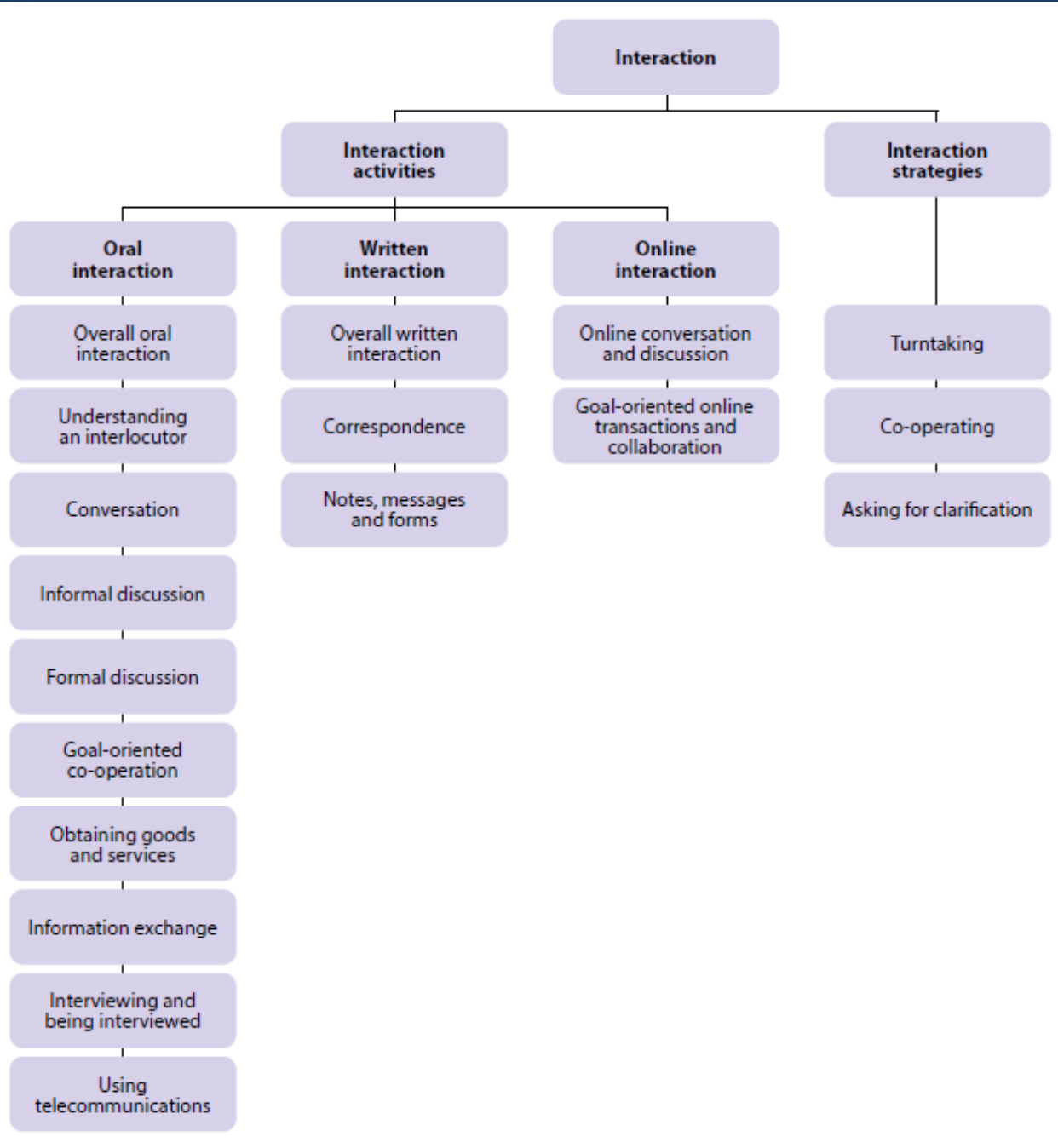
スピーキング能力の社会的な側面

コミュニケーションは‘ダイナミック (rather than static)であり、話者間の意味の交渉によって左右され、コンテクストに特有である。’

Savignon (1983: 8-9)



- 社会的コンテクストの中で話者同士が共同でやり取りを作り上げる (co-construction of interaction)
- コミュニケーションは話者同士が共同で達成(a shared achievement)



評価基準

IELTS™

SPEAKING: Band Descriptors (public version)

Band	Fluency and coherence	Lexical resource	Grammatical range and accuracy	Pronunciation
9	<ul style="list-style-type: none"> speaks fluently with only rare repetition or self-correction; any hesitation is content-related rather than to find words or grammar speaks coherently with fully appropriate cohesive features develops topics fully and appropriately 	<ul style="list-style-type: none"> uses vocabulary with full flexibility and precision in all topics uses idiomatic language naturally and 	<ul style="list-style-type: none"> uses a full range of structures naturally and appropriately 	<ul style="list-style-type: none"> uses a full range of pronunciation features with precision features throughout

Cambridge B2 First Speaking rating scale

B2	Grammar and Vocabulary	Discourse Management	Pronunciation	Interactive Communication
5	Shows a good degree of control of a range of simple and complex grammatical forms. Uses a range of appropriate vocabulary to give and express views on a wide range of topics.		Is intelligible. Intonation is appropriate. Sentence and word stress is	Initiates and responds appropriately, linking contributions to those of

TRINITY
COLLEGE LONDON

ISE II Speaking and listening rating scale

Score	Communicative effectiveness	Interactive listening	Language control	Delivery
	<ul style="list-style-type: none"> Task fulfilment Appropriacy of contributions /turn-taking 	<ul style="list-style-type: none"> Comprehension and relevant response Level of understanding Speech rate of examiner 	<ul style="list-style-type: none"> Range Accuracy/precision Effects of inaccuracies 	<ul style="list-style-type: none"> Intelligibility Lexical stress/intonation Fluency Effects on the listener

- どんなタスク・評価基準が適切かは、どんな目的のテストなのか、テストスコアを何に使うのか、どんなスピーキング・コンストラクトを測りたいのか次第
- Standardised testsでは、タスクを組み合わせて測りたいコンストラクトの様々な側面をカバーしたり、テストの信頼性を高める

5	<ul style="list-style-type: none"> uses a full range of structures naturally and appropriately uses a full range of pronunciation features with precision
4	<ul style="list-style-type: none"> uses a full range of structures naturally and appropriately uses a full range of pronunciation features with precision
3	<ul style="list-style-type: none"> uses a full range of structures naturally and appropriately uses a full range of pronunciation features with precision
2	<ul style="list-style-type: none"> uses a full range of structures naturally and appropriately uses a full range of pronunciation features with precision

5	Shows a good degree of control of a range of simple grammatical forms. Uses a range of appropriate vocabulary to give and express views on a wide range of topics.		Is intelligible. Intonation is appropriate. Sentence and word stress is	Initiates and responds appropriately, linking contributions to those of
4	Shows a good degree of control of a range of simple grammatical forms. Uses a range of appropriate vocabulary to give and express views on a wide range of topics.		Is intelligible. Intonation is appropriate. Sentence and word stress is	Initiates and responds appropriately, linking contributions to those of
3	Shows a good degree of control of a range of simple grammatical forms. Uses a range of appropriate vocabulary to give and express views on a wide range of topics.		Is intelligible. Intonation is appropriate. Sentence and word stress is	Initiates and responds appropriately, linking contributions to those of
2	Shows a good degree of control of a range of simple grammatical forms. Uses a range of appropriate vocabulary to give and express views on a wide range of topics.		Is intelligible. Intonation is appropriate. Sentence and word stress is	Initiates and responds appropriately, linking contributions to those of

Computer-delivered speaking testsで コミュニケーション能力は測れる?



- やりとりをするコンストラクトを、どうやって実現するか (e.g. Galaczi & Taylor, 2018)
- コミュニケーションのAuthenticityをどうやって再現するか (e.g. May, 2018)

パンデミックにおけるスピーキングテストの発展

Test Review

Test review: Current options in at-home language proficiency tests for making high-stakes decisions

Daniel R. Isbell
University of Hawai'i at Mānoa, USA

Benjamin Kremmel
University of Innsbruck, Austria

LANGUAGE TESTING

Language Testing
1–20
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0265532220943483
journals.sagepub.com/home/ltj

LANGUAGE ASSESSMENT QUARTERLY
2021, VOL. 18, NO. 1, 1–5
<https://doi.org/10.1080/15434303.2020.1866576>

Routledge
Taylor & Francis Group

An Overview of COVID-19's Impact on English Language University Admissions and Placement Tests

Gary J. Ockey

Iowa State University, Ames, Iowa, USA

Assessment in Education:

Principles
 Policy &
 Practice

ASSESSMENT IN EDUCATION: PRINCIPLES, POLICY & PRACTICE
2021, VOL. 28, NO. 4, 343–349
<https://doi.org/10.1080/0969594X.2021.2004530>

Routledge
Taylor & Francis Group

EDITORIAL

Use of innovative technology in oral language assessment

可能な解決案

a) ビデオコール・デリバリー



Trinity ISE Online
<https://www.trinitycollege.it/inglese/esami-in-videoconferenza/>

c) バーチャル環境

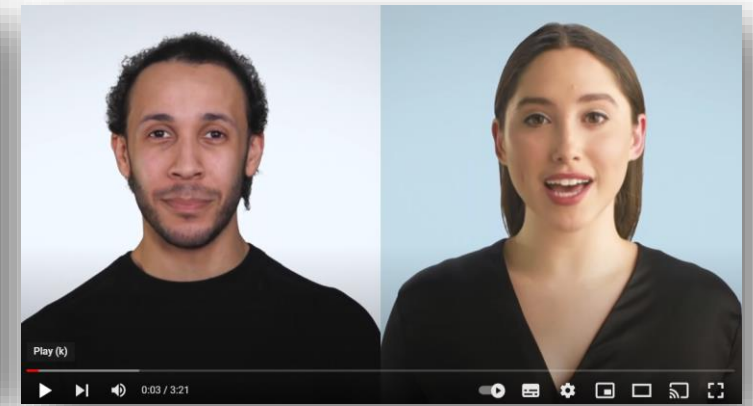


(Ockey et al. 2017)

b) 音声対話システム



Matsuyama (2022)
<https://www.teai-waseda.jp/assessment/>



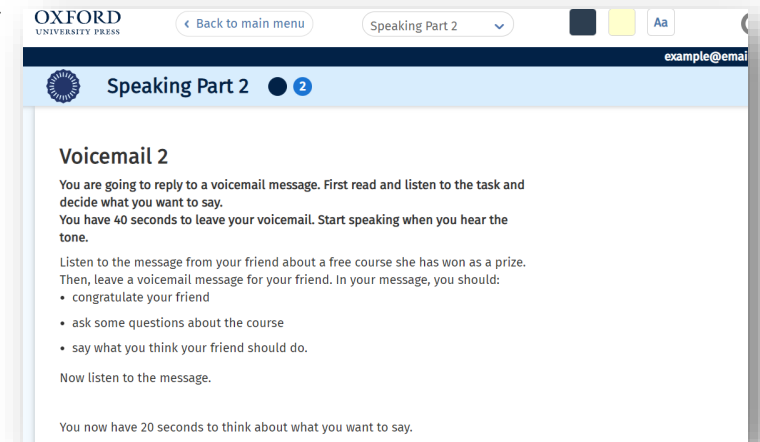
GPT-3: Two AIs talk about becoming human
<https://www.youtube.com/watch?v=jz78fSnBG0s>

d) virtual reality (VR) & augmented reality (AR)



Mondly AR
<https://www.youtube.com/watch?v=9P0t9JI22y8>

e) 注意深くデザインされたsemi-direct タスク



Oxford Test of English
<https://fdslive.oup.com/www.oup.com/elt/general/content/global/ote/demo-v3/#/exercise/speaking/2/1>

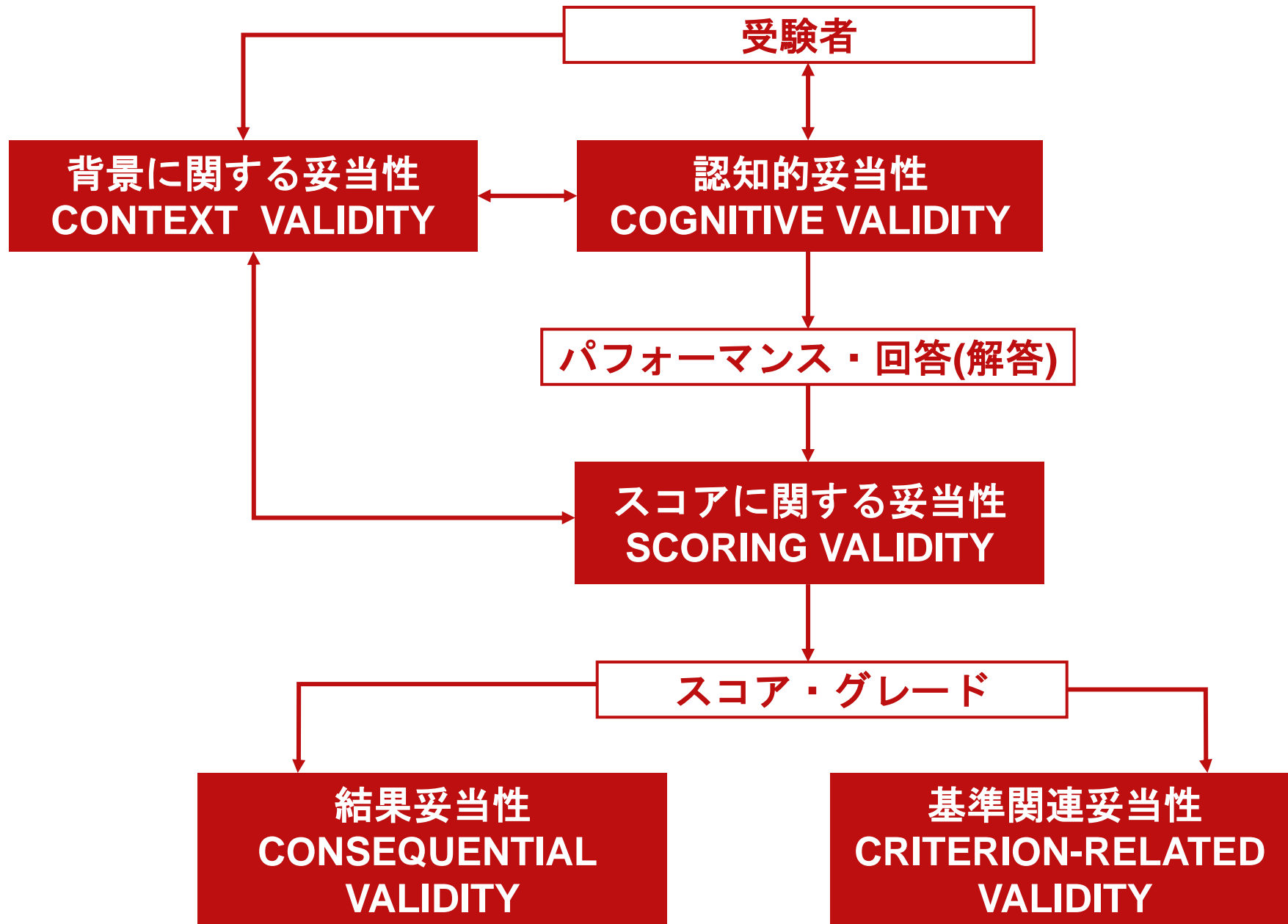
スピーキングテストはどのように
作られているのか？



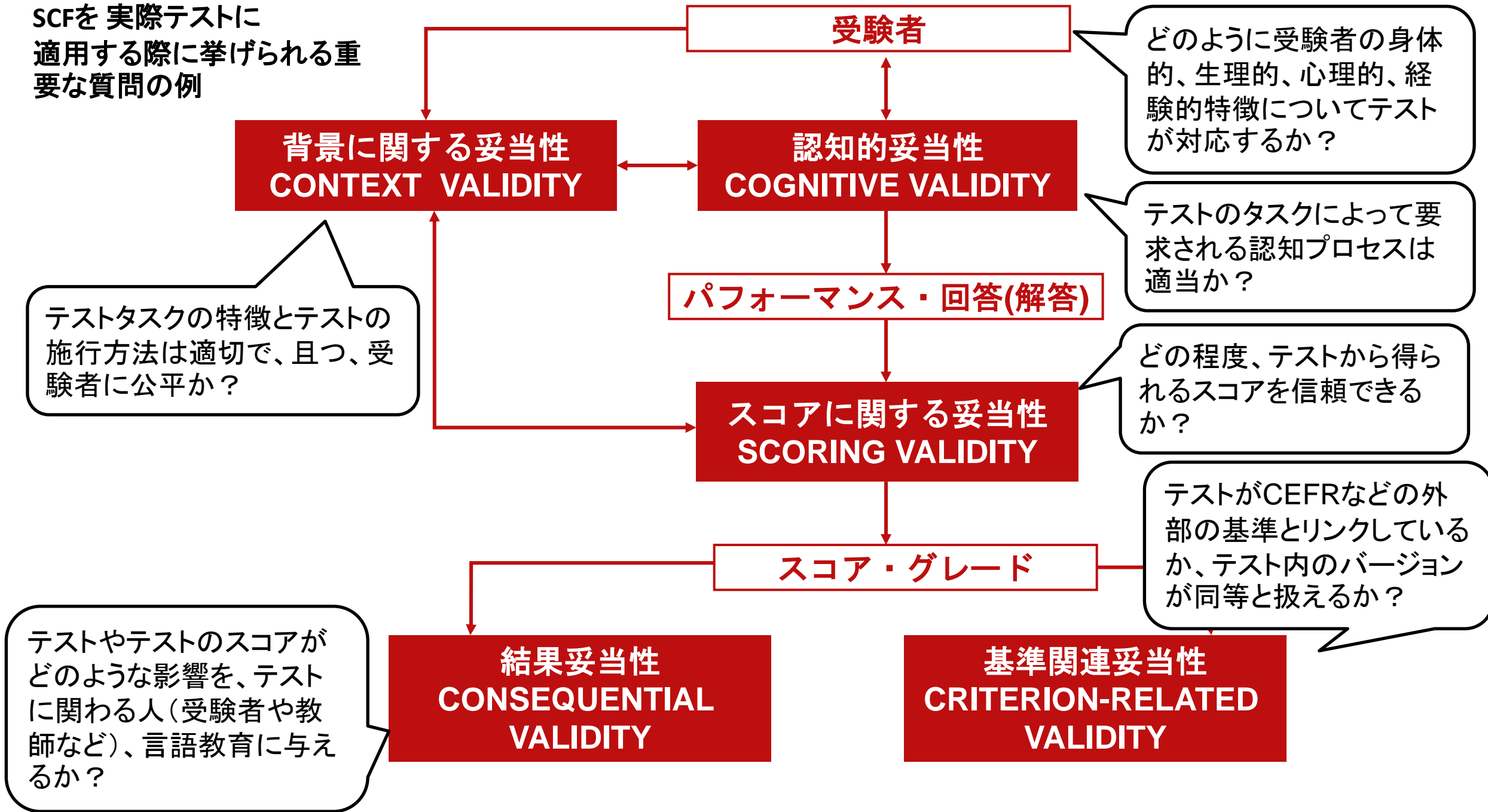
テストの妥当性を検証するための Weir (2005)の「社会的認知的枠組み」 (Socio-cognitive Framework)



- Weir, C. J. (2005). Language testing and validation: An evidence-based approach. Palgrave Macmillan.
- Taylor, L. (ed.) (2011). Examining speaking, Cambridge University Press.



SCFを 実際テストに
適用する際に挙げられる重
要な質問の例



「社会的認知的枠組み」の主な特徴

- 受験者の能力を認知的なプロセスの現われとして捉えていて、タスクを行う時の言語使用を純粋な言語的現象というより社会的現象として捉えている(CEFRの考え方と共通)。
- フレームワークは1つのテストのために妥当性の証拠を集めるための統合されたアプローチを示している。様々な妥当性の構成要素がどのように時間的・概念的に合わさっているかを描いている。
- 矢印は主な仮定される関係(どの要素がどの要素に影響を与えるか)を表しており、時間枠は上から下へと流れている。
- テスト作成の最初の段階で妥当性検証が行われるべきcontext validity と cognitive validity と、テスト作成の後半に検証されるであろうscoring validity, consequential validity, criterion-related validityから成り立つ。

TEAPスピーキングテスト開発の例 (4つの開発・妥当性研究: 2010-2014)



- 文献調査
- 言語機能調査
 - 高校教員(n=167) & 大学教員 (n=24); O'Sullivan et al. (2002)のチェックリスト
- 様々なステークホルダーとの意見・情報交換
- テストの仕様、評価基準、試験官ガイドラインのドラフト
- ミニ・トライアル
- 小規模トライアル
- 大規模パイロット
- 採点方法トライアル

受験者とインタビューアーからのフィードバック

受験者の発話サンプルから、引き出された言語機能の分析

背景に関する妥当性
CONTEXT VALIDITY

認知的妥当性
COGNITIVE VALIDITY

受験者のパフォーマンスの言語・談話の特徴の分析による評価基準・記述の検証

スコアの統計分析による評価基準・記述の検証

スコアに関する妥当性
SCORING VALIDITY

採点者からのフィードバック & 採点後のフォーカスグループによる採点プロセス・採点の焦点の検証

テストの概要

タスク形式		レベル	ターゲットする言語機能 ← 高校、大学教員のニーズ分析
1	短いQ&A	A2/B1	<ul style="list-style-type: none">• 個人的な情報を与える
2	ロール・プレイ	B1	<ul style="list-style-type: none">• インタラクションを始める• 情報・意見を尋ねる• コメントする
3	モノローグ	B1/B2	<ul style="list-style-type: none">• 同意・反対する• 理由を述べる• 詳しく述べる
4	長いQ&A	B2	<ul style="list-style-type: none">• 意見を述べる• 理由を述べる• 比較する• 推測する• 詳しく述べる

ターゲット & 実際に引き出された言語機能

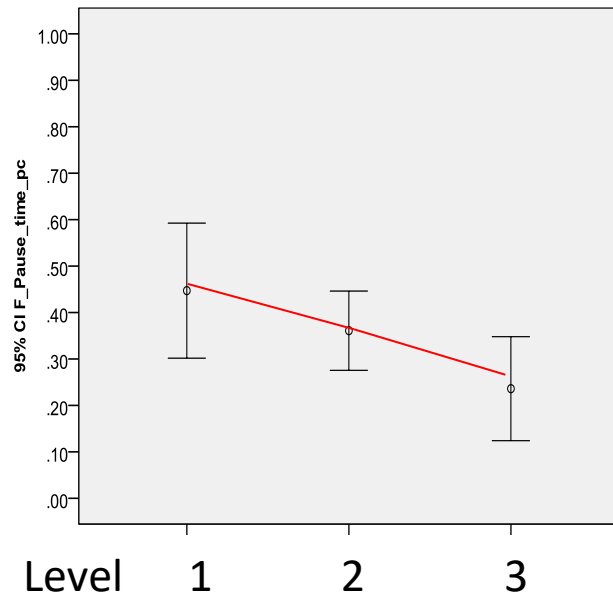
	パート 1		パート 2	
	ターゲット	引き出された	ターゲット	引き出された
個人的な情報を与える(現在)	✓	✓		
個人的な情報を与える(過去)	✓	✓		
個人的な情報を与える(将来)	✓	✓		
意見・好みを述べる		✓		
詳しく述べる		✓		
意見を尋ねる			✓	✓
情報を尋ねる			✓	✓
コメントする			✓	✓
許可を求める				✓
挨拶する		✓		✓
感謝する		✓		✓
意味の交渉をする: 理解を示す				✓
インタラクションを始める			✓	✓
相手の会話参加を促す				✓

	パート 3		パート 4	
	ターゲット	引き出された	ターゲット	引き出された
個人的な情報を与える(現在)				✓
意見・好みを述べる			✓	✓
詳しく述べる	✓	✓	✓	✓
理由を述べる	✓	✓	✓	✓
比較する			✓	
推測する			✓	✓
同意する	✓	✓		
反対する	(✓)			
感謝する		✓		✓
意味の交渉をする:理解を確かめる				✓
意味の交渉をする:聞き返す				✓

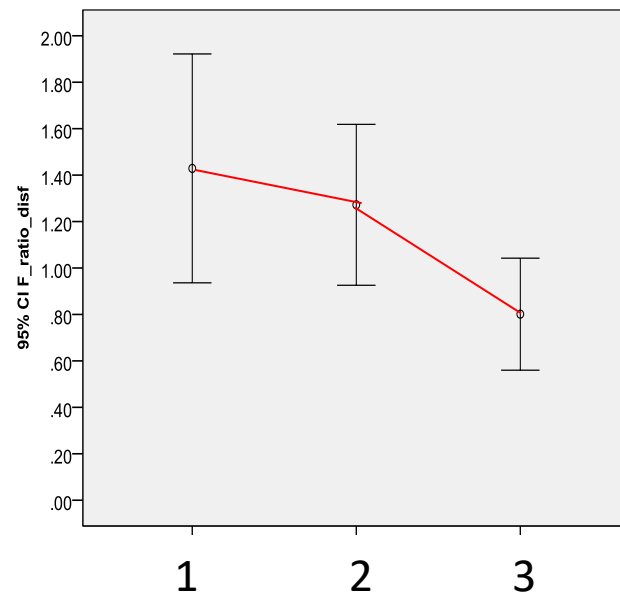
Corresponding rating category	Focus	Measure	Parts of the test applied
a. Grammatical range and accuracy	Complexity	Ratio of subordinate clauses to AS-units	1, 2, 3, 4
		Number of words per AS-unit	1, 2, 3, 4
	Accuracy	Percentage of error-free AS-units	1, 2, 3, 4
b. Lexical range and accuracy	Range	Lexical frequency coverage (K1+ K2 words)	1, 2, 3, 4
		Academic Word List coverage	1, 2, 3, 4
	Accuracy	(Lexical accuracy was not measured in this analysis)	
c. Fluency	Hesitation	Number of unfilled pauses (utterance initial) per 50 words	1, 2, 3, 4
		Total pause time as a percentage of speaking time	3
	Disfluency	Ratio of repair, false starts, and repetition to AS-units	1, 2, 3, 4
Temporal	Speech rate in Part 3	3	
	Articulation rate in Part 3	3	
d. Pronunciation	L1 influence	Number of words pronounced with noticeable L1 influence (katakana-like) as percentage of total words produced	1, 2, 3, 4
e. Interactional effectiveness	Length of response	Average words per response	1, 4
	Number of extra questions	Number of separate questions asked that were not on required list in Part 2	2
	Back-channelling and comments	Number of instances of back-channelling and comments in Part 2	2
f. Others – the amount of talk	Length of long turn	Total number of words produced in Part 3	3
	Total production	Total amount of production across all parts of the test, measured in words	1, 2, 3, 4
		Total number of AS-units produced across all parts of the test	1, 2, 3, 4

流暢さの指標の例

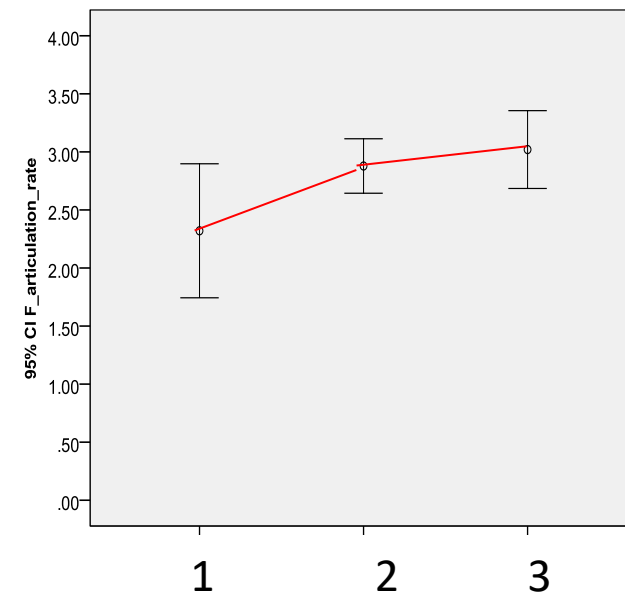
発話の始めのポーズの数
(50語における)



言い直し、繰り返しの割合
(ASユニットの数に対して)



Articulationのスピード



Nakatsuhara, F. (2014). *A Research Report on the Development of the Test of English for Academic Purposes (TEAP) Speaking Test for Japanese University Entrants – Study 1 & Study 2.* www.eiken.or.jp/teap/group/pdf/teap_speaking_report1.pdf

Nakatsuhara, F., Joyce, D., & Fouts, T. (2014). *A Research Report on the Development of the Test of English for Academic Purposes (TEAP) Speaking Test for Japanese University Entrants – Study 3 & Study 4.* www.eiken.or.jp/teap/group/pdf/teap_speaking_report2.pdf

スピーキング能力向上につながる スピーキングテストとは？



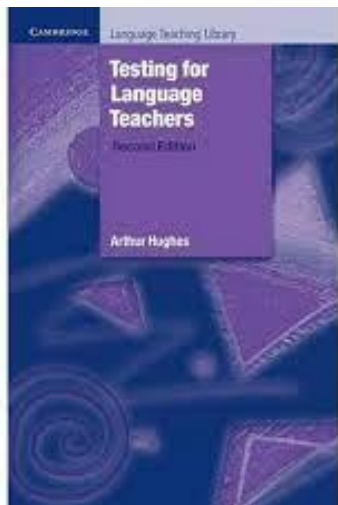
テストの波及効果: テストの教授・学習に与える影響

ポジティブ・ネガティブな波及効果

- 「どんなに素晴らしいテストでも完璧にコンストラクトを測ることはできない。なぜなら、コンストラクトの理論によって含まれていないといけない何かを省いているか (construct under representation) 省かなければいけない何かが含まれてしまうか (construct-irrelevant variance)、またはその両方のことが起きてしまうからだ。」 (Messick, 1989, p.36)
- 「その二つの可能性をなるべく減らすことが、ポジティブな波及効果を促すのに不可欠」 (Messick, 1996)

「有益な波及効果を生むテストを導入するのには時間もお金もかかります。しかし、そんなコストにかける余裕がないと結論づける前に、有益な波及効果を生むテストを導入しない場合のコストはどうなるか、自問してみてください。有益な波及効果を生むテストの導入コストと、そのテストがないために教師と生徒が本来の学習目標を達成するのに不適切な活動を行っていることによる労力と時間の浪費を比較すると、有益な波及効果を生むテストを導入しないままにいる余裕なんてない、と判断するのではないのでしょうか。」

Hughes (2003, p.56) Testing for Language Teachers



テストが学習や 教育に良い波及 効果を与える 条件

(Cheng, 2005; Green, 2007; Hughes, 2003; Wall 2005)

- **発達を促進したい能力をテストすること**: テストとターゲットの場面での言語能力の要求が、かなり多くオーバーラップしていることが必要
- 直接的なテストを使うこと
- Criterion-referenced テストを使うこと
- 幅広く、そして予測がつかないサンプリングを行うこと
- **テストが生徒と教師に熟知されていること**
- テストでの成功が大事だと考えられていること
- テストでの成功が難しい、しかし成功が可能であると考えられていること
- **教師に補助を行うこと; 新しいテストのコンテキストや方法、有効な教授法の知識について教師をトレーニングするのが不可欠**
- **適切な教材のサポートは、無料で行われるべき**

研究の紹介

- やりとりをする能力をどうやって測る？リソースがなくてもCBTって作れる？
- 英語民間試験が英語教育リフォームを推進するためには、テストに関わる全てのステークホルダーの理解が大事。その第一歩の研究。

Eliciting pragmatic and interactional competence in semi-direct speaking tests

語用論的言語能力とやりとりの能力をローテクなCBTで
引き出せるか



Nakatsuhara, F., May, L., Inoue, C., Willcox-Ficzere, E., Westbrook, C., & Spiby, R. (2021). *Exploring the potential for assessing interactional and pragmatic competence in semi-direct speaking tests*, Aptis Validation Report. British Council.
https://www.britishcouncil.org/sites/default/files/nakatsuhara_et_al_2021_-_assessing_interactional_and_pragmatic_competence_in_semi-direct_speaking_tests_v2.pdf

リサーチ・デザイン(やりとりの能力の部分のみ)

やりとりの能力を測るテストの仕様の開発

やりとりのタスクを2つ作成



パイロットと修正・変更



ベンチマーキングタスク一つ&2つのやり取りタスクの実施 (N=48) 回答の録音

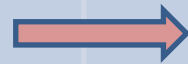
中国: 24 x B1-Cレベルの受験者

オーストリア: 24 x B1-Cレベルの受験者



受験者からのフィードバック

アンケート (N= 48)



フォローアップ・インタビュー (N=12)

やりとりのタスクの作成

■ やりとりの能力のどの特別な要素をターゲットにするかの決定

(e.g. Nakatsuhara, May, Lam & Galaczi, 2018; Willcox-Ficzere, 2019)

■ 主な原則:

- Authenticなコンテキストをビデオ・イメージで与える
- 「聞き返し」をタスクに入れる (“Sorry I didn’t get the last point. Can you say that again, perhaps using different words or giving an example?”)
- テストの指示、インプット言語をCEFRレベルでコントロールする (EVP/EGP)
 - 受験者: B1, B2, C
 - ターゲットのアウトプットレベル = B2 (インプット言語 = B1)



■ 「社会的認知的枠組み」のテンプレートを使った**テスト仕様**の作成

(e.g. Weir, 2005; O’Sullivan et al., 2020)

TEST	Aptis Research Pilot Test		COMPONENT		Speaking	Task	IC (a)/(b)
Features of the Test							
Skills focus	The task is designed to tap into four features of IC: <i>responding to a partner, negotiating towards a joint outcome, interactive listening and negotiating meaning</i> . More specifically, the task measures candidates' ability to: <ul style="list-style-type: none"> disagree and put forward a different point of view effectively and provide justification effectively link their own contribution to the partner's work towards a decision by trying to persuade the partner acknowledge partner's views demonstrate they have been listening carefully/attentively through responding appropriately to the partner's idea be able to clarify/rephrase their points 						
Task level (CEFR)	A1	A2	B1	B2	C1	C2	
Task description	Candidates will first watch a video in which a conversational partner expresses views, and then will be asked to express a differing opinion while appropriately responding to and persuading the partner, also using interactive listening skills. Key words from the partner's points as well as the points that the candidates are required to make will appear on screen. After the completion of the response time, an automated video clip will be played to ask a clarification question, to which the candidate will be required						

やりとりのタスクの仕様

Nature of information	Only concrete	Mostly concrete	Fairly abstract	Mainly abstract		
Relevant domain	Public	Occupational	Educational	Personal		
Topic	From topic list for B2					
Topic familiarity	Familiar			Unfamiliar		
Discourse mode	Descriptive	Biographical	Expository	Argumentative	Instructive	Quasi-Interactive

Skills focus

The task is designed to tap into four features of IC: *responding to a partner, negotiating towards a joint outcome, interactive listening and negotiating meaning*. More specifically, the task measures candidates' ability to:

- 相手の意見に反対し、異なった観点を効果的に伝え、理由を述べる
- 自分の発話を相手の発話に効果的にリンクする
- 相手を説得し、意見の一致を促す
- 相手の意見に理解を示す
- 相手のアイディアに適切に回答することを通じて、注意深く聞いている姿勢を見せる
- 自分のポイントを言い換えて分かりやすく再度伝える

Video prompt	to switch between aural and written input modes, written information on the screen should be of a supportive nature throughout the task. After the completion of the response time (i.e. 75 seconds), an automated video clip is played to ask a clarification question, to which the candidate is required to respond. A conversational partner's talk: Between 140-160 words; No key information in the first 30 words; Start with some contextual information and then share two main points; The prompt to finish with 'What do you think?' There will be a standardised clarification request: "Sorry, I didn't get the last point. Can you say that again, perhaps using different words or giving an example?"										
Length of written prompt	Two bullet points for a conversational partner's views. Two bullet points for the candidate's views. Each of the candidate's bullet point is no longer than 12 words.										
Lexical level	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	TBC
Grammatical range	Most lexical items up to B1 according to the English Vocabulary Profile list										
Content knowledge	Most grammatical structures up to B1 according to the English Grammar Profile list										
Cultural specificity	General										
	Neutral										
	Specific										

	Elaborating*	Modifying/commenting*	Reciprocating
	Justifying opinions	Asking for opinions	Deciding
	Comparing	Persuading	
	Speculating	Asking for information	
	Staging	Conversational repair	
	Describing	Negotiation of meaning	
	Summarising		
	Suggesting		
Rating scale for task	A task-specific holistic rating scale is used for the task. The rating scale is a 7-point scale from 0-6. A B2-level performance is required to achieve score bands 3-4. A score of 5 or 6 is awarded for performances beyond B2 level, with a 5 describing performance equivalent to a C1 level, and 6 for performances at a C2 level. (Suggestions for descriptors and rating methods to be made after the research)		

*functions targeted: those highlighted in grey to be finalised after piloting

IC(a)

In a classroom, Jan starts talking to you.



IC(a)

Use the two points to persuade Jan:

- **Group presentations only work if all team members work hard**
- **Probably unfair, if all team members receive the same score**



Now Jan has a question for you.

You have 30 seconds to answer after the beep.

IC(a)

In this part, you will watch the video of your classmate Jan, who will argue for **the use of group presentations**.

Jan will make **2 points** in her talk.

After you watch Jan's talk, you will be given **the 2 points that you're going to make**. You will have:

- **60 seconds to prepare** your argument, and
- **75 seconds to speak**.

Please talk to Jan in a way that **shows that you understand Jan's points**, and **use the two points to try to persuade** Jan to change her opinion.

After you have finished speaking, you will hear **a question from Jan**. You will be given **30 seconds to answer** the question.

IC(a)

Use the two points to persuade Jan:

- **Group presentations only work if all team members work hard**
- **Probably unfair, if all team members receive the same score**



You now have **75 seconds to talk to Jan**, using the two points. Remember to show your understanding of her points.

Speaking time: **00:75**

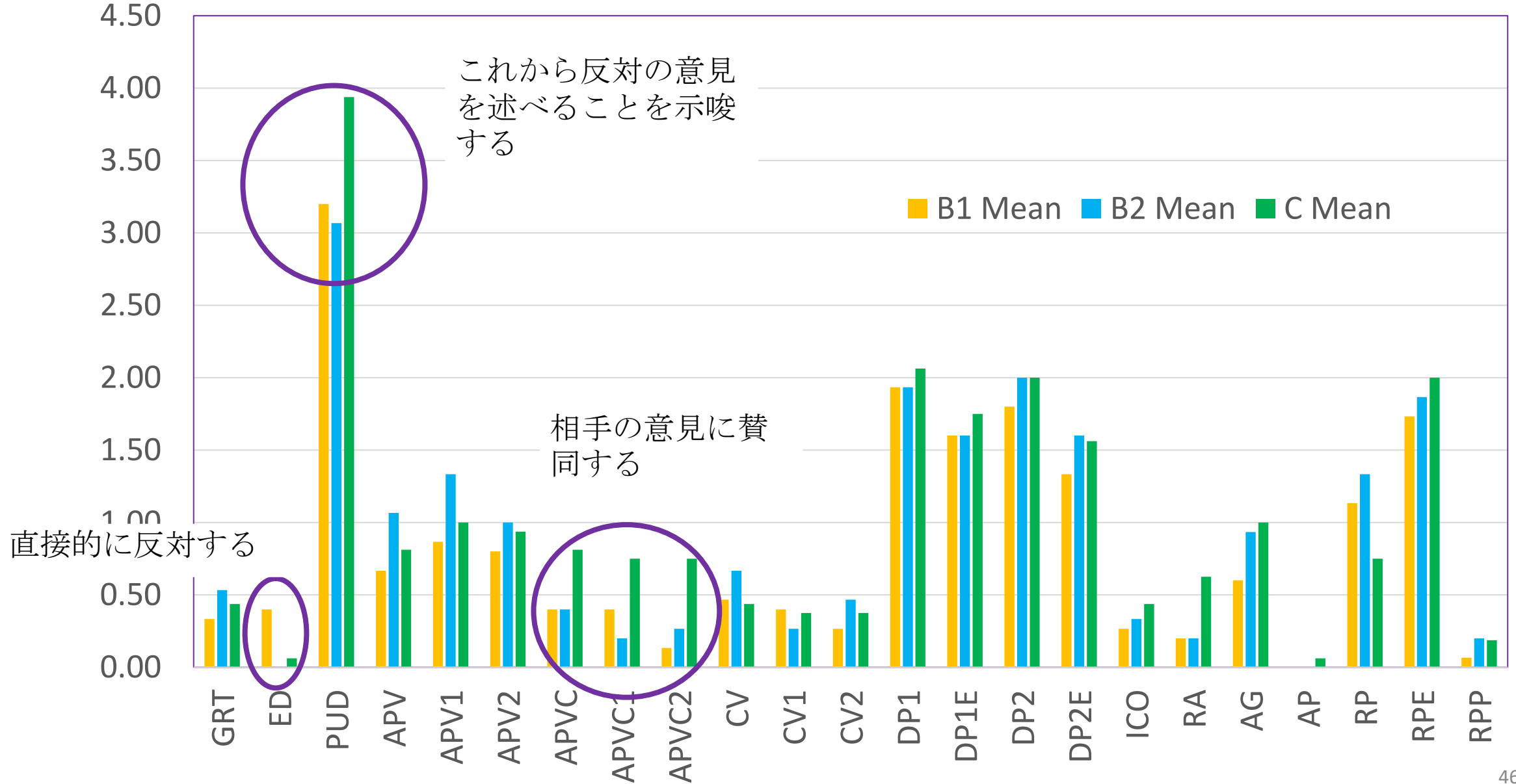
分析

- **受験者のレベルを判断:** ベンチマークタスク(Aptis Task 4)を2-3人の採点者で採点→3レベル (B1, B2, C)
- **Sequential & 言語分析:** テストの回答の書き起こし & (コーディングの信頼性: 1日のワークショップと信頼性チェック二回の後)分析:
 - インタラクションのMovesの記述統計 (例: 相手の意見への理解を示す、これから反対意見を述べることを示唆する、1つめのポイントを使って反対意見を述べる)
 - 3レベルの特徴を質的に検証
- **フィードバックアンケートの記述統計 & インタビューのテーマ分析**

B1 sample (C18) Task A- groupwork

Well, Jan.	これから反対の意見を述べるのを示唆する
I have understand your points .	相手の意見全体への理解を述べる
You mean that (..) teamwork skills could help us (..) work together when we have a presentation. And because (..) we are all friends, we could work together and (..) help each other (..) during our work.	相手の意見1への理解を述べる
And it means that it could be (..) real life. Maybe we could learn more from others	相手の意見2への理解を述べる
But in my opinion ,	これから反対の意見を述べるのを示唆する
group presentations only work if all team members work hard.	1つ目のポイントを使って反対意見を述べる
because when we have different points and, (er) (..) and discuss with, (..) with it (..) it will (..) spend many time to (..) solve the problem	1つ目のポイントについて例を挙げたり詳しく述べる
and probably unfair (..) if all team members receive the same score.	2つ目のポイントを使って反対意見を述べる
For example, one person work.	2つ目のポイントについて例を挙げたり詳しく述べる
My last point is that it's probably unfair to everyone if all team members receive the same score.	最後のポイントをもう一度述べる
For example, someone just (..) do the simple work, but (..) he, (..) but he receive the same score with someone who work harder. And, (..), or who (..) do all, most work. (..)	最後のポイントについて例を挙げたり詳しく述べる

結果: IC interactional moves



フィードバック・アンケートからの抜粋

	IC tasks
タスク内で誰とコミュニケーションをとっているかが明確だった。	93.8%
会話の相手の内容を理解するのに、ビデオが(とても)役立った。	93.8%
タスクにあるような会話の相手を説得させる場面は日常良く起きる。	89.6%

結論

- ローテクなsemi-direct speaking tasks でも、「やりとりの能力の要素」をきちんと狭く定義して選べば、その能力を引き出すタスクを作れる（例: [acknowledging an interlocutor's view, clarifying/exemplifying a point in order to resolve a breakdown in communication](#)）
- 受験者は、誰と話しているかを明確に認識できると感じた。また、言語分析においても、受験者の発話が特定の相手に向けて、特定の観点を伝えようとしていることが検証された。

Innovation in English Language Education in Japan: Reforming University Entrance Examinations



日本での英語教育の改革：大学入試のリフォームに向けてステークホルダーにテストを熟知してもらい、サポートニーズを知る

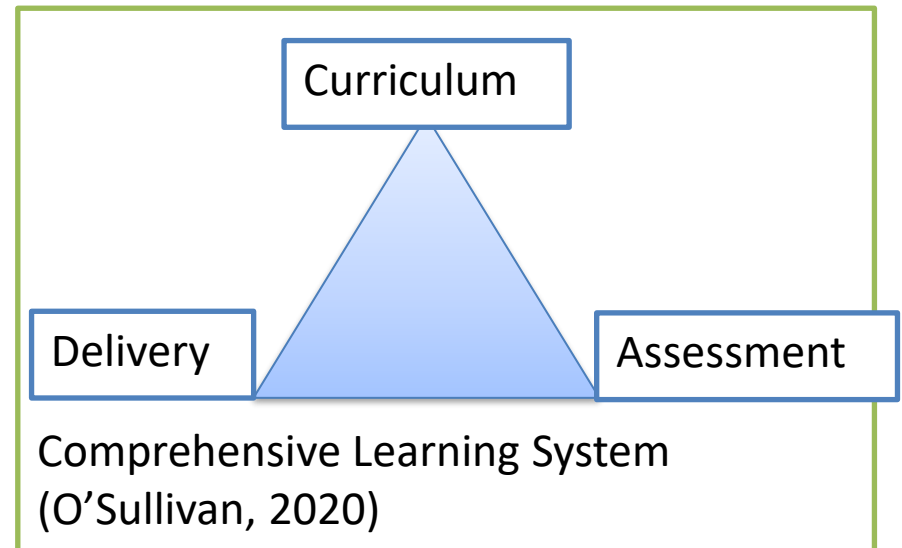
48

Nakatsuhara, F., Inoue, C., Nitta, R., Aoyama, T. (in press). *Innovation in English Language Education in Japan: Reforming University Entrance Examinations*. Cambridge English Research Notes. Forthcoming online publication.

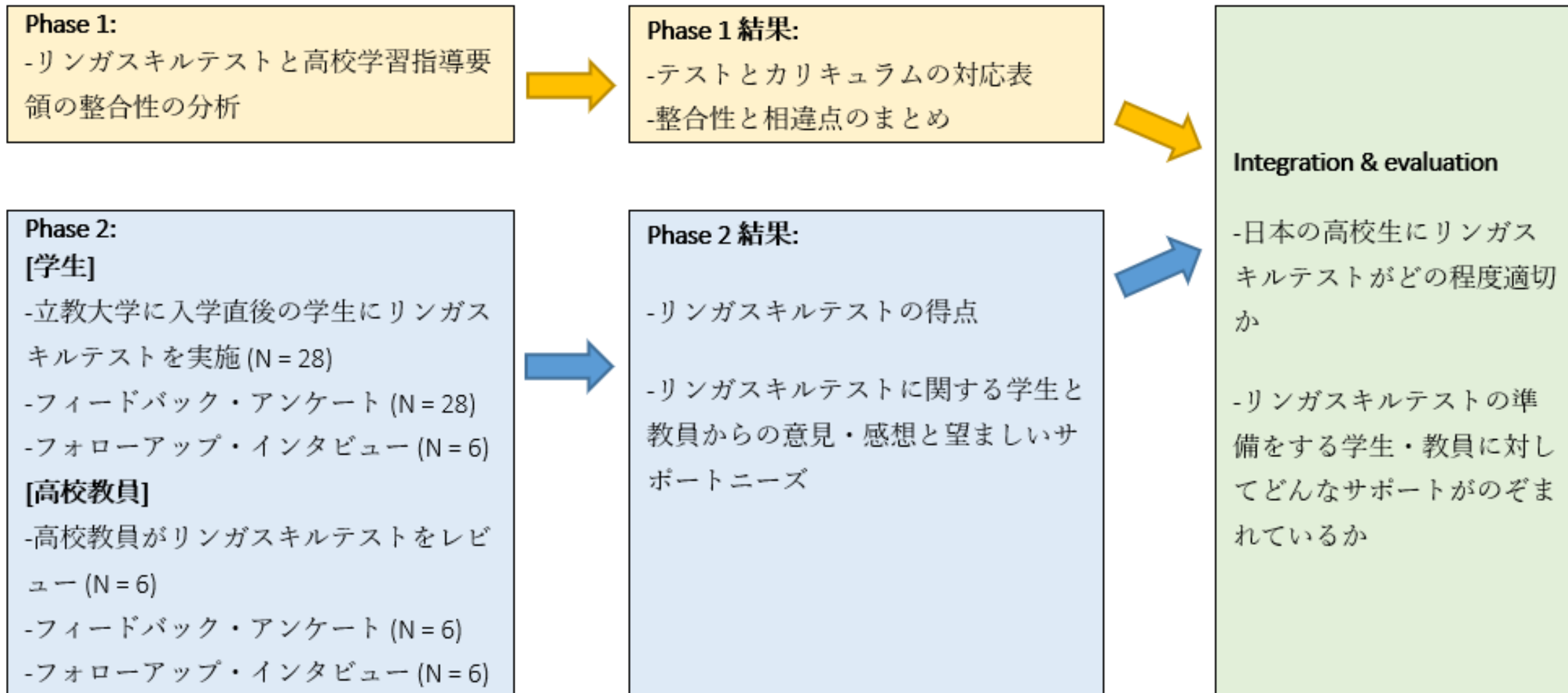
研究の背景

大学入試における英語4技能試験の導入の見送り・断念。大学個別対応への方針転換。

- 「大規模な試験改革が教育システムに良い影響を与えるには、教員へのトレーニング・教材提供・広報活動がしっかり行われることが必須」(Green, 2016, p.135)
- 「試験改革を進めるには全てのステークホルダーへのコミュニケーションを構築するモデル、それを遂行する計画が必須」(O'Sullivan, 2020)
- 「『言語試験リテラシー』という言葉は、リテラシーがある識者が他のステークホルダーに対し、どんなテストが『良いテスト』なのか教えてあげる、というような欠如を前提としたモデルを想像してしまう。それは、全てのステークホルダーのグループが協力しようという動きに逆効果。」(Baker, 2020)
- 「日本での英語民間試験導入で最も影響を受けるであろう『学生』と『教員』の声を体系的に聞き、教員へのトレーニング、学生・教員への教材提供を行う上で、何が必要かを他者が決めるのではなく、彼らが何が本当に必要なのかのニーズを探るべき」(Nakatsuhara, 2020)



リサーチのデザイン



- Q1. [学生] 以前練習したことがあるか・[教員] 高校3年生は経験があるか
 Q2. [学生]問題形式で求められていることは、すぐに分かったか・[教員]高校3年生は分かると思うか
 Q3. [学生]どのくらい難しかったか・[教員]高校3年生は難しいと感じるか

タスク1：質問を聞いて短い回答をする。

タスク2：表示された文を音読する。 ←学習指導要領との相違点

タスク3：提示されたトピックについて一分間話す。

タスク4：提示された図や表について一分間話す。 ←経験が比較的少ない ←難しい (68%学生、67%教員)

タスク5：提示されたトピックについて考え、聞こえてきた質問に答える。 ←難しい (75%学生と100%教員)

Instructions

Listen Think Speak 54 seconds

Your English-speaking friend wants to buy a cycling jacket and has asked for your advice.

This table shows product reviews for two cycling jackets.

Look at the table and then leave a message for your friend, recommending which cycling jacket to buy.

	Be Seen Jacket	Stay Dry Jacket
Waterproof fabric	★★★★★	★★★★★
Colour	★★★★★	★★★★★
Comfort	★★★★★	★★★★★
Price	\$75	\$110
Overall rating	7/10	9/10

Instructions

Listen Think Speak 29 seconds

A journalist is writing a magazine article about people's attitudes to technology and the internet. He wants to find out your opinion about technology and the internet.

He will ask you questions about:

- useful technology in daily life
- having the newest technology
- social media
- online advertising
- depending on technology

望ましいサポート・教材（欲しい・どちらでも良い・必要ない）	学生	高校教員
過去問題集のような練習問題冊子（紙媒体）とリスニングの音源及び解答と解説	92.9%	83.3%
過去問題集のようなウェブ上での練習問題（コンピューター使用）及び解答と解説	75.0%	100.0%
過去問題集のようなウェブ上での練習問題（スマートフォン・タブレット使用）及び解答と解説	75.0%	83.3%
複数のスピーキング回答例とその解説	82.1%	83.3%
スピーキング練習用アプリ	71.4%	100.0%
スピーキングの添削（フィードバック）サービス（例：自分のスピーチを録音してコメントをもらうオンラインでのスピーキング指導）	85.7%	66.7%
オンデマンド配信のリンガスキル対策動画	50.0%	50.0%
リンガスキルテストに特化した自宅で行うオンラインの英語コース（双方向型）	21.4%	33.3%
リンガスキルテストに特化した会場（学校・塾など）で行う英語コース（対面）	28.6%	16.7%
自宅で行う模擬テスト	50.0%	83.3%
テスト会場で行う模擬テスト	42.9%	16.7%
その他、リンガスキル受験対策として欲しいサポート・教材がありましたら記述してください。		

本日の話のまとめ

様々なスピーキングテストのタスク

- やりとり、一方的な発話、機械的な再生
- 引き出すべき言語機能(情報を与える、やりとりをする、やりとりを管理する)

スピーキングテストは何を測っているのか？

- 認知的な側面、社会的な側面
- テストの目的によって適するコンストラクトは異なる

スピーキングテストの開発

- テスト開発・妥当性検証のための「認知的社会的枠組み」
- テストの仕様、テストが施行される前に多角的な妥当性検証が必要

スピーキング能力向上につながるスピーキングテスト

- テストの波及効果
- コンピューターデリバリーのテストの活用
- カリキュラム、授業内容、テストの一貫性が大事
- 全てのステークホルダーが歩み寄って、テストの内容と重要性をシェアする必要性

ありがとうございました。

fumiyo.nakatsuhara@beds.ac.uk

