

# オンライン英語ディスカッション可視化の実践に基づく 学生の発話行動の量的分析

松岡 真由子<sup>1</sup> 大竹 翔子<sup>2</sup> 水本 武志<sup>3</sup> 森下 美和<sup>4</sup>

<sup>1</sup> 追手門学院大学基盤教育機構 〒567-8502 大阪府茨木市西安威 2 丁目 1-15

<sup>2,4</sup> 神戸学院大学グローバル・コミュニケーション学部 〒650-8586 神戸市中央区港島 1-1-3

<sup>3</sup> ハイラブル株式会社 〒170-0005 東京都豊島区南大塚二丁目 26 番 12 号 鈴音ビル 2 階 203 号室

E-mail: <sup>1</sup>m-matsuoka@otemon.ac.jp, <sup>2,4</sup>{sotake, miwa}@gc.kobegakuin.ac.jp, <sup>3</sup>t.mizumoto@hylable.com

**あらまし** 本稿では、オンラインの英語教育におけるグループディスカッションの定量化・可視化システムを用いた学生の行動の定量分析について報告する。グループディスカッションは効果的であるものの「良さ」の定義や客観的な評価が困難であった。そこで学生の発話行動を可視化するシステムを用いて授業実践を行った。学生のアンケートから経験を積むほど自己や他者の行動の即時フィードバックを求めていることがわかった。また、実践で得られたディスカッションの定量データから、ターンテイクと自己評価との関連性の高さが明らかになり、ターンテイクによって英語ディスカッションの自然さの分類ができる可能性が示唆された。

**キーワード** フィードバック, ディスカッション経験, ターンテイク, 自己評価

## Quantitative Analysis of Students' Online English Discussions: Data from A Web Conference Visualization System

Mayuko MATSUOKA<sup>1</sup> Shoko OTAKE<sup>2</sup> Takeshi MIZUMOTO<sup>3</sup> and Miwa MORISHITA<sup>4</sup>

<sup>1</sup> Institute of Liberal Arts, Otemon Gakuin University 2-1-15 Nishiai, Ibaraki-shi, Osaka, 567-8502 Japan

<sup>2,4</sup> Faculty of Global Communication, Kobe Gakuin University 1-1-3 Minato-jima, Chuo-ku, Kobe-shi, Hyogo, 650-8586 Japan

<sup>3</sup> Hylable Inc. 2-203, Suzuoto Bldg., 2-26-12, Minami Otsuka, Toshima-ku, Tokyo, 170-0005 Japan

E-mail: <sup>1</sup>m-matsuoka@otemon.ac.jp, <sup>2,4</sup>{sotake, miwa}@gc.kobegakuin.ac.jp, <sup>3</sup>t.mizumoto@hylable.com

**Abstract** Online group discussion can be an effective way to help students practice their communication skills. There is, however, difficulty in deciding how “good” a discussion is in an objective manner. In this study, we utilized Hylable, a web conference visualization system, in order to both qualitatively and quantitatively analyze student discussions. This system was implemented as part of a classroom exercise. From student surveys, we found that students more experienced with discussions tend to want immediate feedback for themselves and their peers. Furthermore, from the quantitative data, the amount of turn-taking was related to how much a student believed they contributed in a discussion. Possibilities of using turn-taking as a way of determining if a natural discussion took place or not will be discussed.

**Keywords** feedback, discussion experience, turn-taking, self-evaluation

### 1. はじめに

ディスカッションは、私達の日常生活において重要なコミュニケーションの方法の一つである。相手との言語（ここでは話し言葉）を介したやり取りは、日常的

な事柄に関するトピックや時事問題、解決策を要する問題など様々なテーマについて行なわれる。ディスカッションパターンは1対1でのやり取りから複数人でのやり取りに至るまで多岐に渡る。ゆえに、ディスカ

松岡真由子・大竹翔子・水本武志・森下美和, “オンライン英語ディスカッション可視化の実践に基づく学生の発話行動の量的分析”, 言語学習と教育言語学 2021年度版, pp. 67-76, 日本英語教育学会・日本教育言語学会合同編集委員会編集, 早稲田大学情報教育研究所発行, 2022年3月31日.  
Copyright © 2021-22 by Matsuoka, M., Otake, S., Mizumoto, T., & Morishita, M. All rights reserved.

セッションは一方通行で話すプレゼンテーションよりも多くのスキルを学ぶ重要な機会を与えるものである。

本稿で扱う英語でのディスカッションは、母語でのそれと比べて難しい。例えば Yanagi & Baker (2016) は、日本人の学生が海外に留学したときに一番参加が難しいのがグループディスカッションであると報告している[1]。英語の熟達度の低さも原因であるが、ディスカッションを円滑に進める質問やトピックに対する意見の出し方など、やり取りに不慣れなことや経験不足も原因として考えられる。文部科学省(2014)は、「今後の英語教育の改善・充実方策について 報告～グローバル化に対応した英語教育改革の五つの提言～」([https://www.mext.go.jp/b\\_menu/shingi/chousa/shotou/102/houkoku/attach/1352464.htm](https://www.mext.go.jp/b_menu/shingi/chousa/shotou/102/houkoku/attach/1352464.htm))で、小・中・高等学校が連携して4技能を活用した実際のコミュニケーションを行う言語活動の充実化を図ること、英語を用いてコミュニケーションを図る体験の必要性を示している[2]。特に大学との接続が求められる高等学校では、「国際社会の多様性に対応した目標・内容を設定し、幅広い話題について発表・討論・交渉などを行う言語活動の高度化を図ることが適当である」としている。

しかしながら、使用言語に限らずディスカッションは次の2つの点で運営が難しい。1点目は評価方法である。従来の授業では、教員評価と自己/ピア評価が行われている。教員評価は時間の制約から机間巡視して学生の様子を観察し評価することが一般的だが、この方法では教員がいないときの行動を観測できないので全体を見た評価とは言えない。自己/ピア評価は、評価の妥当性が学生のメタ認知能力に依存するため、必ずしも信頼できる結果とは限らない。2点目は良いディスカッションの定義が不明確である点である。良いディスカッションとは何かと尋ねられれば、グループの構成員個々の発話が多いことや構成員同士のやり取りが多いことなど、その答えは人によって様々である。西村(2018)は、ディスカッションを「参加者が積極的に意見を交換し、意見の相違点や新たな課題を確認することを通じて、正しい結論やより良い解決策を見つける共同作業」と述べている[3]。つまり、意味のある形式で話し手と聞き手の両方がやり取りを作り上げる力が問われる作業である。このような、やり取りを作り上げる力を「やり取りをする力(interactional competence; 以下ICとする, Galaczi & Taylor, 2018)」という[4]。仮にこのICを駆使して進めるディスカッションが「良い」ディスカッションであるとするれば、それは具体的にどのようなディスカッションなのであるか。教育現場では日々ペアやグループでの話し合いが活動として取り入れられている。話し合い後により良いディスカッションを目指して振り返りやフィード

バックがなされても、何らかの定量的な基準なしでは従来のように教員や学生の主観に頼らざるを得なくなる。

英語の授業における英語ディスカッションは、近年急激に普及したオンライン授業でさらに困難になっている。なぜなら、受講者がブレイクアウトルームに分かれて話し合いを行うと、教員とは完全に隔離されてしまうため、実質的に教員が観察できる時間がより限られるからである。また学生も他のグループの会話を聞くことができないため、自分たちの進行を周囲と比べるような認知ができない。このため、原田他(2020)はオンライン授業になると学生が安心して英語ディスカッションができるまでに対面より長い時間がかかると報告している[5]。

本稿では、著者の一部が開発したオンライン会議を可視化するシステム Hylable [6]を利用して、英語ディスカッションを可視化する授業を実践し、得られたデータの分析結果について報告する。Hylable はリアルタイムで会話の状態をフィードバックでき、発話やターンテイクの量を記録できるため、英語ディスカッションの課題であった(1)全ての英語ディスカッションの評価と、(2)定量的な「良い」ディスカッションの設計が解決する可能性がある。実際の英語ディスカッションの授業で収集したデータを用いて、ディスカッションデータを量的、質的側面からの分析と、学習者の認知に焦点を当てて報告する。具体的にはディスカッション時の参加者の行動データの分析や、発話内容との関係、学生の求めるフィードバック(ここでは他者から与えられる外的フィードバックを指す)についての調査を報告する。

本稿の構成は以下の通りである。まず2章で研究の背景についてまとめる。次に3章でディスカッションの可視化システムについて説明し、4章で実験条件について述べる。5章で収集したデータの3つの観点からの分析結果について述べ、6章でその結果について考察する。最後に7章で本稿をまとめる。

## 2. 研究の背景

第二言語/外国語教育において、ディスカッションスキルの向上は母語のそれと同様に重要視されており、スピーキング研究の枠組みでディスカッションスキル向上に向けた研究が数多く行なわれてきた。例えば方略トレーニングによるディスカッションスキルの変化や向上を報告した研究がその一例である。Lam & Wong (2000)は、香港人高校生58名を対象としたディスカッション時に使用する方略トレーニングの効果を報告している[7]。まずディスカッションで使用される方略についての気づきを高める導入から、方略を使用する実

実践トレーニングを行い、修正、改善をしながら方略使用の練習を繰り返した。その結果、グループ内での発話量ややり取りが増加し、ディスカッションでの自身の発話を明確化する行動が多くみられるようになったことを報告している。また Lam(2009)では、香港人高校生の英会話クラスで実施したメタ認知的方略の指導によって、英語ディスカッションでのパフォーマンスが向上したことが示された[8]。これらの研究は対面でのディスカッションについて検討したものであるが、モバイル端末を利用した語学学習 (Mobile Assisted Language Learning; MALL) として、オンライン上で学習者同士が1対1でやり取りをし、第二言語のスピーキング力が有意に向上し認知的変化がみられたことを示した研究もある[9]。本邦では、対話システム技術を活用して言語学習者のレベルや理解に合わせて質問を適応的に変更し、学習者の発話を引き出してスピーキング熟達度を評価するシステムの開発[10]が行なわれている。このシステムはあくまでも interviewer (AI) と interviewee (学習者) とのやり取りを評価対象としており、本稿で取り上げる複数人でのディスカッションは対象とはならない。

従来、スピーキング評価には、複雑さ、正確さ、流暢さ (CAF; [11]) の3つの指標に加えて、語彙[12]が項目として取り入れられている。しかしながら、小集団でのディスカッションでは、個人の発話能力や自分の意見を話し相手の意見を聞くスキルだけでなく、会話を継続させる IC も求められる。例えば相手の発話を精緻化するための質問や発言をすることや、やり取りが滞った際にやり取りを再開させたり促したりするような発言をすることなどが IC の一部として挙げられる。

1対1でのやり取りや複数人でのディスカッションは、話し手と聞き手双方が協力しながら行なうコミュニケーションである。個人を対象としたスピーキング評価では測りきれない力が IC の一部となっている[13]ため、ディスカッションの評価項目は個人のスピーキング対象のそれとは異なるものでなければならない。Stones(2021)は評価者の解釈は様々であること、評価に使用するルーブリックの項目が分かりづらいことを挙げ、IC の新たな評価方法について提案している[14]。しかしながら、ルーブリックを用いた評価者の自己判定に頼る評価方法であることには変わらず、その評価方法には客観性に乏しい部分が残る。

前述の点を踏まえると、ディスカッションの評価には話し手個人の力だけでなく、話し手と聞き手双方が創り上げる力を客観的に評価する指標を取り入れた、個人とグループ全体それぞれに対する客観的指標を用いた評価がふさわしい。このような背景から、小集団

でのディスカッション時の発話を正確に捉える客観的指標をフィードバックする対面議論可視化システム (第3章で述べるシステムの対面版) を活用した研究がある。松岡他(2019)は大学生が3~4人を1グループとして英語ディスカッションを行ない、議論評価システムで得られたグループ全体のターンテイク数を量的変数として分析した[15]。結果、グループ全体でのターンテイク数が少ないグループはメンバーの英語熟達度が高くメタ認知も高いことが明らかとなった。また森下・松岡(2020)は、対面議論可視化システムで得られた個人の発話量と発話量に対する自己評価との関係性について検証し、個人の発話量が多いほどそれに対する自己評価が高くなることを示した[16]。

このような、個人、グループ全体それぞれのレベルにおける定量的指標を英語のディスカッション評価に組み込むことによって、客観性の高いフィードバックが可能となる。また、従来英語教育の分野で行われてきたように、グループの発話内容を質的に分析し、ディスカッションが円滑に行われるための問いかけや発言 (例: 相手の発言を精緻化する言い換え) を指し示すようなフィードバックも取り入れると、量的、質的観点からのより良いフィードバックとなり得る。

ここで合わせて考えなければならないのが、学習者がフィードバックされる評価をどのように解釈し、次のパフォーマンス向上に活用するのかということである。フィードバックとは、エージェント (他者、教員) から個人 (学生) に対して与えられる、パフォーマンスや理解度に関する情報であり、指導や学習プロセスの一部として重要なものである[17]。教員と学生がフィードバックの情報を介して、学生のパフォーマンスを向上させるよう尽力するプロセスが想定される。しかしながら、教員側と学生側では、フィードバックや評価に対する見解が異なっている[18]という調査報告もある。つまりフィードバックの受け手となる学生が求めるフィードバックが、必ずしも教員が与えるそれと同じとは限らないことを意味するのである。さらに、学生側の課題として、フィードバックを十分に理解する力が不足していることも考えられる。

以上の背景から、本稿では英語のディスカッションデータを量的、質的双方から分析するとともに、学習者の認知的側面としてフィードバックに対する認知に焦点を当てて行なった調査結果を報告する。

### 3. 可視化システム

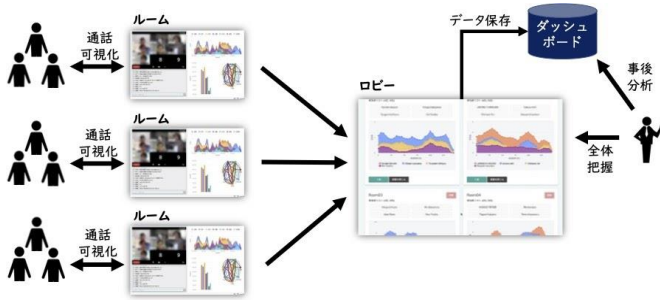


図 1. 可視化システム全体像

本項では、第3著者が開発したオンライングループディスカッションの可視化システム[6]を使用して英語ディスカッションを可視化した実践を報告する。本システムの全体像を図1に示す。本システムを使用する学生は、通常のウェブ会議システムのブレイクアウトルーム（同時にウェブ会議に参加している参加者を一時的に分割する「ブレイクアウト機能」によって複数の小グループに分けられた、各グループが行うウェブ会議のこと）のようにグループに分かれてウェブ会議に参加できる。ただし、各参加者の音声はサーバで発話区間検出され、定量化される。具体的には、次の3つの観点で定量化される。なお、音を検出したタイミングや継続時間に基づいて計量するため、使用されている言語を区別しない。

- 総発話量: 会議開始時刻から各学生が発話した発話時間の合計。棒グラフで表示する。
- 発話量の時間変化: 特定の長さの区間を指定し、それをスライドしながら、参加者ごとの発話時間の占める割合を計算した時系列データ。積み上げ折れ線グラフで表示する。
- ターンテイク: 発話区間を開始時刻の順にソートし、隣り合う発話者の組み合わせを数え上げたデータ。重み付き有向グラフで表示する。

これらの定量化されたデータは学生自身にリアルタイムでフィードバックされるので、学生は自分のグループのディスカッションの状態を視覚的に把握しながら議論を行える。この可視化データは、学生がディスカッションの状況をメタ認知するための足場かけとして作用する。一方、教員はこれらの会議ルームで収集された発話データを一覧できる「ロビー」を閲覧することで、全てのディスカッションを同時に把握することが可能である。また、これらのデータは蓄積されダッシュボードから閲覧できるので、授業後に分析や評

価も可能である。

本システムの特長は、自動音声認識技術を使わず、検出された音声区間の量やタイミングの情報を使用する点にある。特に英語初学者は英語でディスカッションするよう指示しても、日本語が混ざることもあり、また発音が明瞭でないことが多い。そのような状況では音声認識が困難になるため、自動音声認識技術に基づくテキストベースの分析は困難である。一方、発話の量やタイミングは発話の内容にかかわらず分析できるため、どのようなディスカッションでも同じ尺度で定量化することが可能である。

### 4. 条件

本項では、学生の英語グループディスカッションの条件について説明する。

ディスカッションの対象者は、関西圏にある私立大学で2020年度前期、後期の英語会話クラスを受講した1, 2年生のうち、調査協力に同意した計104名であった。そのうち次項の5.2, 5.3の結果については後期クラスを履修した43名のデータを分析対象とした。

参加者はZoom上で集合した後、4~5名で1組のグループにランダムに振り分けられ、前項で説明したオンライングループディスカッションの可視化システムHylableに移動した。ディスカッションのテーマについては、事前に授業内で関連するリーディング、語彙学習のテキストで扱われたものであり、参加者はディスカッション実施の1週間前からディスカッションで話す内容がある程度準備しておくよう指示を受けていた。当日はグループごとに英語での話し合いを15分程度行った。ディスカッション中、参加者はリアルタイムで自身やグループの他の参加者の発話量をグラフで見ることが出来た。

ディスカッション後、参加者はオンラインで質問紙調査の項目に回答をした。質問の内容は参加者の属性、PC環境、英語オンラインディスカッションの経験回数、当日の活動への参加度（自己評価）、そしてフィードバックに関して問うものであった。参加者は自己評価をグループディスカッションの事後レポート（後述）を見る前に行った。

当日のパフォーマンスについては、グループごとに可視化された事後レポートが各参加者にオンライン上で提示された。参加者は自分のグループの発話量やターンテイクをまとめたデータを見て、自分がどの程度グループ内でディスカッションに貢献していたかを確認することができた。

## 5. 結果

### 5.1. 学生の求めるフィードバック

本項では、学生の求めるフィードバックとディスカッション時の他者の様子（発言や行動等）から得られる気づきに関する質問紙調査の結果を報告する。回答者は102名であった。オンラインでの英語ディスカッションの経験数によって参加者をグループに分けて結果を比較した。未経験、経験回数5回未満、経験回数5回～10回、経験回数10回以上の4グループを設定した。

学生の求めるフィードバックについては、参加者に「英語のディスカッション能力の向上にはどのようなフィードバックが有効だと思いますか.」という質問に対し、計8つの選択肢から複数回答を可能とし回答を求めた。この選択肢は、第2著者が過去の英語会話の授業で実際に使用したフィードバックの形式を基に設定した。個人のパフォーマンスに関しては、成長（発言量ややり取りの増加）、会話評価（発言の内容に対しての即時的評価）、行動（グループ内での役割）、参加度（グループ内での発言量）に対するフィードバック、そして事後の録音確認（ディスカッション後に参加者個人が録音を確認する）、録音を基にした事後フィードバック（ディスカッション後に与えられるコメント）が設定された。グループ全体のパフォーマンスに関しては、ロールモデルの提示（ディスカッション後にうまくディスカッションができたグループをモデルとして示される）、教員の即時的フィードバック（教員がディスカッションに参加しその場でアドバイスを求める割合は、両方とも56%であった。また、発言量（発言の時間量）ややり取り（ターンテイク数）の増加といった、自分の成長が確かめられるようなフィードバックを求める割合は44%であった。

参加者のうち、10回以上の英語オンラインディスカッションを経験している学生（計9名、以下経験者とす）と未経験者（計50名）の結果を選んで比較を行った。経験者は、総じて選んだ選択肢の数が少なく、個人の発言に対する即時的評価とグループ内での役割や今後の活動へのアドバイスを求める割合は、両方とも56%であった。また、発言量（発言の時間量）ややり取り（ターンテイク数）の増加といった、自分の成長が確かめられるようなフィードバックを求める割合は44%であった。

一方で未経験者50名はほぼ全ての選択肢を選ぶ傾向にあった。最も回答の割合が多かったのが自分の成長が分かるフィードバック（56%）であり、次いで個人の発言に対する即時的評価（52%）であった。そして今後の活動へのアドバイス（46%）、録音を基にした事後フィードバック（46%）と続いた。10回以上の経験者と概ね同様の傾向が見られたが、ディスカッションの録音を基にした事後フィードバックを求めていることが異なっている点となった（経験者は22%）。

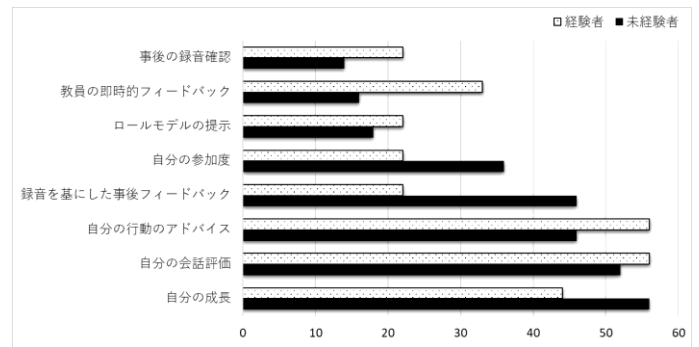


図2. 英語ディスカッション能力向上に有効なフィードバックの割合

ディスカッション時の他者の様子（発言や行動等）については、参加者に「同じグループの人のdiscussionのやり方をみて、自分の行動を変えたいと思ったことは何ですか.」という質問に対し、自由記述で回答を求めた。なお記述のコーディングは、第1著者と教育心理学を専攻する大学院生（博士課程）との二者で行った。まず第1著者がKH Coderを用いて共起ネットワークを確認し、生成されたグループを基にコーディングの項目を作成した。計8つのグループが確認されたが、そのうちの2つは非常に似た内容を含んでおり、部分的につながりが発生していた。そのため、二者で協議を行い最終的に以下の7項目を設定した。なお、括弧内の語は図3の項目名を示している。

1. 自己の発言量の増加や話題提供（自己）
2. 相手への質問やコメントの増加（他者）
3. 役割に対する積極性（役割）
4. 英語を話す／聞く態度（態度）
5. ディスカッションの方法や進め方（方法）
6. 流暢さの向上（流暢さ）
7. 言語スキルの向上（言語）

上記の項目でコーディングを行った後、評定者間信頼性を求めたところ $\kappa=0.82$ となり、高い一致率を確認することができた。不一致のコーディングについては二者で協議し決定した。結果は図3の通りである。

経験者（N=52）で最も多かった回答は、1. 自己の発言量の増加や話題提供に関する気づき（38.5%）であった。次いで2. 相手への質問やコメントの増加に関する気づき（21.2%）であり、未経験者（N=50）もほぼ同様にこの2つの項目についての回答が約6割を占めた。特にこの2つの項目についてはHylableの発話量の推移（波線グラフ）やターンテイク数などのデータの量的側面と関係するところである。経験の有無に関わらず、活発なやり取りをディスカッションの理想とする傾向が推察される。

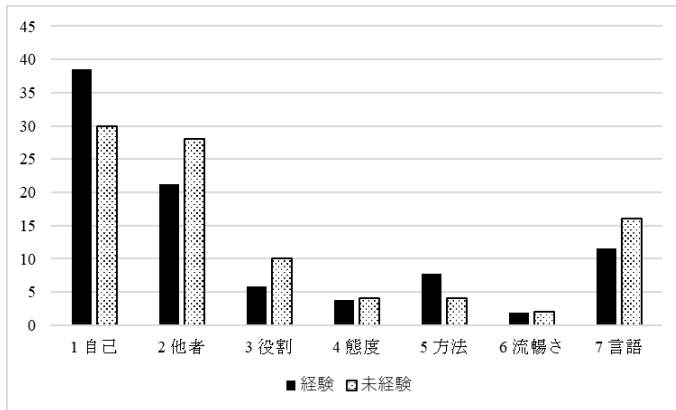


図3. 他者行動から得られた気づきの割合

3番目に多かった項目は、7. 言語スキルの向上に関する気づき（11.5%）であり、未経験者の方が割合はやや大きかった（16%）。言語スキルの向上とは、例えば瞬時に使う（口から出てくる）ことができる表現を増やすことや、言い回しのバリエーションを豊かにすることなどを指す。

4番目に多かった項目は、経験者と未経験者で異なった。経験者はグループ全体のディスカッションの進め方についての気づき（7.7%）を得ていた一方で、未経験者は個人の役割に対する積極性に関する気づき（10%）を得た結果となった。ディスカッションの進め方については、役割に対する積極性につながる項目ではあるが、向けられた視点が異なると判断し別項目とした。

## 5.2. 発話行動の分析

本項では、一般化線形モデル（GLM; Generalized Linear Model）を用いて、学生の発話の割合と平均ターンテイク数から貢献度の自己評価を推定した結果を報告する。データは2020年後期の英語会話授業受講者（1, 2年生）43名のものを使用した。モデルでは相互作用は有意でなかったため、発話量とターンテイク数は主効果のみ分析した。発話量はディスカッションの中で個別の学生が発話した時間（msec）を総ディスカッション時間（msec）で割って算出した。ターンテイクは、「自分が話した後に他人が話した」場合に1ターンとし、自分から他人へのターンテイクの発生回数をターンテイク数とした。自己評価は0（参加できなかった）から5（たくさん話せた）までの5段階評価とした。また、この分析には比較的人数が多い1年生（30名）のデータで行われた。2年生は比較的少ない人数であったため分析はしていないものの、表1, 2では同じような傾向が見られた。

表1, 2は学生の自己評価と発話量、そしてターンテ

イク数の関係性をまとめたものである。分析の結果、自己評価を発話量から推定した場合（ $(F(1, 41)= 12.83, p < .001)$ ）、発話量が多いほど学生の自己評価が高い事（ $t(41)= 3.58, p < .001$ ）が分かった。また、ターンテイク数が多いほど、学生の自己評価が高い（ $t(41)= 3.58, p < .001$ ）という事も分かった。自分の発話量だけではなく、他人の反応も自己評価に反映されていたと見られる。

表1. 発話量に対する自己評価と実際の発話量

自己評価	1年生（2回目）		2年生（2回目）	
	発話量	人数	発話量	人数
0 参加できなかった	0	0	0.04	1
1 全く話せなかった	0	0	0	0
2 少し話せた	0.23	10	0.12	4
3 まあまあ話せた	0.25	12	0.32	4
4 やや多めに話せた	0.32	4	0.3	4
5 たくさん話せた	0.54	4	0	0

表2. 発話量に対する自己評価とターンテイク数

自己評価	1年生（2回目）		2年生（2回目）	
	ターンテイク数	人数	ターンテイク数	人数
0 参加できなかった	0	0	56	1
1 全く話せなかった	0	0	0	0
2 少し話せた	356.1	10	160	4
3 まあまあ話せた	318.33	12	185.5	4
4 やや多めに話せた	432.75	4	255	4
5 たくさん話せた	762.25	4	0	0

## 5.3. 発話テキストの分析

本項では、録音音声の書き起こしデータに焦点を当て、英語で話している割合やその内容に関する分析結果について報告する。データは5.2の結果と同様、2020年後期の英語会話授業受講者（1, 2年生）43名のものを使用した。

全10グループの録音音声について、マイクロソフト社が提供する Microsoft Azure Speech to Text の日本語モードで音声認識を行ったところ、英語発話の場合は自動で英語表記になるはずが、発音の問題などで誤認

識され、カタカナ表記になった。Microsoft Azure の英語モードでも音声認識を行ったが、すべてアルファベット表記のため、英語と日本語の判別はテキストからは困難だった。そこで、冒頭 5 分のみ録音音声につき、業者委託による書き起こしに変更し、英語と日本語の切り分けについて高い信頼性を確保することとした。

書き起こしデータを確認したところ、a. 雑談しているだけのグループ、b. 準備した英文を順に読み上げているだけのグループ、c. 英語でディスカッションがしっかりできているグループのおおむね 3 種類が存在することが分かった。英語と日本語の割合については、a. 雑談グループはほとんど日本語、b. 読み上げグループは準備した英文以外はほとんど日本語、c. ディスカッショングループはほとんど英語を使っている傾向が見られた。b. 読み上げグループでは、お互いに内容を理解できていないためか、他のグループに比べて相槌などのバックチャネルやターンテイクが少ない傾向が見られた。また、「次、俺か」、「次、どうぞ」のように、順番を確認するような発言が目立ち、ディスカッションが成り立っていない様子が認められた。

さらに詳しく分析するため、すべての書き起こしデータについてグループごとに単語数を集計したところ、表 3 のような結果となった。ここでの単語の分割には、形態素解析ツール Nagisa[19]を使用した。このツールには、深層学習モデルを用いた系列ラベリング手法が実装されており、同モデルは日本語・英語を含む大量のコーパスを使って学習されているため、話し言葉や英語が混ざった文であってもうまく単語への分割ができた。

表 3. グループの特徴および品詞ごとの単語数と英語を使用していた割合

#	グループの特徴	学年	英語	日本語	英語の割合	総ターンテイク数
1	a.雑談	1年生	36	1772	2%	3200
2	a.雑談	1年生	99	664	13%	1252
3	a.雑談	1年生	25	1541	2%	2057
4	a.雑談	1年生	44	1764	2%	2773
5	b.読み上げ	1年生	222	575	28%	830
6	b.読み上げ	1年生	72	816	8%	1418
7	b.読み上げ	1年生	386	511	43%	621
8	c.ディスカッション	2年生	576	137	81%	1162
9	c.ディスカッション	2年生	779	110	88%	449
10	c.ディスカッション	2年生	622	32	95%	962

これらの結果は、書き起こしデータから筆者らが読み取った傾向とおおむね合致していた。しかしながら、3 種類のグループのうち、b. 読み上げグループについては、準備した英文の量などにより、各グループ (#5,

6, 7) 間で、英語の単語数における差が比較的大きいことが分かった (222 語 vs. 72 語 vs. 386 語)。一方、日本語の単語数においてはそれほど大きな差がなく (575 語 vs. 816 語 vs. 511 語)、準備した英文を読み上げる以外はほとんど日本語で話していたことを反映する結果となった。また、英語の単語数だけを見ると、グループ #7 とグループ #8 の差はそれほど大きくないものの (386 語 vs. 576 語)、英語を使用していた割合については、グループ #7 はグループ #8 の半分程度であり (43% vs. 81%)、b. 読み上げグループと c. ディスカッショングループの違いは歴然としている。

各グループの書き起こしから、特徴的な箇所をピックアップして以下に挙げる。

<読み上げグループ 5 の例>

- A: 先どうぞ。  
 B: 私ですか。My New Year's resolutions is 合ってるんかな、to submit assignment on time and to get up early arise and 何やったっけ? And もう 1 個。Exercise. 終わりです。  
 C: Exercise ばり分かる。  
 B: 今サークルでな、(## ##)。  
 C: (## ##)好きやもんな。今週はないけどな。  
 B: 悲しい。どうぞ。  
 C: A さん、いこう。  
 A: はい。Enjoying campus life. If the COVID-19 settles down, I want to go abroad. Since I'm 20 years old, I hope I can grow up in many ways.  
 B: いける、C ちゃん?

<読み上げグループ 6 の例>

- A: 何しゃべるんやっただけ。  
 B: 新年、どう過ごしたか、みたいな。How did you spend the New Year みたいな感じちゃう。  
 C: 早くコロナ終わってほしいけどな。  
 B: ほんまにな。おまえ、日本語。もう日本語。せっかく俺が英語に振ってやったのに、すぐ日本語。おかしいわ。  
 C: いや、ほんまに。  
 B: ずっとバイトしてたわ、年末年始。

<読み上げグループ 7 の例>

- A: 最初のトピック。Think of three New Year's resolutions からやりましょう。  
 B: 誰からいきますか。どうしよう。私からいきますね。Think of three New Year's resolutions. First, get a good night's sleep and don't eat too many sweets. And third, don't put off assignments. Last year, I stayed up

until midnight to do a lot of assignments. So this year, I want to live a healthier life.

A : 次, やりますね. I want to get the driver's license during the spring vacation. Second, I have to study English hard and improve my TOEIC score. Last, I want to do a part-time job and earn much money. Thank you. 終わりです.

C : 次, いきます. First, I want to be able to speak more English. Second, I want to take a picture of nature. Second, I want to...

D : これ, 声聞こえてますか.

一同 : 聞こえてます.

D : 皆さん, めっちゃ, ちゃんと英語でやるんですね.

B : めっちゃ, 考えた. 昨日.

<読み上げグループ 8 の例>

A : これ全員で 5 人かな?

B : スタートします?

C : OK. So 英語, 今から?

B : 英語.

C : OK. So who would like to say your New Year's resolutions?

A : I have three resolutions.

C : You have, already? OK. Would you tell us?

A : Always smile. I think in 2020, we were hit by many disasters, and I can't always be smile, so in 2021, I want to be smile and be positive. Do you have any resolutions?

C : Are there other two New Year's resolutions?

<読み上げグループ 9 の例>

A : OK. 始めます. First of all, Happy New Year, everyone. And today, we're going to start discussions. OK. I will ask you three questions, and please answer each question. First, think of three New Year's resolutions and discuss them with your group. What is your New Year's resolutions for B-san?

B : I spent the entire New Year's holiday at home. I didn't do anything specially, didn't go to see the first sunrise and didn't go on New Year's visit. That's all.

A : Thank you. And how about C-san. What is your three New Year's resolutions?

C : I want to change my lifestyle. Particular, I want to take care of my health. So that I will get at least seven hour of sleep every day.

<読み上げグループ 10 の例>

A : OK. So maybe everybody are here. So let's get started the discussion.

B&C : Yes.

A : The topic is New Year's and COVID-19 these days. So number one is about our resolutions in New Year's.

D : OK.

A : So did you come up with some ideas already? Not yet?

B : Yes. My New Year's resolution is improve my English conversation skills. And I want to use academic English because this year, sorry, I want to improve my English skill this year.

A : OK. I also think of that. This year, due to the COVID-19, we couldn't be outside. That means we couldn't communicate with others in English much. I think I want to keep my health the best in this year, also, and I want to improve my English skill too. And I want to find another expertise, such as management or economic or something like law. These are my resolutions in this year. So how about others?

雑談グループはほとんど日本語で話しており, ターンテイク数が多かった ( $F(2,6)=6.01, p<.05, \eta^2=.63$ , ANOVA, グループ間での有意差なし). 読み上げグループは英語の割合にはばらつきがあるが, 日本語の単語数には大きな差はなかった. 準備した英文を読み上げる以外はほとんど日本語だったためと考えられる. ディスカッショングループは概して英語の割合が高かった ( $F(2,6)=53.4, p<.01, \eta^2=.94$ , ANOVA, ディスカッショングループとの間で有意差あり; 雑談-ディスカッション ( $p<.01$ ) / 読み上げ-ディスカッション ( $p<.01$ ) / 雑談-読み上げ ( $n.s.$ )). 英語の発話量とターンテイク数は反比例している傾向 ( $r=-.68, p<.05$ ; 負の相関) が見られ, 相槌を打てるほど内容を理解できていないか, 英語での適切な相槌を使いこなすことができず, 自分の意見の読み上げにとどまっている可能性が示された.

## 6. 考察

5.1 で示した学生への質問紙調査の結果から, 学生は英語ディスカッションの経験を積むほど(1) 自分の行動に対する評価やアドバイスを求めており, (2) 行動評価の視点が自分自身からグループ全体に広がっていることが分かった. 複数のグループが同時進行する英語ディスカッションの場面では, 教員が各学生に対してこのようなきめ細かいフィードバックを行うことは現実的ではない. その点, 本稿で実践したグループ全体の英語ディスカッションを可視化したデータをフィードバックする方法を使えば, 自分の行動の状況とグループ全体の状況が把握できるため, 学生の求めているフィードバックの提供が期待できる. ただし, 本



システムで提供するのデータそのものであるので、データから自身のメタ認知の促進に繋がる情報を得るには、フィードバックデータの読み方のレクチャー等が必要であろう。

5.2 の発話データの分析からは、自己の発話量だけでなくターンテイクも多い学生ほど多く話せたという自己評価をしていることが分かる。すなわち、英語ディスカッションにおいては実際の発話量に加えて相手とのやりとりも「多く話したという感覚」に影響を与えている。このことから、特に初学者で多く話すことができたという実感を持つことが必要な段階では、ターンテイクの多さが「良い」ディスカッションの指標の一つとなっていると考えられる。ただし、ディベートや交渉などより高度なディスカッションにおいては、単純なターンテイクの量よりもその場面に応じた指標を組み合わせたディスカッションの「良さ」の設計が必要だろう。

5.3 の書き起こしデータの比較からは、同じ英語を使って会話しているグループでも、準備してきた原稿を読み上げたグループと実際にディスカッションを行ったグループは、ターンテイクの量で分類できる可能性が示唆された。似たアプローチで、和田他 (2021) はグループワークにおける発話量と画面共有をしている画像の変化量から、沈黙している区間の理由を、各自が個人の作業をしているからなのか、全員の作業が止まってしまったからなのかを判別する方法について報告している [20]。このように、会話データと他の情報を組み合わせることによって、学生に対する個別化された形成的評価や、目的に合わせた「良い」ディスカッションを設計できると期待できる。

これらのことから、英語ディスカッションを定量化して可視化データをフィードバックすることは、学生の求めるニーズを満たすだけでなく、学生の自己評価と合う「良い」ディスカッションの設計を行うデータに使えると言える。

本稿で用いた分析手法は言語を区別しないため、日本語で参加しても発話として計測される。この理由のひとつは、とくに英語初学者が日本語混じりの文法かつ日本語と似た発音で話したときの、英語との区別の難しさにある (例えば「スタートする?オーケー」が日本語混じりの英語か、日本語か一意に定まらない。) そのため、言語自動判別問題の解決を避け、自動で即時フィードバックを行うことを優先した。

「良い」ディスカッションは、多くの場合は語彙や文法など言語的な情報に基づいて評価されることが一般的だが、このアプローチは協同学習の態度の観点に着目していると解釈できる。なぜなら、たとえ学習者が日本語混じりの英語や日本語しか話せなくても、デ

ィスカッションに参加した事実を評価しているからである。日本語の雑談は「流暢に話せていた場合の会話パターン」になるので、本人たちの潜在的な会話能力を示しているといえる。そのため、英語で会話したときのターンテイクの目標として設定することができる。

## 7. まとめ

本稿では英語のディスカッションデータを量的、質的双方から分析し、さらに学習者の認知的側面としてフィードバックに対する認知に焦点を当てた調査結果について報告した。その結果、特にターンテイクが参加できたという自己評価との関連が高いことを示した。また、書き起こしデータと比較することで、自然な英語ディスカッションと英文読み上げの区別が付く可能性が示唆された。アンケート結果からは、このような可視化データをフィードバックすることは、学生が求めているフィードバックの種類にも合っていると見える。

今後は、学生のメタ認知をより促すような授業設計やフィードバックの方法などに関する検討をしていく。また、実験では特定の大学の一時期の学生のみを対象としていたので、参加者の数や学年を広げて、この結果がどの属性を持つ学習者にまで一般化できるのかを調査していく必要がある。とくに、本稿では初学者の英語ディスカッションを対象としたため、英語ディスカッション中の日本語の使用は避け難かった。より英語に慣れた学生を対象に含めることで、使用言語と自己評価と関係も分析できるだろう。

## 謝 辞

本稿は、2020年度神戸学院大学教育改革助成金『タマゴ型マイクロホンアレイを用いた英語ディスカッションの評価:メタ認知,協同学習態度との関連性から』(研究代表者:大竹翔子)の助成を受けている。

## 付 記

本稿は、筆者たちが2021年6月16~18日に開催された International Association for Language Learning Technology (IALLT) 2021ならびに2021年8月7~8日に開催された全国英語教育学会第46回大会(いずれもオンライン)において口頭発表した内容を加筆・修正したものである。

## 文 献

- [1] Yanagi, M. & Baker, A. A. (2016). Challenges experienced by Japanese students with oral communication skills in Australian universities.

- TESOL Journal, 7 (3), 621-644.
- [2] 文部科学省(2014). 今後の英語教育の改善・充実方策について報告～グローバル化に対応した英語教育改革の五つの提言～, [https://www.mext.go.jp/b\\_menu/shingi/chousa/shotou/102/houkoku/attach/1352464.htm](https://www.mext.go.jp/b_menu/shingi/chousa/shotou/102/houkoku/attach/1352464.htm) (2021-12-09 閲覧)
- [3] 西村智朗(2018). ディスカッションの方法. [http://www.ritsumei.ac.jp/ir/ir-navi/common/pdf/technic/technic\\_text\\_09\\_2019.pdf](http://www.ritsumei.ac.jp/ir/ir-navi/common/pdf/technic/technic_text_09_2019.pdf) (2022-02-28 閲覧)
- [4] Galaczi, E., & Taylor, L. (2018). Interactional competence: Conceptualisations, operationalisations, and outstanding questions. *Language Assessment Quarterly*, 15(3), 219–236.
- [5] 原田康也・赤塚祐哉・坪田康・鍋井理沙・森下美和(2020). オンライン授業における学生間のインタラクション(相互作用)と全人的な交流機会の担保, 日本認知科学会第 37 回大会発表論文集, 303-312.
- [6] 水本武志・中村祐希(2021).オンライングループディスカッションの可視化技術と実践モデル, 日本教育工学会 春季全国大会(オンライン開催), 6-S031-2.
- [7] Lam, W.Y.K., & Wong, J. (2000). The effects of strategy training on developing discussion skills in an ESL classroom, *ELT Journal*, 54(3), 245–255.
- [8] Lam, W.Y.K. (2009). Examining the effects of metacognitive strategy instruction on ESL group discussions: A synthesis of approaches. *Language Teaching Research*, 13(2), 129-150.
- [9] Akkara, S., Anumula, V. & Mallampalli, M. (2020). Impact of WhatsApp Interaction on Improving L2 Speaking Skills. *International Journal of Emerging Technologies in Learning (iJET)*, 15(3), 250-259.
- [10] 佐伯真於・松浦瑠希・鈴木駿吾・宮城琴佳・小林哲則・松山洋一(2021).IntelLLA: 適応的な質問戦略を有するスピーキング能力判定会話エージェント, 人工知能学会研究会資料 言語・音声理解と対話処理研究会 93 回,15-20.
- [11] Housen, A., & Kuiken, F. (2009). Complexity, accuracy and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461-473.
- [12] Skehan, P. & Foster, P. (2008). Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the Ealing research. In Van Daele, S., Housen, A., Kuiken, F., Pierrard, M. & Vedder, I. (Eds.). *Complexity, Accuracy, and Fluency in Second Language Use, Learning, & Teaching*. pp. 207-226. Brussels: Contactforum.
- [13] McNamara, T. F. (1997). ‘Interaction’ in second language performance assessment: Whose performance? *Applied linguistics*, 18(4), 446-466.
- [14] Stone, T. (2021). Developing a Rating Scale for Assessing Interactional Competence. Oral presentation at Japan Association for Language Teaching (JALT) 2021, A virtual conference.
- [15] 松岡真由子・大竹翔子・森下美和・海老原由貴(2019). タマゴ型マイクロホンアレイを用いた英語ディスカッションの評価—メタ認知, 協同学習態度との関連性から—, 全国英語教育学会第 45 回弘前研究大会, 口頭発表.
- [16] 森下美和・松岡真由子(2020).ウェブ会議システムの自動データ化機能を利用したディスカッションの分析, 2020 年度大学英語教育学会 (JACET) 関西支部大会(インターネットによるオンデマンド動画配信), 口頭発表.
- [17] Hattie, J., & H, Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- [18] Carless, D. (2006). Differing Perceptions in the Feedback Process. *Studies in Higher Education*, 31(2), 219-233.
- [19] 池田大志(2018). RNN を用いた日本語単語分割/品詞タグ付けツールの紹介, 言語処理学会第 24 回年次大会ワークショップ, 岡山コンベンションセンター.
- [20] 和田秀飛・堀越泉・石井優奈・田村恭久(2021). 発言量と資料編集量を用いた資料作成型グループディスカッションの状況可視化, 日本教育工学会, 2021 年秋季全国大会(オンライン開催)口頭発表.