



# 文学能力モデルに基づいた評価の有用性の検討： 相関を用いた文学読解定期試験結果の ベイズ推計による分析

西原 貴之（広島大学）

ntakayk@hiroshima-u.ac.jp

# 研究の目的

現在統一化が進められつつある文学能力モデル（The Innsbruck model of literary competence：IMLC）は、文学教材を使った英語教育を实践した授業の期末試験において、学習者のパフォーマンス評価及びその結果の使用（学習者の到達度の確認やフィードバック）に有用に機能するかどうかを検証するためのエビデンスを得ること

# 研究の目的

具体的には…

1. IMLCの下位要素（下位能力）区分は適切か
2. IMLCに基づいた文学読解パフォーマンス評価のスコアは、文学読解と関連しつつも強い相関は期待されない別のパフォーマンス指標と并別のなものとなっているか

文学教材を使った英語教育を実践した授業の期末試験のスコアに対して、ベイズ統計学（簡易版）の手法を使い、

- (a) IMLCの下位要素のスコア間の相関分析
  - (b) IMLCの下位要素のスコアと「授業の平常点」の相関分析
  - (c) IMLCの文学独自の要素と「TOEICのリーディング・セクション」との相関分析
- を行い、IMLCが、期末試験（及び授業）において学習者のパフォーマンスの評価及びその結果の使用（学習者の到達度の確認やフィードバック）に有用に機能するかどうか、検討する。

# 研究の目的

西原（2013a）との違い：西原（2013a）と同一のデータに対して、新しい分析手法を用いて、新しい理論的枠組みの中で考察し直す研究

1. 近年、従来の頻度主義的な統計手法とは別の手法として注目されているベイズ統計学の手法を用いて（ただし、本発表ではその簡易版の手法を使用）データの再解釈を行うこと
2. 教室内での評価研究（classroom-based assessment：CBA）及び、IMLCという新しい枠組みの中に位置づけた研究であること

※西原(2013a)：(1) 西原（2013b）（またはNishihara（2015））で提案した文学能力モデルにおける文学読解独自の要素の特性の理解を深めること、(2) 英語文学教材を使用した英語授業の特性の理解を深めること、を目的とした研究

# 本発表の流れ

1. CBAにおける構成概念の重要性
2. 従来の文学能力モデルの問題点と現在統一化されつつあるモデル（IMLC）
3. 本研究の意義とねらい
4. 調査方法
5. 分析1：各設問タイプのスコアの相関分析
6. 分析2：各設問タイプのスコアと学生の平常点の相関分析
7. 分析3：各設問タイプの成績とTOEICのリーディング・セクションのスコアの相関分析
8. 結論
9. 今後の課題

# CBAにおける構成概念の重要性

- これまで言語テスト研究は大規模な標準テストを対象にしてきた。しかし、近年教室内での評価を対象としたclassroom-based assessment (CBA) の研究が徐々に進められつつある (e. g., Bachman & Damböck, 2017; Black & William, 2018; Chapelle & Voss, 2014; Fox, Abdulhamid, & Turner, 2021; Giraldo, 2020; Kane, 2021; Kane & Wools, 2020; Turner, 2012) 。
- CBAは、定期試験はもちろんのこと、授業内での様々な形での学習者のパフォーマンス評価を含む (※本研究はCBAの中でも期末試験に絞って議論を行う)。これらは、学習者の到達度を確認したり、学習者の得意なスキルと不得意なスキルを診断したり、フィードバックを与えるなど、様々な用途に役立てられている (Kane & Wools, 2020) 。学習者のパフォーマンス評価を通して学習をさらに促進させることはCBAの重要な働きの1つと考えられている (Fox, Abdulhamid, & Turner, 2021; Kane, 2006; Kane & Wools, 2020; Turner, 2012) 。

# CBAにおける構成概念の重要性

- 学習者の到達度を確認し、さらに有益なフィードバックを与えるためには、彼らのパフォーマンスを適切に評価することが不可欠である。
- このことを行うためには、構成概念の定義が不可欠である。William (2010) は、構成概念の定義がしっかりとされていない場合は、テスト結果から妥当な解釈を行うことが困難になると述べている。
- テストにおける構成概念の重要性は大規模な標準テスト研究において広く認識されてきたが (e. g., Chapelle, 2012; Chapelle & Lee, 2021; Messick, 1989; William, 2010) 、CBAにおいてもその重要性が指摘されている (Giraldo, 2020) 。

# CBAにおける構成概念の重要性

- Kane (2006) は、教師は教室で学習者のパフォーマンスや到達度を解釈するためには、“the conceptual frameworks” (Kane, 2006, p. 48) が必要と指摘している (Kane & Wools, 2020)。
  - [t]he conceptual frameworks provide a basis for evaluating student performance and for anticipating various factors that might influence student learning and performance (Kane, 2006, p. 48)
  - 文学教材を使用した英語授業の期末試験においては、学習者は文学読解のパフォーマンスの評価を行い、その結果を使用してその授業での到達度を確認すると同時に、学習者の英語文学読解能力を向上させるためのフィードバックを与えることが求められる (Lee & Goldman, 2015)。
- このことを有益に行うためには、その構成概念またはthe conceptual frameworkとして文学能力モデルの定義が重要となる (e. g., Lee & Goldman, 2015)。

# 従来の文学能力モデルの問題点と現在統一化されつつあるモデル（IMLC）

応用言語学、言語学、文学理論、文学テスト研究で様々な文学能力モデルが提示されてきた（e. g., Bachman, 1990; Culler, 1975; Grinder & Elgin, 1973; Henning, 1992; Spiro, 1991）。しかし、これらのモデルには以下の問題点があった。

## 文学能力モデルの問題点（Nishihara, 2022）

1. 文学能力の記述が断片的である
2. 文学能力の育成を上級外国語学習者に限ってしまった
3. 一般的な言語能力や読解能力と文学読解独自の要素（genre-specific traits）の双方をバランスよく取り入れることができていない
4. 文学能力に含めるべき下位要素に関して統一的な見解が得られなかった

→実際、文学能力モデルは学習者の文学読解パフォーマンスの評価（そして指導）への活用はほとんどなされなかった。結果として、Hall（2015）やLee（2011）が指摘するように、文学教材を使った授業における学習者のパフォーマンス評価の研究が大きく遅れる事態を引き起こした。また、CBAでGiraldo（2020）が、文学教材を用いたテストのデザインを提案しているが、文学能力モデルに基づいたものとはなっておらず、取り上げられている要素も断片的である。

# 従来の文学能力モデルの問題点と現在統一化されつつあるモデル（IMLC）

文学能力に関しては、2000年代に入り久しく研究の進展が停滞していたが、近年、上記の問題点を改善しつつ、モデルを統一化しようとする動きがある。

The Innsbruck model of literary competence (IMLC) (Atler & Ratheiser, 2019)

- (1) reading competence and general linguistic competence in English (processing texts)
- (2) emphatic competence: “the readers’ ability to personally relate to the characters in a text, their actions and reactions, emotions, thoughts, motives; their ability to connect with a text on an emotional level such as in ‘How would I behave in a similar situation?’ or ‘I think I would feel miserable under these circumstances.’” (pp. 380–381)
- (3) aesthetic and stylistic competence: “the ability to experience and appreciate a literary text and the literary principles of construction, i.e. the stylistics, it follows.” (p. 381)
- (4) cultural and discursive competence: “the learners’ ability to identify and work with the specific cultural freight and framing of a text and the discourses these pertain to.” (p. 380)
- (5) interpretive competence: “the ability to infer meaning from a work of literature” (p. 382)

# 従来の文学能力モデルの問題点と現在統一化されつつあるモデル（IMLC）

## IMLCの特徴

- 5つの下位要素はお互いに関連し合っている
- (1) が (2) ~ (5) の下位要素の土台と考えられている (Paran, 2006; Zyngier, Fialho, & do Prado Rios, 2007)
- (1) の上に(2) ~ (5) の下位要素がバランスよく伸長することで、文学能力が発達すると考えられている
- このモデルは、一般的な言語能力や読解力を含めた上で文学能力を多面的 (Paran, 2006) にとらえており、Nishihara (2022) が指摘した1点目と3点目の問題点の改善が見られる
- 従来の文学能力モデルでは、文学作品は英語力自体がかなり高くないと読むことができないという前提があったのに対して (e. g., Henning, 1992)、このモデルでは初級レベルの学習者から文学能力を育成していくことが可能であるとしており、実際の指導例も示している (他にも、Chesnokova & van Peer (2016) や Scott & Huntington (2007) など、入門期の英語学習者であっても、目標言語で文学作品を読む力を育成することが可能であることを示す研究が報告されている) (Nishihara (2022) が指摘する3点目の問題点に対しての改善)

# 従来の文学能力モデルの問題点と現在統一化されつつあるモデル（IMLC）

Paran, Spöttl, Ratheiser & Eberharter (2021) は、今後の文学テスト研究をIMLCに基づいて議論しており、モデルに含めるべき下位要素について見解の統一が進められつつある。Atler and Ratheiser (2019) よりも少し前に、Nishihara (2015) が類似した文学能力モデルを提案しており、IMLCは文学能力の構成要素を統一する動きの中の1つの到達点として評価できよう。なお、Nishihara (2015) は、文学教材を使った英語授業の到達度テストの中で以下の4要素を評価する必要があるとしている。

## Nishihara (2015)のモデル

1. literal meaning of the text.
2. interpretation of the text.
3. effects of distinctive language use in the text
4. learners' personal meaning and affective responses to the text

# 従来の文学能力モデルの問題点と現在統一化されつつあるモデル（IMLC）

## Nishihara (2015)のモデルの特徴

- IMLC同様に最初の要素が他の要素の土台と考えている
- 文学読解独自の要素である2つ目から4つ目の要素もAtler & Ratheiser (2019)と共通である
- Nishihara (2015)のモデルにはIMLCで言うところのcultural and discursive competenceに相当する要素が欠けている。これは、このモデルが一般英語教育に特化したモデルであるのに対して、Atler and Ratheiser (2019) は外国語または第2言語における専門的な英語文学教育も視野に入れたモデルであることに起因している (Paran, Spöttli, Ratheiser, & Eberharter, 2021)。実際、Nishihara (2015) は、文学読解において文化的な側面の重要性は認識しているものの、これらの知識はどちらかということ文学の専門的な教育で取り上げられるべきものであるとして、上記のリストに含めなかった (p. 119)。

# 本研究の意義とねらい

- Chappelle (2012) やMislevy (2009) は、言語能力を例として挙げながら、構成概念が心理学的に正確であるかどうかということよりも、スコアの解釈に役立つものであるかどうかという点が重要であるとしている。有用なモデルは、学習者のパフォーマンスの評価を元にして学習者の授業における到達度を記述したり、学習者にさらなる学習を導くフィードバックを与えることを可能にする (Kane & Woolls, 2020)
- IMLCは、CBAにおいて、学習者の英語文学読解パフォーマンスをさらに向上させる上で役立つかどうか、その反証データも含めて証拠を集めていかなければならない。
- 西原 (2013a) のデータを再分析することを通して、IMLCが学習者の英語文学読解パフォーマンス評価に利用可能であることを今後検討するための1つのエビデンスを提示する。本発表では、以後IMLCの用語に準じてNishihara (2015) の下位要素を指すこととする。
- 上述の通り、Nishihara (2015) 及びこれに基づいた西原 (2013a) では、cultural and discursive competenceに関するデータを収集していない。また、後述するようにemphatic competence に関して、そのパフォーマンス評価に失敗している。本発表ではIMLCのうち残りの3要素の観点から、このモデルの期末試験における有用性を検討する。

# 本研究の意義とねらい

## 西原（2013a）の研究結果

自身のモデル内に含まれる4つの下位要素（IMLCで言うところのcultural and discursive competenceを除いた4要素）に関する設問に文法に関する設問を加えて、合計5タイプの設問から構成される文学テストを作成し、そのスコアの相関を分析した。結果として、字義的内容理解の設問と文法の設問の間でのみ強い相関（ $r=.54$ ）が確認され、それ以外の設問間では大きな相関は確認されなかった（字義的内容理解の設問とinterpretive competenceの設問の間で $r=.32$ の相関が見られたのが最高であった）。そして、以下の点を指摘した。

1. 文学読解の中心作業が字義的内容理解であったとしても、それだけで文学テストを作成することは問題がある
2. 英語文学を読む際に重要となるその他の側面に関する能力についても別個に問う必要（Lee & Goldman, 2015）がある（もちろん、文学読解をinterpretive competence、aesthetic and stylistic competence、emphatic competenceのみで考えることは否定している）
3. 文学能力を様々な下位能力の集合体と見なし、これらの下位能力すべてを期末試験で扱う必要がある（Paran, 2006）

# 本研究の意義とねらい

文学教材を使った英語教育のCBAの一環として実施する期末試験においてIMLCは（期末試験をIMLCに基づいて構成することは）学習者の英語文学読解パフォーマンス評価及びその結果の使用（到達度の把握やフィードバック）に有用に機能するかどうか、ということを通して検討する。

1. IMLCの下位要素区分は適切か
2. IMLCに基づいた文学読解パフォーマンス評価のスコアは、文学読解と関連しつつも強い相関は期待されない別のパフォーマンス指標と并別的なものとなっているか

→上記の2点を、西原(2013a)のデータをベイズ統計学（簡易版）の手法を使って再分析することを通して検証

→上記2点目については、IMLCの下位要素のスコアと「授業の平常点」、IMLCの下位要素のスコアと「TOEICのリーディング・セクション」、の相関分析を通して検証

→IMLCのうち、西原(2013a)で扱わなかったcultural and discursive competence、うまく測定することができなかったemphatic competenceを除いた下位要素に基づいて検証

# 調査方法

## データを収集した授業

- 県立広島大学国際文化学科2年対象の選択科目「英語VIc」
- 同時帯に「英語VIa」と「英語VIb」という科目があり、それぞれ別の教員が担当した。その内容は各教員に任されており、本研究の発表者が担当した「英語VIc」では、英語小説教材の読解練習を行うこととした
- このことはシラバスおよび授業のイントロダクションで学生に周知してあるため、学習者は授業の内容を知った上で本科目を受講している。したがって、なんらかの形で英語小説の読解力を向上させることに関心を持っている学習者集団である

# 調査方法

## 調査参加者

- 県立広島大学国際文化学科2年対象の選択科目「英語VIc」を履修した37名（2年生33名、再履修生4名）に対する期末試験の1部
- 参加者の英語力はCEFRで言うところのA2からB2レベルと考えられる。実際、19名が同時期に実施されたTOEIC-IPを受験している。その結果は、582.1点を平均（標準偏差112.1）として、最高得点は835点、最低点は400点であった
- 受講者には英語関連の内容を専攻したいという学生とそうでない学生（日本語を含む東アジア系の言語、社会学など）が混在している。しかし、受講者は全員が英語文学を読むことに関心がある点で共通している。
- 一般的な英語の授業の中で英語文学小説の読解方法や読解方略について体系的な指導を受けた学習者はいなかった

# 調査方法

## 分析に使用するデータ

- (a) 本科目の期末試験のスコア
- (b) 本科目の平常点
- (c) TOEIC-IPのリーディング・セクションのスコア

→(a) と (b) のデータについては、調査参加者全員のデータを使用

→(c) については、授業期間中にTOEIC-IPを受験した19名分のデータを使用。したがって、このスコアとの相関を分析する際には、期末試験のスコアもTOEIC-IPを受験した19名分のもののみを使用。

※これら3種類のデータの研究目的での使用については、すべての調査参加者から書面にて許可を得た

# 調査方法

## 調査材料 : Hayase & Egashira (2010)

- 各章で異なる作品が取り上げられており、学習者は様々なアメリカ文学の作品に触れることができる（複数の章にまたがって扱われている作品もある）
- 各作品が部分的に取り上げられており、その部分に基づいて字義的内容理解や文法、interpretive competenceに関する設問が設けてある

# 調査方法

## 授業実践内容

- 毎回補助プリントを作成し、テキストに設けられている設問で問われていない箇所についても理解をはかった
- 教科書の設問が英語で作成されていたため、補助プリントの設問も英語で作成した。設問の理解で学習者がつまづかないように、極力シンプルな英語で作問した。
- IMLCに基づき、補助プリントには、字義的な内容理解に関する設問に加えて、interpretive competence、aesthetic and stylistic competence、emphatic competenceに関する問いも含めた
- このプリントは授業時間に配布し、学習者は授業時間内にテキストの設問および補助プリントの設問に解答した
- 解答した後、学習者はペアで設問の回答について意見交換を行い、その後で全員でその答えを討議した。また、内容理解に加えて適宜語彙や文法の確認も行った。

※一般英語の授業であるため、テキストで問題文として出題されている箇所が理解できているかどうかという点を重視し、cultural and discursive competenceについては扱わなかった。

# 調査方法

## 期末試験の記述

- 設問の文言や形式は授業時と同一ではないものの、一度授業で学習したテキストを使用したテストである。よって、学習者にとってはテストに含まれる英文は再読となる。今回のテストは授業の到達度テストであり、現時点での彼らの英語文学読解能力を測定した熟達度テストではない。
- 試験は90分間で実施し、100点満点とした。問題の構成は語彙や文法に関する一般的な問題が20点、文学作品の読解問題が80点であった。
- 文学作品の読解問題は、大問を3つ出題した。テキストは、Charlotte Gilmanの “The Yellow Wallpaper”、Kate Chopin の *The Awakening*、Edgar Allan Poe の “The Black Cat” とし、それぞれの配点を 38点、22点、20点とした。
- 設問の理解で学習者がつまづかないように、極力シンプルな英語で作問した。
- 解答言語は彼らの母語である日本語を指定した。
- 今回のテスト結果はこの授業を担当した教師1人のみによるローカルな使用が意図されたものであった (Kane, 2006; Kane & Wools, 2020) ため、すべての解答は授業実践者である発表者一人が採点を行った。

※本研究では、Charlotte Gilmanの “The Yellow Wallpaper” を扱った箇所の設問のスコアをデータとして使用する。字義的内容理解を中心としつつ、interpretive competence、aesthetic and stylistic competence、emphatic competenceについての設問を組み入れた。また、これらに加えて文法に関する設問も入れている（本発表では扱わない他の大問も同様の設問構成とした）。

# 調査方法

## 設問タイプ of 具体例

- 文法問題 (grammar (G)) : 英文中で使用されている文法項目や語彙、熟語 (授業で扱った項目) の理解を確認するための設問

Why, I wouldn' t have a child of mine, an impressionable little thing,  
( ) in such a room for worlds.

Q. Choose the best choice for the parenthesis.

(i) lives

(ii) live

(iii) to live

(iv) lived

- 字義的な内容理解の設問 (Reading competence and general linguistic competence in English (RC & GL)) : 登場人物の行為の理由を問うもの、ある表現が指す内容を尋ねるもの、など多岐にわたるが、このカテゴリーの設問はすべて英文中に明示的に書いてある事柄について問うものとなっている。文内の内容理解、複数の文にまたがる内容理解の両方が含まれるが、本発表では一括して扱う。以下は、ある表現が指す内容を尋ねる設問である。

And dear John gathered me up in his arms, and just carried me upstairs and laid me on the bed, and sat by me and read to me till it tired my head.

Q. To what does 'it' at underlined position refer?

# 調査方法

## 設問タイプの具体例

- interpretive competence (I)に関する設問：本文中に明示されていない事柄について問う設問とした。設問自体の答えに加えてその答えを支持するような理由を記述することを課した（設問としては2問に分けた）

And it is like a woman stooping down and creeping about behind that pattern.

Q. (i) To what does “a woman” refer? (ii) Also provide the reason why you think so.

- aesthetic and stylistic competence (AS)に関する設問：特定の言語的特徴が本文中でどのような効果をもたらしているのかを答えさせるものとした。

Q. What effects do the ill-formed paragraphs and sentences (for example, their extreme shortness and excessive use of “and” ) achieve?

# 調査方法

## 設問タイプの具体例

- emphatic competenceに関する設問：登場人物に対する情意的な反応を自由に答えさせる問いとした。自分の考えそのものとなぜそのような考えを持ったのか理由を記述することを課した（設問としては2問に分けた）
- Q. If you were in the narrator' s shoes with the same husband, (i) would you feel happy? (ii) Also provide the reason why you feel so.

※残念ながら、今回の試験では、emphatic competenceに関する設問については、学習者のパフォーマンスの優劣を評価することに失敗した。そこで本発表では、以後このタイプの設問結果については扱わず、他の4タイプの設問結果についてのみ分析を行うこととする。出題された文章、具体的な設問、設問のカテゴリー分けのさらなる詳細については、Nishihara (2015) を参照されたい。

# 調査方法

## 分析で用いるその他のデータ

- 本科目の平常点：参加者全員分のデータを使用  
英語IVcの授業成績全体から、期末試験のスコアを除いたもの

### ※英語IVcの成績

1. 出席10%
2. リーディング課題提出10%（TOEICの設問形式に準じた課題の提出点とその課題のスコア）
3. Eラーニングの進捗状況10%（TOEICに準じた内容で、課題にどの程度の時間取り組んだかを評価）
4. 小テスト20%（前回の授業で扱った作品の語彙と文法の復習）
5. 期末試験50%

※「本科目の平常点」は、学生の勤勉さと一般的な英語力の指標として機能すると考えられる

→平常点との関係性を見ることで、文学読解独自の側面（interpretive competence、aesthetic and stylistic competence、emphatic competence）として測定した設問のスコアが単に学生の勤勉さや一般的な英語力の指標となっていないかどうかを確認

# 調査方法

## 分析で用いるその他のデータ

- TOEIC-IPのリーディング・セクションのスコア：授業期間中に受験した19名分のスコアを使用

→このスコアとの関係性を調べることで、文学読解独自の側面が一般的な英語力と併別されたものとなっているかどうかを確認

# 調査方法

## 分析と調査課題の関係

- 分析1：ベイズ統計学による各設問タイプ間の相関分析→研究課題1
- 分析2：ベイズ統計学による各設問タイプと平常点の相関分析→研究課題1 & 研究課題2
- 分析3：ベイズ統計学による各設問タイプとTOEIC-IPのリーディング・セクションのスコアの相関分析→研究課題1（データを19名分に絞った上で再確認） & 研究課題2

※研究課題1：文学能力モデルの下位要素区分は適切か

※研究課題2：文学能力モデルに基づいた文学読解パフォーマンス評価のスコアは、文学読解と関連しつつも強い相関は期待されない別のパフォーマンス指標と弁別的なものとなっているか

# 調査方法

## 分析方法

本発表におけるベイズ統計学による相関分析

- マルコフ連鎖モンテカルロ法（MCMC法）によって発生させた乱数に基づく分析（通常はRを使って実施される）ではなく、JASPを用いた簡便法による分析
- ベイズファクター $BF_{10}$ に基づいた分析
- 母相関係数 $\rho$ として、帰無仮説を「 $H_0: \rho=0$ 」、対立仮説を「 $H_1: \rho \neq 0$ 」とした。 $\rho$ の事前分布として、JASPのデフォルトであるベータ分布 $B(1, 1)$ を使用
- 「 $BF_{10}$ が1より大きいときは対立仮説が（相対的に）正しく、 $BF_{10}$ が1より小さいときは帰無仮説が（相対的に）正しいと判断」（清水・山本, 2020）される。Lee and Wagenmakers (2013) によると、 $BF_{10}$ は、1未満は証拠なし、1より大きく3未満は乏しい支持、3以上10未満は中程度の支持、10以上30未満は強く支持、30以上100未満はとても強く支持、100以上は非常に強く支持、と判断される。

※ベイズ統計学の基本的な考え方や従来の統計学との違い等については、時間の都合上本発表では省略。詳細は豊田（2016）など関連書籍を参照のこと。

# 分析I：各設問タイプのスコアの相関分析

## 記述統計

|      | 期末試験の各設問タイプ |         |     |     |
|------|-------------|---------|-----|-----|
|      | G           | RC & GL | I   | AS  |
| 満点   | 6           | 22      | 4   | 2   |
| 最小値  | 1           | 4       | 0   | 0   |
| 最大値  | 6           | 22      | 4   | 2   |
| 平均値  | 3.7         | 15.9    | 2.9 | 1.6 |
| 標準偏差 | 1.4         | 4.2     | 1.0 | 0.6 |

# 分析I：各設問タイプのスコアの相関分析

各設問タイプのスコアの相関係数

|         | G | RC & GL                  | I                         | AS                       |
|---------|---|--------------------------|---------------------------|--------------------------|
| G       |   | .54 (61.3)<br>[.25, .71] | .14 (0.3)<br>[-.19, .43]  | .20 (0.4)<br>[-.13, .48] |
| RC & GL |   |                          | .32 (1.3)<br>[-.003, .57] | .23 (0.5)<br>[-.10, .50] |
| I       |   |                          |                           | .11 (0.3)<br>[-.21, .41] |
| AS      |   |                          |                           |                          |

表中の ( ) 内はベイズファクター  $BF_{10}$ 、[ ]内は相関係数の95%確信区間

# 分析I：各設問タイプのスコアの相関分析

## 各設問タイプのスコアの相関係数

- 「文法」と「字義的内容理解」の間で $r = .54$ というとても強い正の相関 ( $BF_{10}=61.3$ ) が確認された
- これらの変数とinterpretive competence、aesthetic and stylistic competenceとの相関は小さかった（最も大きなものでも「字義的内容理解」とinterpretive competenceの間に $r = .32$  ( $BF_{10}=1.3$ ) の相関係数が確認されたのにとどまった)
- interpretive competenceとaesthetic and stylistic competenceの相関も小さかった ( $r = .11$ ,  $BF_{10}=0.3$ ) 。

→interpretive competenceとaesthetic and stylistic competenceは、文学能力の異なる側面を測定

→それらは「文法」や「字義的内容理解」とも異なる側面を扱っている。

→頻度主義的な相関分析に基づいた西原 (2013a) と同じ結果

→IMLCの下位要素の区分の適切さを支持

# 分析2：各設問タイプのスコアと学生の平常点の相関分析

## 平常点の記述統計

| 満点 | 最小値  | 最大値  | 平均値  | 標準偏差 |
|----|------|------|------|------|
| 50 | 19.1 | 48.9 | 34.7 | 8.3  |

## 各設問タイプのスコアと授業の平常点の相関係数

|     | G                        | RC & GL                   | I                        | AS                       |
|-----|--------------------------|---------------------------|--------------------------|--------------------------|
| 平常点 | .46 (11.2)<br>[.15, .67] | .60 (323.8)<br>[.32, .76] | .12 (0.3)<br>[-.21, .41] | .22 (0.5)<br>[-.11, .49] |

# 分析2：各設問タイプのスコアと学生の平常点の相関分析

## 各設問タイプのスコアと授業の平常点の相関係数

- 「文法」と「字義的内容理解」に比べて、interpretive competenceとaesthetic and stylistic competenceは値が低くなっていた。具体的には、ピアソンの積率相関係数が「文法」とは $r=.46$  ( $BF_{10}=11.2$ )、「字義的内容理解」とは $r=.60$  ( $BF_{10}=323.8$ )であるのに対して、interpretive competenceとは $r=.12$  ( $BF_{10}=0.3$ )、aesthetic and stylistic competenceとは $r=.22$  ( $BF_{10}=0.5$ )となっている。

→一般的な英語力が関係している「平常点」は一般的な英語力が関係しており、理論上は「文法」と「内容理解」との高い相関及びinterpretive competenceとaesthetic and stylistic competenceとの低い相関が予想されたが、今回の分析結果は、この予想を支持。

→interpretive competenceとaesthetic and stylistic competenceは文学読解独自の指標として機能したと言え、IMLCによる評価が単に学生の勤勉さや一般的な英語力の指標になっていないことを支持。

→相関分析結果から、通常の英語学習に関して動機づけが低い学習者や英語が苦手な学習者であっても文学読解では著しい能力を発揮する可能性がある

→逆に、授業への取り組みがまじめであったり、一般に英語が得意な学習者の中には、文学読解独自の側面を苦手とする者がいることも考えられる。このような学習者にとっては、文学能力モデルに基づいた授業及び評価、フィードバックを通して、読解力の幅を広げることが期待できよう。

→加えて、文学教材を使用したテスト（または授業）では、設問（発問）が「文法」と「字義的内容理解」だけにならないように注意しなければならない

# 分析3：各設問タイプの成績とTOEICのリーディング・セクションのスコアの相関分析

TOEIC-IPを受験した19名の学習者の記述統計

|      | 期末試験の各設問タイプ |         |     |     | TOEIC-R |
|------|-------------|---------|-----|-----|---------|
|      | G           | RC & GL | I   | AS  |         |
| 満点   | 6           | 22      | 4   | 2   | 445     |
| 最小値  | 1           | 9       | 0   | 0   | 185     |
| 最大値  | 6           | 22      | 4   | 2   | 385     |
| 平均値  | 3.8         | 17.4    | 2.9 | 1.7 | 269.5   |
| 標準偏差 | 1.3         | 3.4     | 1.2 | 0.6 | 53.7    |

# 分析3：各設問タイプの成績とTOEICのリーディング・セクションのスコアの相関分析

各設問タイプのスコアとTOEICのリーディング・セクションの相関係数

|         | G | RC & GL                  | I                        | AS                        | G                         |
|---------|---|--------------------------|--------------------------|---------------------------|---------------------------|
| G       |   | .64 (16.1)<br>[.22, .83] | .13 (0.3)<br>[-.32, .51] | .13 (0.3)<br>[-.32, .51]  | .62 (11.5)<br>[.20, .82]  |
| RC & GL |   |                          | .25 (0.5)<br>[-.21, .60] | .09 (0.3)<br>[-.36, .49]  | .37 (0.9)<br>[-.11, .67]  |
| I       |   |                          |                          | -.02 (0.3)<br>[-.44, .41] | -.06 (0.3)<br>[-.47, .37] |
| AS      |   |                          |                          |                           | -.08 (0.3)<br>[-.48, .36] |
| TOEIC-R |   |                          |                          |                           |                           |

# 分析3：各設問タイプの成績とTOEICのリーディング・セクションのスコアの相関分析

## 各設問タイプのスコアとTOEICのリーディング・セクションの相関係数

- データの数が減ったことにより、多少の数値の変動はあるものの、各設問タイプ間の相関関係に関しては、分析1と同様な結果となった（分析3で用いたデータは分析1で用いたデータの一部であるため、このことは驚くことではないが）
- 「文法」と「字義的内容理解」に関しては、 $r = .64$ という強い正の相関（ $BF_{10}=16.1$ ）が確認されたのに対して、これらの変数とinterpretive competenceとaesthetic and stylistic competenceとの相関は小さく（最も大きなものでも「字義的内容理解」とinterpretive competenceの間に $r = .25$ （ $BF_{10}=0.5$ ）の相関係数が確認されたのにとどまった）、かつ、interpretive competenceとaesthetic and stylistic competenceの相関も $r = -.02$ （ $BF_{10}=0.3$ ）と小さかった。

→IMLCの下位要素の構成については再度支持

TOEIC-Rのスコアに関しては、分析2と同様に、理論上は「文法」と「字義的内容理解」との高い相関及びinterpretive competenceとaesthetic and stylistic competenceとの低い相関が予想

- 「文法」とは $r = .62$ と大きな正の相関（ $BF_{10}=11.5$ ）が確認され、この予想を支持
- しかしながら、「字義的内容理解」とは、 $r = .37$ （ $BF_{10}=0.9$ ）という結果であり、予想に反して低い相関となった

→TOEICで読む文章と文学教材ではジャンルが大きく異なっており、同じ字義的内容理解であってもその内実は大きく異なっていることがその原因かもしれない。このことについては今後さらなる調査が必要。

- TOEICはinterpretive competenceとaesthetic and stylistic competenceとはほぼ無相関（それぞれ  $r = -.06$ （ $BF_{10}=0.3$ ）と $r = -.08$ （ $BF_{10}=0.3$ ））。

→このことについては予想通りの結果であり、TOEICの代用にはなっていないことを支持。

# 結論

- 研究課題1：IMLCの下位要素区分は適切か  
→分析結果1、2、3の結果から支持
  - 研究課題2：IMLCに基づいた文学読解パフォーマンス評価のスコアは、文学読解と関連しつつも強い相関は期待されない別のパフォーマンス指標と并別のなものとなっているか  
→分析結果2、3の結果から支持
- IMLCはCBA（本発表では期末試験）において学習者の文学読解パフォーマンスの評価、及びその使用（授業到達度の確認や学習者へのフィードバックなど）に関して有用性を支持

# 今後の課題

- しかしながら、今回の調査だけでは、個々の教室を超えてIMLCが、文学読解パフォーマンスの評価及びその使用に有用に使用できるかどうか一般化することはできない。今後も様々な環境の中で異なる文学教材、異なる設問形式を多様な学習者に対して与える中で、このモデルの有用性の証拠を集めていくことが必要。
  - 本研究では扱うことができなかったemphatic competence及びcultural and discursive competenceの取り扱いについても今後検討していくことが必要。
  - 同時に学習者の文学読解をつぶさに観察し、IMLCが必要な要素を取り入れたものとなっているかどうかという点（構成概念が“underrepresentation”（Messick, 1989）の状態になっていないかどうかの確認）も継続的に検証し、モデルの充実をはかっていく作業も必要
- Probst (2004) は、学習者がどれほど自分と違う他人の意見に対して寛容でいられるかということは文学読解パフォーマンスと密接に関連していると述べているが、IMLCにはこの要素を反映できるような下位要素は含まれていない。
- Schmidt (1983) は、文学コミュニケーションには、作品の受容だけでなく、創作、出版社による媒介、受容後のポスト処理（レビューを書くなど）、といった要素も関わることを指摘している。このことは、IMLCのスコープを拡張する必要性を示唆していると言える。
- 学習者の文学読解パフォーマンスの発達段階を評価できるような仕組みづくりも求められる（Chapelle & Lee, 2021; Lee & Goldman, 2015）。CEFRにおけるcan-doディスクリプタ（Council of Europe, 2018, 2020）と文学能力モデルの統合などを行うことが必要（e. g., Adler & Ratheiser, 2019）。
  - 本発表では議論しなかったものの、設問形式（e. g., Carter & Long, 1990; Paran, 2010）が文学読解パフォーマンスの測定に与える影響についての研究も行われる必要がある。



ご清聴  
ありがとうございました。

本発表は本研究はJSPS科研費JP19K00790の助成を受けたものです。