

# 日本語学習者の Academic Speaking の 評価基準の構築と妥当性の検証

横山 千聖  
(2019年10月3日受理)

Assessing Second-Language Learners' Academic Speaking Skills in Japanese:  
Development of Valid Assessment Criteria

Chisato Yokoyama

**Abstract:** An assessment system for academic speaking in Japanese as a second language (L2) was developed. The L2 English academic speaking ability is tested using high-stage tests such as Test of English as a Foreign Language (TOEFL) and International English Language Testing System (IELTS). However, no such tests are available for L2 Japanese learners. To fill this gap, the present study aims to develop assessment criteria for argumentative speech by nonnative Japanese speakers. In so doing, we developed a rating scale using the TOEFL Independent Speaking Rubrics and the IELTS Speaking Assessment Criteria, and analyzed the fit between the descriptions in the rating scale and the linguistic characteristics of the data, such as fluency, complexity, and accuracy. The results revealed that overall fluency was positively correlated with complexity and accuracy, but negatively correlated with fluency. Therefore, there is a trade-off in effects between fluency and accuracy and between fluency and complexity. Examinees who received high scores tended to be more accurate with a high degree of complexity, but their speech rate was slow. In this sense, the rating criteria were accurate for complexity and accuracy, but the fluency description may need to be modified.

Key words: Academic Speaking, Japanese Learner, Oral Opinion Statement, Rating

キーワード：アカデミックスピーキング，日本語学習者，意見陳述，評価

## 1. はじめに

留学生30万人計画の実施などから、日本国内の留学生の人数は増加の一途を辿っている。それらの留学生は大学などの高等教育機関や、また将来的にビジネスシーンでプレゼンテーションをしたり、議論を行ったりすることが求められるようになる。そのためには日本語のスピーキング能力の向上が必要となってくる。

本論文は、課程博士候補論文を構成する論文の一部として、以下の審査委員により審査を受けた。

審査委員：畑佐由紀子（主任指導教員）、白川博之、  
永田良太、仁科陽江

近年、日本語では Academic Writing については田中・阿部（2014）などにより Academic Writing の特徴や構造、また日本語学習者に対する評価の研究が進められてきているが、日本語の Academic Speaking については、基礎研究や評価基準についても、研究が進められてきていないのが現状である。さらに、英語圏では Academic Speaking の評価基準に関する研究が進められてきているが、日本語にはそのような評価基準は存在しない。日本語学習者の Academic Speaking を評価するため、本稿では日本語での Academic Speaking の評価基準を作成すること、また、その妥当性を検証することを目的とする。

## 2. 先行研究

### 2.1 Academic Speaking の定義

Academic Speaking の定義は様々な研究者が行なってきた。本稿では、まず初期から外国語教育について研究が進められてきた英語での Academic Speaking の定義を述べ、次に日本語に関する研究で議論されている Academic Speaking の定義についてまとめる。そして、これらの定義の比較検討を行った上で、本稿での Academic Speaking の定義を行う。

#### 2.1.1 英語における Academic Speaking の定義

Jordan (1997) は、Academic Speaking は大学等での「学習」が中心の概念であると主張している。また、Using English for Academic Purpose For Students in Higher Education (2019) によると、Academic Speaking とは日常生活のスピーキングとは異なり、明白なものであり構造的であると述べられている。Academic Speaking の使用場面には、De Chazal (2014) が、講義、プレゼンテーション、ゼミ、チューターとの面談、指導教員との面談、グループプロジェクト、議論の中での他の学生達とのやりとりを挙げている。このように、Academic Speaking とは、高等教育機関などでの学習時の、論理的な発話のことを指すと言える。

#### 2.1.2 日本語における Academic Speaking の定義

英語に対して日本語では、三宅 (2003) が、Academic Speaking に関する詳細については言及してはいないものの、日本留学試験の報告書をもとに、アカデミック・ジャパニーズとは何かについて考察しており、学術研究と日常生活に関する日本語の2種類に分類されると述べている。また、ボイクマン (2013) は山本 (2004) や門倉 (2006) の考え方を踏まえ、Academic Speaking は論理的思考とコミュニケーション力の2種類に分類することができると主張している。また、門倉 (2003) も、日本語ではアカデミック・ジャパニーズの定義は曖昧で、生活のためと学術のための2つの分野に分類されていると述べている。さらに、Academic English はアカデミック・ジャパニーズの定義とは異なり、クリティカルシンキング (批判的思考) やロジックなどが重視されていると指摘している。

#### 2.1.3 Academic Speaking の定義のまとめ

日本語の Academic Speaking は、大学の日常生活でのコミュニケーション能力など、キャンパス・ジャパニーズも含めて考えられているのが現状であるが、キャンパス・ジャパニーズは道聞きや情報を得るために質問するなど、キャンパス以外の日常生活で行わ

れるトピックとも重複している。Academic Speaking とは、発表や議論など高等教育機関で行われる学業としての意味として定義されるべきであると考ええる。そこで、本稿では、Jordan (1997) や門倉 (2003) の述べている Academic English の概念に基づき、日常生活での使用を目的とするキャンパス・ジャパニーズは含めない、大学の学業などで求められる論証や説明のことであると定義する。

### 2.2 母語話者の Academic Speaking の特徴

本節では、英語と日本語それぞれの母語話者の Academic Speaking について、これまで明らかにされてきた特徴について述べる。そして、それぞれの特徴から明らかになったことについてまとめる。

#### 2.2.1 英語母語話者の Academic Speaking の特徴

Ranta (2009) は、ELFA という英語母語話者のアカデミックな場面での発話を収集したコーパスと MICASE という英語学習者の同様の発話を収集したコーパスを対象に、統語構造、特に条件節について分析を行った。その結果、母語話者は仮定形の would もしくは would have という構造を Academic な場面で頻用していることが分かった。また、Mauranen (2009) は、Ranta と同様の2種類のコーパスデータの中のモノログを用いてレトリックと接続表現について分析を行った。その結果、議論の談話の中では母語話者は再提示や言い換えを多用していたことが明らかになった。さらに、Mauranen (2012) では上記2つの論文と同様のコーパスを用いて、語彙頻度等について分析を行った。その結果、母語話者は難易度の高い語彙を使用する傾向があった。また、母語話者が複雑な談話構造を用いるのに対し、学習者は単純な談話構造や文法を使う傾向が観察された。また、母語話者はメタ表現を用い、言い換えや強調などを行なっていることが分かった。

#### 2.2.2 日本語母語話者の Academic Speaking の特徴

日本語母語話者の Academic Speaking の特徴を探る研究は少ないものの、外国語学習者の口頭産出能力を測る ACTFL-OPI の評価基準を借用し、分析を行なった研究がいくつか存在する。

荻原 (2009) は日本語母語話者48名を対象に、日本語学習者の日本語会話能力を測定する ACTFL-OPI と同じ手法を用いて日本語母語話者にインタビューを実施した。そして、インタビュー会話での意見表明時における談話の論理展開を分析した。インタビューの内容は「オリンピックにおけるドーピングについて」、「外国人と比較した日本人の働きぶりについて」という内容であった。分析方法については、意見談話から中心となる主張を抽出して、「中心文」として設定

し、その中心文が談話中の何文目に該当するかを算出して、中心文が談話の初めに位置する「冒頭型」であるか最後に位置する「末尾型」であるかについて分析した。その結果、「冒頭型」は39.6%、また「末尾型」は47.9%であるということが明らかになった。

また、長坂・木田・木山（2005）は、ACTFL-OPI データの中の意見陳述談話の構成要素について分析を行った。対象者は日本語母語話者13名であり、意見の内容、構成について分析した結果、内容については「意見成分」と「意見補強成分」があり、さらにその構成については3つの種類があることが分かった。まず、個人的主張から一般的な概念へと話を展開していく「ズームアウト」、次に、主張を行いその後論点の詳細を行うという「ズームイン」、そして様々な観点から検討を行う「複数の視点の提示」に分類されたことを報告している。

### 2.2.3 母語話者の Academic Speaking の特徴のまとめ

母語話者の Academic Speaking の分析結果を考察すると、母語話者の研究は談話構造や談話の構成要素に関する分析が主に行われてきていることが分かる。また、それに伴った接続表現やメタ表現についての分析も多くなされてきたことが明らかになった。しかし、特に日本語母語話者の先行研究では、討論やインタビューなど他者の介入やインターアクションが多くあるため、他者の発話に依存する結果であることに留意しなければならない。また、母語話者の中でもいわゆる評価が高い Academic Speaking とそうでないものがある。2.2.1や2.2.2の対象者はその点において、どのような対象者であったかは定かではなく、母語の論証や説明において明確な特徴を示しているかは不明である。

## 2.3 第二言語学習者の Academic Speaking の能力評価基準の分析

本節では、これまで作成されてきた第二言語学習者の Academic Speaking の能力を測定するスタンダードやテスト評価について概観し、それらの特徴をまとめる。

### 2.3.1 IELTS の概要と目的

International English Language Testing System（以下、IELTS と呼ぶ）はイギリスで考案された、英語を母語としない人を対象とする英語の使用能力を測定するための試験である。海外移住、または留学を希望する人が英語力を測定されるための基準となる試験でもある（渡慶次, 2014）。

#### 2.3.1.1 IELTS の Academic Speaking の評価基準

IELTSが発行している IELTS Speaking Assessment Criteria (2018) によると、IELTS の Academic Speaking

の評価基準は9段階の評価に分かれており、理解可能でやりとりが成立する初級レベルの1や2、単文や複数の文を成立させられるが発音や文法などに誤りを生む3や4のレベル、談話を成立させることができ、レベルが上がっていくが、一貫性などにまだ問題や癖が生じるレベルの6や7、発音等に僅かな母語の影響が出るが一貫性や流暢性があり文法が正確な、上級、超級レベルにあたる8、9といったレベル分けが行われている。評価基準は、流暢性と一貫性を測る Fluency and coherence、統語構造や語彙を測る Lexical resource、文法のバリエーションや正確さに関する Grammatical range and accuracy、発音に関する Pronunciation という4つで構成されている。

### 2.3.2 TOEFL の概要と目的

Test of English as a Foreign Language(以下、TOEFL とする)は、大学レベルの英語を使用および理解する能力を測定するための試験である。英語を母語としない人の英語の使用能力を測定するためのアメリカ発祥の試験である。PC を用いて行う試験である（渡慶次, 2014）。

#### 2.3.2.1 TOEFL の Academic Speaking の評価基準

TOEFLが公開している TOEFL iBT Speaking Section Scoring Guide (2014) によると、Speaking の試験は2種類あり、Speaking 能力のみを測定する Independent Speaking Test と、他の Listening, Reading などの技能と複合した Integrated Speaking Test を実施している。評価基準は、質問の要求に答えられているか、一貫性があるかどうかに関する概要を測定する General description、発音やリズムやスピードなどの話し方に関する Delivery、文法や語彙、接続表現などの言語運用に関する Language Use、タスク内容との合致度、話の展開を基準とする Topic development の4つの指標をもとに設定されている。

評価のレベルは4段階あり、レベル1は課題に限定的にしか答えられず文法や語彙が限られており、発音も聞き手の努力を要し、課題に基本的なことにのみ答えられるという特徴を持つ。レベル2は、課題に対応可能で基本的な文構造を使う。しかし、トピックの発展性がなく、一貫性がなく分かりにくさが生じるのが特徴的である。レベル3は、話は首尾一貫しており、音声や文法も癖や誤りはあるものの理解に大きな影響がない。しかし、時折アイデア間の関係がすぐにははっきりしないことがある。レベル4は、質問にはっきりと答えられ、ベース良く話し、難易度の高い語彙や文法を使いこなし、根拠やアイデア間の繋がりが明確であるという特徴を持つ。以上のような4つのレベルに分けられている。

TOEFL と IELTS はいずれもアカデミックな論証をする能力を測るものであり、実際に専門的な話をするを前提とはしていない。IELTS の基準が流暢さ、統語構造、文法、発音など、言語運用に重点を置くのに対し、TOEFL は、言語運用や流暢さの他にトピックの一貫性や話の展開といった論証の要素を評価基準に入れている点で異なる。

### 2.3.3 CEFR の概要と目的

Council of Europe(2001)によると、Common European Framework of Reference for Languages: Learning, teaching, assessment (以下、CEFR と言う) は1990年代初頭に欧州評議会がヨーロッパ共通の言語学習や評価について打ち出した共通枠のことである。言語が話せることだけでなく、異文化でのコミュニケーション能力に主眼を置いている。さらに、教育のカリキュラム作成や教科書、評価の基準にもなっている。

#### 2.3.3.1 CEFR の評価基準

CEFR のレベル分けに関しては、A1, A2が初級学習者、B1, B2が中級学習者、C1, C2が上級の熟達した学習者と分類されている。このように6レベルに分かれており、Cレベルは高い能力が求められる。これらの評価基準は「～ができる」というコミュニケーション能力が基準となっている (Council of Europe, 2001; 文部科学省, 2018)。

#### 2.3.4 ACTFL-OPI の概要と特徴

ACTFL-OPI は、発話の流暢性の向上と教育と学習の向上を目指している (牧野・鎌田・山内・齊藤・荻原・伊藤・池崎・中島, 2001)。

##### 2.3.4.1 ACTFL-OPI の評価基準

ACTFL-OPI は合計で11レベルに設定されており、特に初級レベルに関して詳細に分類が行われている。上級レベルでは Academic Speaking の論証などが求められている。評価表は、正確さ、流暢さ、質問内容など評価の基準があるが (American Council on the Teaching of Foreign Languages, 2012)、概要にとどまっており、詳細な記述はされていない。

### 2.3.5 日本語における Speaking の指標

日本語では、Academic Speaking に特化した評価基準はまだ開発されておらず、一般的な Speaking の評価基準も少ない。Academic Speaking については、国際交流基金が考案した JF スタンダードが僅かに言及している。

#### 2.3.5.1 JF スタンダードの概要と目的

JF スタンダードは、ヨーロッパで考案された CEFR を基にして作成されたスタンダードのことであり、基本的にコースデザインや授業設計、評価など、教師が教育目的として活用することを主な目的として

いる (国際交流基金, 2018)。

#### 2.3.5.2 JF スタンダードの評価基準

CEFR を基にしているため、CEFR と同様、「何をどのくらい遂行できるか」という行動志向型である (国際交流基金, 2018)。

### 2.3.6 評価からの Academic Speaking 能力のまとめ

TOEFL, IELTS は大学留学を主眼に置いているため、Academic Speaking 能力を測ることを目的としている。両者の相違は、レベルがどのくらい細分化されているか、また言語使用を中心として評価基準を設定するか、発話内容や論理展開を含めるかによる違いがある。一方、CEFR を基準とした JF スタンダードは教育のための使用目的の要素が大きく、さらにコミュニケーションと機能を重視していることが分かる。また、ACTFL-OPI はタスク達成型という側面が強く、正確さなどにも言及しているが概要にとどまっていると言えるであろう。

2.1や2.2の母語話者との Academic Speaking の特徴と合わせて鑑みると、TOEFL や IELTS には、構造や根拠の複雑さなど Academic Speaking に必要な評価基準が含まれている。

### 2.3.7 第二言語学習者の Academic Speaking の特徴

2.3.1から2.3.6までは学習者の評価について概観したが、ここでは第二言語学習者の Academic Speaking の特徴について質的に分析を行った研究について概観する。

Vercellotti & Packer (2016) は英語学習者のモノログでの文構造の複雑さを習熟度別に分析した。分析方法は、AS-Unit を用いて句構造を分析した。AS-Unit とは、Foster, Tonkyn, Wigglesworth (2000) によると発話の単位のことで、Writing の T-Unit の文の単位から発展した単位のことであり、AS-Unit は、発話者の発話の主節 (従属節が付随することもある) または独立副節 (従属節が付随することもある) のことである。副節とは、発話の文脈により省略されていた箇所が復活することで節になる句のことを指し、これらの指標を用いて分析を行なった。その結果、非定型節 (不定詞、分詞、動名詞) は中級レベルの学習者は用いることができていなかったことが明らかになった。

また、荻原・齊藤・増田・米田・伊藤 (2001) は、ACTFL-OPI の超級、上級と日本語母語話者を対象に「話題の種類」と「発話内領域 (身近な出来事、一般的出来事、抽象的事象)」を項目として、対応、発話構成、談話構成、文法について分析した。その結果、発話構成や抽象表現、接続表現などで上級学習者間に差が見られたという。

鈴木 (2006) は、ACTFL-OPI の発話をコーパスにした KY コーパスのデータ (超級15名, 上級14名) を対象に上級超級学習者の論述談話について発話構造の特徴を分析した。その結果, 超級は複数の理由付けや裏付けの正当性を論じることができ, さらに言い換えや繰り返しを用いて, 総合的に談話を関連づけることができるということが明らかになった。

#### 2.4 先行研究のまとめと残された課題

以上のような先行研究の知見から, 多言語での Speaking 能力の評価基準が開発されていることが明らかになったが, その評価基準は教育的側面が強いものが多い。さらに, TOEFL や IELTS のような学術的な言語運用を想定した指標は英語では確立しているが, 日本語での Academic Speaking の能力を測定する指標は確立されていないことが明らかになった。また, 第二言語学習者の Academic Speaking の一部の特徴は明らかにされてきているが, 特定の文法項目や特定の談話構造など, 分析項目に偏りがあり, Academic Speaking の流暢さや文法, 語彙や正確さなど全体像を明らかにしているとは言い難い。そこで, 本稿では以下の課題を設定する。

【課題】日本語での Academic Speaking の評価基準を作成し, その妥当性を検証する

### 3. 日本語の Academic Speaking の評価基準の作成と妥当性の検証

#### 3.1 目的

日本語の Academic Speaking の評価基準を作成すること, また, その妥当性を検証することを目的とする。

#### 3.2 対象者

本稿では, 科学研究費助成事業「学部・留学生教育研究のためのアカデミック・ジャパニーズスピーキングコーパスの開発」(課題番号: 17K02847 基盤研究 (C)) で収集したデータの一部を用いた。具体的には, 英語を母語とする学習者と中国語を母語とする学習者のデータから, Excel のランダム係数を用いて無作為にそれぞれ15名ずつ, 計30名を選び, これらの対象者のデータを評価資料として採用した。

性別は男性13名, 女性17名であった。年齢幅は19歳から35歳で平均年齢は25.2歳であった。英語母語話者は4つの大学の学習者から協力を得た。レベル測定については, A 大学ではプレースメントテストで判定した。結果は10名が全5レベル中3, 4レベルと中上級レベルであった。B, C, D 大学では Spot90 というレベル判定テストを実施, その結果74~84点の学習者が

5名であった。中国語母語話者に関しては, 2つの大学の協力を得た。学習者の詳細は交換留学生2名, 研究生6名, 修士生4名, 博士生3名であり, SPOT90 というレベル判定テストを実施したところ75点以下の学習者が7名, 75~85点の学習者が8名であった。

#### 3.3 対象データ

科学研究費助成事業「学部・留学生教育研究のためのアカデミック・ジャパニーズスピーキングコーパスの開発」(課題番号: 17K02847 基盤研究 (C)) の音声データは, TOEFL と IELTS の Speaking の課題内容を参考に作成され, 課題に対する回答を収集したものである。課題には2種類の論証課題と1種類の描写課題があるが, 本稿では, 論証課題の発話を評価データとして用いた。論証課題は, 「A: テクノロジーは私たちの生活を複雑にしたいと思いますか, 思いませんか。どうしてそう思いますか。できるだけ詳しく, 例を使って, 自分の意見を説明してください。」といったような二者択一問題10題と, 「B: 文化が違う国に旅行に行った時は, その国の習慣に従った方が良いという意見があります。あなたはこの意見にどの程度賛成しますか。」というような解答に幅を持たせた問題10題で構成されていた。

対象者には対面形式もしくは PC 上でランダムに抽出された課題 A, 課題 B それぞれについて2問ずつ話させたため, 対象者のデータは4つの談話からなる。よって分析対象としたデータの総数は4談話×30名=120データである。

なお, 課題について話してもらう前に「この調査では大学教員に話したり, 大学で公衆の面前で話したりするときのような話し方で話してください。」と指示を行い, Academic Speaking であるという前提を説明した。

#### 3.4 日本語の Academic Speaking の評価基準作成

日本語での評価基準作成のために, 23で検討したように教育的側面やタスク達成型の側面を示している CEFR, JF スタンダードや ACTFL-OPI ではなく, 言語面や論理展開を重視している IELTS, TOEFL に着目した。IELTS の流暢さ, 統語構造や語彙, 文法, 発音などの言語運用を, また, TOEFL Speaking Rubric (Independent) のトピックの一貫性や話の展開を評価基準の要素に含めた。TOEFL Speaking Rubric (Independent) と IELTS の評価基準を参考に, 「概要」, 「話し方」, 「語彙」, 「文法」, 「話の展開」という項目を設けた。「話し方」, 「語彙」, 「文法」は IELTS の記述を, 「概要」と「話の展開」は TOEFL の記述を和訳し, 「話し方」に関して細かい記載をしている英語に特化した記載を改訂した。また,

5, 6段階の「概要」と「話の展開」に関する記述がTOEFLにはないため、IELTSの他のレベルの記述も参照しながら8, 9スコアに合わせて独自に評価基準を作成した。

「概要」とは、課題に答えられているか、妥当な根拠を示しているかという点に関する評価観点であり、「話し方」とは発話のアクセントやイントネーション、言い淀みやペース等の観点を指す。また、「語彙」とは語彙の使用に関する評価観点、「文法」は文法に関する評価の観点であり、「話の展開」はアイデア間の繋がり、一貫性やその繋げ方についての評価観点のことである。また、TOEFL Speaking Rubric (Independent) の4, 3, 2, 1スコアの基準と本稿の基準を合わせ、その基準に更に上位の者を評価できるようにIELTSの9, 8スコアを参照し、6, 5スコアを追加して6段階評価とした。

### 3.5 対象データの評価者

評価者は日本語母語話者2名であった。評価者の詳細は、大学院博士課程後期在学中の大学院生1名(筆者)と日本語教育を専門とする教員1名である。

### 3.6 評価方法

評価者がそれぞれ個別に評価を行なった。各学習者の4つのデータについて音声データを聞き、各項目の「概要」、「話し方」、「語彙」、「文法」、「話の展開」について6段階評価を行い、さらにその総合得点を計算し、それぞれ一致率を測定した。

### 3.7 評価結果の一致率

評価基準をもとに評価者2名が別々に評価を行ったところ、「概要」の一致率が0.87, また「話し方」は0.83, 「語彙」は0.95, 「文法」は0.89, 「話の展開」については0.87, また総合得点は0.93であった。評価基準の一致率は一般に9割以上を満たすことが必要であるとされているが、その基準を満たさなかった観点がかった。

そこで、評価がずれた点について確認し、評価基準の説明について再考した後、再評価を行い突き合わせた。その結果、「概要」の一致率は0.95, また「話し方」は0.91, 「語彙」は0.97, 「文法」は0.94, 「話の展開」については0.94, 総合得点は1.00と信頼性のある一致率を得た。

表1は、全学習者、英語を母語とする学習者、中国語と母語とする学習者の中央値、四分位範囲、最大値、最小値を示したものである。四分位範囲とは、データ数を四分等分した際にデータを小さい順から数えたときの75%の値から25%の値を引いて算出した範囲のことを指す。本稿のスピーキング評価は評価基準をもとに評価を行なっているため、統計的には正規分布を想定しない順序尺度である。このような場合、ばらつきの

表1 全体と各母語の学習者の中央値、四分位範囲、最大値、最小値の一覧

項目	総合評価	評価項目				
		概要	話し方	語彙	文法	話の展開
中央値 (全体)	15.88	3.25	3.13	3.13	3.00	3.13
中央値 (英語)	13.75	3.00	2.75	2.75	2.75	2.75
中央値 (中国語)	19.25	3.75	3.75	4.00	4.00	3.75
四分位範囲 (全体)	5.88	0.94	1.25	1.38	1.25	1.25
四分位範囲 (英語)	7.75	1.50	1.75	1.25	1.38	1.50
四分位範囲 (中国語)	6.38	1.13	1.38	1.13	1.50	1.13
最大値 (全体)	24.75	5.00	5.00	5.00	5.00	4.75
最大値 (英語)	21.25	4.50	5.00	3.75	4.50	4.00
最大値 (中国語)	24.75	5.00	5.00	5.00	5.00	4.75
最小値 (全体)	6.00	1.25	1.00	1.25	1.00	1.00
最小値 (英語)	6.00	1.25	1.00	1.25	1.00	1.00
最小値 (中国語)	13.50	3.00	2.75	2.50	2.25	2.50

指標には標準偏差ではなく、四分位範囲が用いられる。

表1から分かるように、英語を母語とする日本語学習者は評価項目の「話し方」の評価の差が大きいという傾向がある。一方、中国語を母語とする日本語学習者の評価の最大値や最小値は高く、高評価の傾向にあるが、評価項目の「文法」や「話の展開」では学習者間で差があるということが言える。

両者の評価のばらつきを見ると、英語を母語とする日本語学習者の最大値と最小値の差は15.25で、四分位範囲は7.75であった。中国語を母語とする日本語学習者は最大値と最小値の差は11.25であった。四分位範囲は6.38と英語学習者と比較すると値は小さかった。

#### 3.7.1 日本語の Academic Speaking の評価基準の妥当性の検証

本稿で作成した評価基準は特定の母語の学習者を想定したものではない。よって、評価基準の妥当性の検証は英語を母語とする日本語学習者と中国語を母語とする日本語学習者のデータを統合した合計のデータで検証を行う。今回作成した評価基準の妥当性を検証するために、文字化したデータをもとに流暢さ(語彙的流暢さ、発話の流暢さ)と複雑さ(語彙的複雑さ、発話の複雑さ)、正確さ(文法的正確さ)の5つの指標を用いて妥当性を検証した。評価基準の妥当性の検証

の方法にこの5つの指標を用いたのは、発話の評価を行うために「流暢さ・複雑さ・正確さ」という大きく3つの指標を用いるのが主流であり (Ellis, 2009), また近年では「語彙の複雑さ」や「発話の複雑さ」などに関してもスピーキングの評価を行っている研究も多く存在する (Michel, 2017)。そのため、本稿では「流暢さ」, 「複雑さ」の指標に「発話」と「語彙」の下位分類を設けることにより, さらに詳細な評価が行えると考えたため, この5つの指標を用いた。本稿で作成した評価基準によって評価した結果が, 従来の5つの発話の評価指標と高い相関を示すことで, 本稿で作成した評価基準の妥当性が立証されることになる。

また, 今回は特に評価基準の「話し方」と「語彙」, 「文法」の言語面に焦点を当てて妥当性を検証した。本稿の目的は「本稿で作成した評価基準の妥当性を検証する」というものだが, 妥当性の検証について, 従来スピーキングに関して研究が進められている「流暢さ・複雑さ・正確さ」という言語面の指標を用いることにより, より精度の高い検証が行えると考えた。その指標を用いて測定できるのは「話し方」と「語彙」, 「文法」であるため, 本稿では優先的に取り上げて検証した。本稿で言語面をまず検証することで, より段階的にミクロからマクロの分析ができると考えたため, 本稿では「言語面」に着目した。

### 3.7.2 流暢さ

流暢さには語彙的流暢さと発話の流暢さを指標に用いた。語彙的流暢さは, 発話の総語数を発話時間 (秒数) で割って算出した。語数は, 今回は名詞, 動詞などを含む内容語を対象とし, 助詞などの機能語は含めないこととした。また, 発話の流暢さは発話の AS-Unit の総数を発話時間 (秒数) で割って算出した。AS-Unit とは, 2.3.7 で触れたように, 発話の単位のことである (Foster, Tonkyn, Wigglesworth, 2000)。

### 3.7.3 複雑さ

複雑さには, 語彙的複雑さと発話の複雑さを指標に用いた。語彙的複雑さは内容語について, 複数回出現した同一の語を1回とカウントする異なり語数を総語数で割って算出した。また, 発話の複雑さは発話の総節数を AS-Unit の総数で割り, 算出した。

### 3.7.4 正確さ

正確さについては, 文法や語彙に関するもので, Error Free AS-Unit, つまり正しい文法や語彙を使用している AS-Unit を AS-Unit の総数で割って, 算出した。

## 3.8 結果

調査の各項目の結果をもとに, 評価表に基づいた評価結果 (話し方, 語彙, 文法) に対して, 流暢さ (発

表2 結果一覧

	得点	概要	話し方	語彙	文法	展開	流暢 (語彙)	流暢 (発話)	複雑 (語彙)	複雑 (発話)	正確さ
得点	1										
概要	0.96	1									
話し方	0.94	0.88	1								
語彙	0.97	0.91	0.86	1							
文法	0.96	0.87	0.89	0.95	1						
展開	0.97	0.97	0.88	0.92	0.90	1					
流暢 (語彙)	0.37	0.35	0.55	0.24	0.33	0.31	1				
流暢 (発話)	-0.49	-0.45	-0.41	-0.49	-0.48	-0.52	0.18	1			
複雑 (語彙)	0.11	0.05	0.11	0.09	0.15	0.14	-0.01	-0.15	1		
複雑 (発話)	0.40	0.37	0.46	0.33	0.35	0.41	0.43	-0.59	0.23	1	
正確さ	0.48	0.49	0.29	0.57	0.50	0.45	-0.18	-0.27	-0.33	-0.08	1

話, 語彙), 複雑さ (発話, 語彙), 正確さの5つの指標で相関関数を用いて分析を行った。表2に結果の一覧を示す。

### 3.8.1 流暢さ

流暢さについては, 語彙的流暢さと発話の流暢さの2つの分析を行った。

#### 3.8.1.1 語彙的流暢さ

語彙的流暢さと評価結果は0.37で低い相関関係を示した。しかし, 本稿で作成した評価基準の詳細である「話し方」との相関は0.55で中程度の正の相関があった。このことは, 発話の中の語彙数 (ここでは内容語) が多いほど, 話し方の評価が高くなることを意味する。

#### 3.8.1.2 発話の流暢さ

発話の流暢さと評価結果には -0.49で中程度の負の相関があった。特に, 発話の流暢さと評価基準の「話し方」には -0.41で負の相関が, 「語彙」は -0.49, 「文法」は -0.48と中程度の負の相関があることが分かった。このことに加え, 発話の流暢さと発話の複雑さには -0.59, 正確さには -0.27の中程度の負の相関や低い負の相関があることを含めて鑑みると, 流暢に話そうとすると意識すればするほど, 発話, 特にここでは節の複雑さや文法的正確さが下がったと言える。このような現象は, 複雑さが上がると流暢さが下がる「トレードオフ現象」であり, 同様の結果が出ている研究も存在する (Ferrari, 2012)。

しかし, 評価基準の構築という観点からすると, この結果は, 実際の流暢さは評価結果とは反対の傾向を示しているということを意味している。本来ならば本

稿で作成した評価基準に基づいた評価結果と指標（流暢さ）に正の相関があれば妥当性を示すことができる。だが、本稿の対象学習者である中上級学習者、特に中級学習者は、今まで学習してきたことを注意深く確認しながら話す傾向があるため、「流暢さ」が低くなり、「複雑さ」や「正確さ」が高くなった可能性がある。また、本稿の対象者の人数に限りがあったことも関連がある可能性もある。

### 3.8.2 複雑さ

複雑さについては、語彙的複雑さと発話の複雑さの2種類について分析を行った。

#### 3.8.2.1 語彙的複雑さ

語彙的複雑さと評価結果には0.11と低い正の相関、また評価基準である「話し方」とも0.11、「語彙」とは0.09、「文法」とは0.15など低い相関関係であることが明らかになった。このことから、語彙的な複雑さ、つまり語彙のパラエティは評価にあまり関係しないのではないかということが考えられる。

#### 3.8.2.2 発話の複雑さ

発話の複雑さについては、評価結果とは0.40と中程度の正の相関であった。また、評価基準である「話し方」とも0.46と中程度の正の相関があることが分かった。発話の複雑さは総節数をAS-Unitの総数で割り出して算出している。節の数が発話の複雑さに関係があり、また評価とも関連があることが推測される。

### 3.8.3 文法的正確さ

正確さについては、評価結果とは0.48と中程度の正の相関を、評価基準の「語彙」は0.57、「文法」は0.50と中程度の正の相関を示した。3.8.1.2でも示したように、正確さが下がると評価結果も下がるということを考慮すると、評価結果には語彙や文法の正確さが関連してくると言える。

## 4. まとめ

本稿では、日本語の Academic Speaking の評価の指標を作成し、その妥当性を評価項目の「話し方」と「語彙」、「文法」という言語面から検証した。その結果、評価には発話の流暢さよりも発話の複雑さ、また文法的正確さが関連してくることが明らかになった。また、発話の正確さと複雑さが、30名という数の少ないデータであっても、指標として有効であることが分かった。

その一方で、流暢さにおいては、流暢であるほど評価が低くなるという傾向が見られた。この原因として、本稿で扱ったデータ数が少なく、安定していないことが一因としてあげられる。また、本稿で対象とした学習者は中級学習者が多く、IELTS や TOEFL に比べ、

能力差の幅が狭いことも影響した可能性がある。中級者だけでは流暢さの差がさほど大きく出るとは限らず、むしろ学習者がどれだけ注意深く話そうとするかが影響する可能性がある。そのため、初級、上級、超級を含めた幅広い習熟度の学習者データを用い、流暢さについて再検討する必要がある。

さらに、発話の流暢さの指標には、本稿で扱ったものだけではなく、一分間のポーズの回数やいい直しの回数など様々なものが提案されているが、どの指標が有効かという点については研究間で異なり、日本語では検討されていない。よって、ほかの指標についても今後検討する必要がある。そして、複雑さと正確さの指標が習熟度の幅を広げても対応可能かどうかについても再検討する必要がある。

これに関して、今後は日本語母語話者を含めて、指標の検討をすべきである。先行研究でも述べたように、日本語母語話者の Academic Speaking の特徴や評価などは解明されていない。CEFR や IELTS などの超級の評価基準を参考にすることにより、日本語母語話者の Academic Speaking も評価、特徴を分析することにより、日本語学習者の Academic Speaking への分析も深まるであろう。

また、Academic Speaking の持つ特殊性が関連している可能性もあると考えられる。具体的には、ACTFL-OPI はタスク達成型で、タスクに応じた内容やコンテキスト、正確さ、発話のタイプに評価基準に関する重点を置いている。しかし Academic Speaking の評価基準は、そうでない Speaking の基準とは低頻度語や慣用句などのより難易度の高い語彙使用や文法使用が求められ、また論証を行う必要があるため、Academic Speaking ではない Speaking の評価より認知的負荷が高いと考えられる。本稿の結果で「発話の複雑さ」や「文法的正確さ」がより高い評価に繋がったのは、そのような要因があるからではないかと考えるため、今後も十分な考察が必要であると考えられる。

最後に、本稿で開発した評価基準には、タスクに対する発話の概要や話の展開などの項目があるが、今回の分析には含めていない。追調査の結果では、総合得点評価と高い相関関係を示すかどうかについては課題が残るため、この点にも留意が必要である。また、論証の構成についても分析し、評価基準にあるような論証構造の明確さや証明の論理性などについても考察を進めていく必要があると考える。

## 【謝辞】

本研究は、2017年度 JSPS 科研費17K02847（基盤研究（C）『学部・留学生教育研究のためのアカデミック・ジャパニーズスピーキングコーパスの開発』研究代表者：半沢千絵美）の助成を受けた。

## 【参考文献】

- 荻原稚佳子 (2009) 「意見述べにおける日本人の論理展開についての一考察」『明海日本語 =Meikai Japanese language journal』 14, 1-11.
- 荻原稚佳子・齋藤真理子・増田眞佐子・米田由喜代・伊藤とく美 (2001) 「上・超級日本語学習者における発話分析：発話内容領域との関わりから」『世界の日本語教育・日本語教育論集』 11, 83-102.
- 門倉正美 (2003) 「アカデミック・ジャパニーズとは何か」『日本留学試験とアカデミック・ジャパニーズ』 1, 123-132.
- 門倉正美 (2006) 「〈学びとコミュニケーション〉の日本語力アカデミック・ジャパニーズから発信」『アカデミック・ジャパニーズの挑戦』ひつじ書房, 3-20.
- 鈴木志のぶ (2006) 「日本語学習者によるアーギュメントの特徴：上級者・超級者間の差異」『Speech Communication Education』 19, 95-112.
- 田中真理・阿部新 (2014) 『Good Writing へのパスポート－読み手と構成を意識した日本語ライティング－』くろしお出版.
- 投野由紀夫 (2013) 『CAN-DO リスト作成・活用 英語到達度指標 CEFR-J ガイドブック』大修館書店.
- 渡慶次正則 (2014) 「TOEFL iBT におけるスピーキング能力とライティング能力の妥当性、信頼性、実現性の検証」『名城大学紀要』 19, 64-76.
- 寅丸真澄 (2006) 「日本語の討論の談話における「意見表明」の構造分析」『早稲田大学日本語教育研究』 9, 23-35.
- 長坂水晶・木田真理・木山登茂子 (2005) 「日本語母語話者の意見表明：OPI にもとづいたインタビューにおける意見述べの分析」『言語文化と日本語教育』 29, 96-99.
- 深川美帆 (2007) 「接続表現から見た上級日本語学習者の談話の特徴－日本語母語話者と比較して－」『言葉と文化』 8, 253-268.
- ボイクマン総子 (2013) 「初級前期からのアカデミック・ジャパニーズ教育－初級前期の口頭発表の実践－」『AJ ジャーナル』 5, 11-19.

- 牧野成一・鎌田修・山内博之・齋藤真理子・荻原稚佳子・伊藤とく美・池崎美代子・中島和子 (2001) 『ACTFL-OPI 入門－日本語学習者の「話す力」を客観的に測る－』アルク.
- 三宅和子 (2003) 「留学生・日本人大学生のアカデミック・ジャパニーズとは」『日本留学試験とアカデミック・ジャパニーズ』 1, 101-112.
- 山本富美子 (2004) 「アカデミック・ジャパニーズに求められる能力とは－論理的・分析的・批判的思考法と語彙知識をめぐって－」『移転記念シンポジウム－アカデミック・ジャパニーズを考える－報告書』 1-6.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- De Chazal, E. (2014). *English for academic Purposes: Oxford handbooks for language teachers*, (pp.243-245). Oxford, UK: Oxford University Press.
- Ellis, R. (2010). Second language acquisition, teacher education and language pedagogy. *Language Teaching*, 43(2), 182-201.
- Ferrari, S. (2012). A longitudinal study of complexity, accuracy and fluency variation in second language development. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency* (pp.277-298). Amsterdam/Philadelphia: John Benjamins.
- Figueras, N. (2012). The impact of the CEFR. *ELT Journal*, 66(4), 477-485.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21(3), 354-375.
- Jordan, R. R. (1997). *English for academic purposes: A guide and resource book for teachers*. Cambridge: Cambridge University Press.
- Mauranen, A. (2009). Spoken rhetoric: How do natives and non-natives face? In E. Suomela-Salmi & F. Dervin (Eds.), *Cross-linguistic and cross-cultural perspectives on academic discourse*, (pp.193-199) Amsterdam/Philadelphia: John Benjamins.
- Mauranen, A. (2012). *Exploring ELF: Academic English shaped by non-native speakers*. Cambridge: Cambridge University Press.
- Michel, M. (2017). Complexity, accuracy and fluency(CAF). In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language*

*acquisition* (pp.2-16) London: Routledge.

Ranta, E. (2009). Syntactic features in spoken ELF: learner language or spoken grammar? In A. Mauranen, & E. Ranta. (Eds.) *English as a lingua franca: Studies and findings*, (pp.84-106) Newcastle upon Tyne: Cambridge Scholars Publishing.

Vercellotti, M. L., & Packer, J. (2016). Shifting structural complexity: The production of clause types in speeches given by English for academic purposes students. *Journal of English for Academic Purposes*, 22, 179-190.

### 【参考 Web サイト】

国際交流基金 (2018) 「JF 日本語教育スタンダード」  
<<https://jfstandard.jp/summary/ja/render.do>>  
(最終閲覧日：2019年6月23日)

文部科学省 (2018) 「各資格・検定試験と CEFR との  
対照表」  
<[http://www.mext.go.jp/b\\_menu/houdou/30/03/\\_icsFiles/afeldfile/2019/01/15/1402610\\_1.pdf](http://www.mext.go.jp/b_menu/houdou/30/03/_icsFiles/afeldfile/2019/01/15/1402610_1.pdf)>  
(最終閲覧日：2019年6月23日)

American Council on the Teaching of Foreign Languages.(2012).ACTFL Proficiency Guidelines, Retrieved June 23, 2019, from

<<https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012/japanese/スピーキング>>

British Council, IDP & Cambridge Assessment English. (2018). IELTS Speaking Assessment Criteria, Retrieved June 23, 2019, from

<<https://www.ielts.org/-/media/pdfs/speaking-band-descriptors.ashx?la=en>>

Educational Testing Service. (2014) . TOEFL iBT Speaking Section Scoring Guide, Retrieved June 28, 2019, from

<[https://www.ets.org/s/toefl/pdf/toefl\\_speaking\\_rubrics.pdf](https://www.ets.org/s/toefl/pdf/toefl_speaking_rubrics.pdf)>

Using English for Academic Purpose For Students in Higher Education. (2019) .Speaking in Academic Contexts, Retrieved June 28, 2019, from

<<http://www.uefap.net/speaking/features/speaking-features-introduction>>