平成27年度研究助成 【サウンド技術振興部門】より

パブリック・スピーキングの定量的な指標の確立を目指して

九州大学大学院
芸術工学研究院　助教
山下　友子

九州大学大学院
芸術工学研究院　助教
冬野　美晴

九州大学大学院
芸術工学研究院　教授
中島　祥好

アイルランド国立大学、
ゴールウェイ校　Lecturer
Mark A Elliott

アイルランド国立大学、
ゴールウェイ校　Senior Lecturer
Stanislava Antonijevic-Elliott

1．はじめに

パブリック・スピーキングとは、大勢の聴衆を相手に自分の考えを論理と感情の絶妙な組み合わせでもって伝えることである。人類の歴史上、偉大な人物の優れたスピーチは、人に感動を与え、ときに人を扇動し、歴史を大きく動かす一端となってきた。情報技術が発展し、グローバル化が進む現代においては、様々なツールを用いて国内外の多くの相手に対して自分の考えを伝えることが可能である。近年、世界中の誰でもが、いつでもどこでも自由に映像・音声を見ることができるようになっており、今や誰もが世界に情報を発信することができる時代である。このような時代の中で、日本人は、論理と感情の両方に訴えるようなスピーチを苦手とする場合が多く損をしている可能性がある。多様な文化的背景や価値観を持つ人々

を相手に自分の考えを効果的に伝え、迅速で正確な意思疎通を図ることができるスキルが今、強く求められている。そこで本研究では、パブリック・スピーキングにおける音声データの定量的な指標と、応用言語学の理論を用いたテキスト分析を統合することで、文理融合型の統合的な分析を行う。このようにして得られた結果はパブリック・スピーキングを苦手とする日本人のための、日本人による実践的な日本語/英語スピーチ指導法の開発につながることが期待される。

2．プロジェクト全体の概要と目的

本研究プロジェクト全体の概要として、音声学、応用言語学、英語教育、画像工学、心理学のそれぞれの分野の研究者や教育者が連携して、各分野の知見を活かした実証的なデータの収集と分析を試みている。得られた実証的な結果を基に、「科学的根拠に基づくパブリック・スピーキング教育」を確立することが本研究プロジェクトの目的である（図１）。

研究チームは、2013年から、暗唱大会・スピーチ・コンテスト、英語教育の現場にて日本人（高校生・大学生）のスピーチ・パフォーマンスを録音・録画して、独自に英語学習者のパブリック・スピーキング・マルチモーダルコーパスを作成して音声・画像分析を行ってきた（Fuyuno et al., 2014）。また、米国における大学の卒業式での祝辞スピーチを用いて、高い水準のスピーチ能力を身につけている話者に共通した要素を明らかにするために画像・音声分析を行ってきた（Yamashita & Fuyuno, 2015）。これまでの研究では、非英語母語話者の学生と、英語母語話者の世界的リーダーとを比べ、相違点が明らかになっている。

今回、カワイサウンド技術・音楽振興財団の助成金を受けて、アイルランド国立大学（ゴールウェイ校）にて、共同研究者の協力のもと英母語話者の大学生を対象に、スピーチ・パフォーマンスを録音･録画して、独自にパブリック・スピーキング・マルチモーダルコーパスを作成して音声・画像分析を行う予定である。2015年

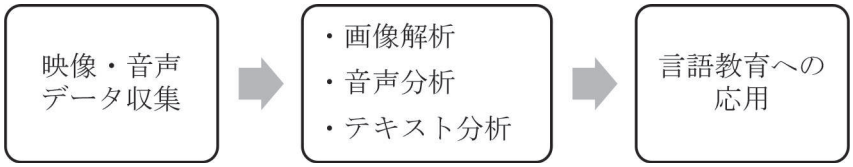

図１　本研究プロジェクト概要

11－12月に、日本人研究者３名が、アイルランド国立大学を訪問する予定である。アイルランド国立大学所属の研究者２名と協力して、大学生のスピーチ・パフォーマンスを録音・録画するために、準備を進めている。現段階では、九州大学、アイルランド国立大学の両大学において倫理審査を通過している。９月頃から研究参加者を募っており、現在20名以上の大学生が参加予定である。

３．予備的調査：東南アジアの大学院生によるパブリック・スピーキング

アイルランド国立大学にて録音・録画を行う前に、予備的調査として、東南アジアの大学院生を対象にパブリック・スピーキングの録音・録画を試みた。東南アジアの大学院生は、本大学の芸術工学府デザイン人間科学国際コースに８月末から２週間ほど受け入れていた学生で、マレーシアやインドネシア出身の学生であった。

録音・録画に参加した学生は、マレーシアの大学院生（女性１名）、インドネシアの大学院生（男性３名、女性２名）であった。録音・録画は、本大学の教室にて行われた。各学生は、まず、英語学習に関するアンケート調査に回答した。インドネシアの大学院生（５名）は英語を第二言語と日常生活で使う学生はおらず、英語を利用する頻度は外国人の友人や教員との会話が月に数回程度であった。マレーシア出身の大学院生（１名）は、言語学専攻であり、小学生を対象に英語を教える仕事に従事していた。

アンケート調査の後、各学生は、英語のスクリプトを手渡された。英語のスクリプトには、Larry Page氏によるミシガン大学の卒業式（2009年）の式辞の一部を用いた（http://googlepress.blogspot.jp/2009/05/larry-pages-university-of-michigan.html）。英語スクリプトの単語数は309単語であった。各学生はスピーチの準備時間として５分与えられた。５分後に、観客（研究者および大学院生３－４名）の前で各学生は、スピーチを行った。スクリプトを参照することは許可されていた。教室の教卓の上に録音機１台、教室の後方にデジタルビデオ録画機１台を設営して、録音･録画を行った。

録音・録画したデータを保存してデータベース化した。学生による音声データとLarry Page本人の音声データ（https://www.youtube.com/watch?v=qFb２rvmrahc）を分析した。

雑音がある部分などは分析対象外として英文スクリプトの同パラグラフ、約60秒を分析対象とした。音声分析ソフトを用いて、0.1秒以上の無音区間（ポーズ）を自動で抽出した。そして、0.1秒以上のポーズで挟まれた区間に、スクリプトの英文を、アノテーションする作業を行った。図２に、Larry Page, インドネシアの学生男性２名（IM01, IM02）を例として示す。

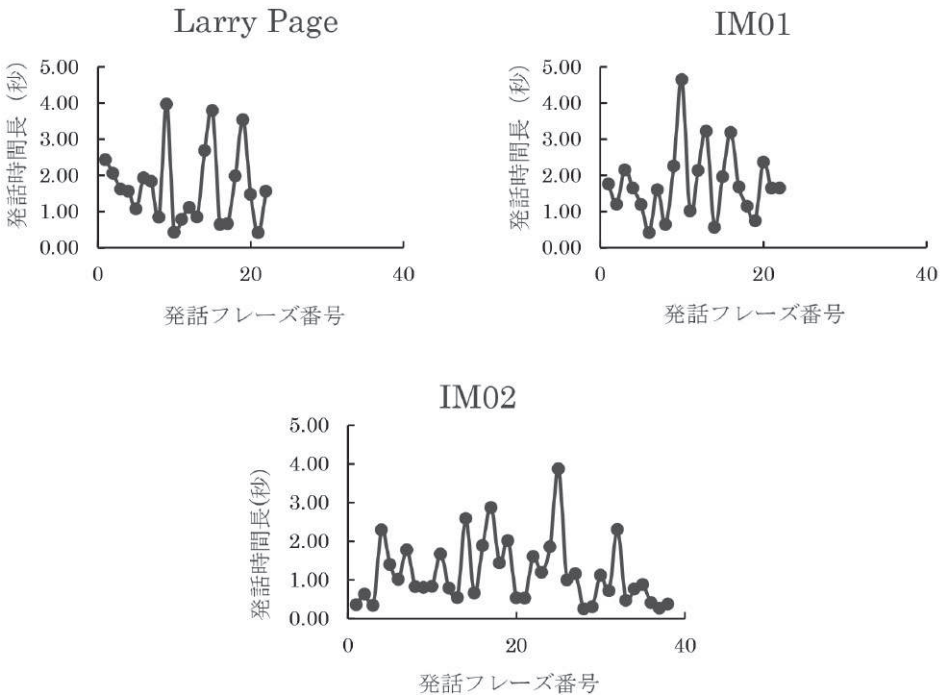

図２　横軸は発話フレーズ番号、縦軸は、発話区間の時間長（秒）である。横軸の発話フレーズ番号とは、0.1秒以上の無音区間（ポーズ）に挟まれた発話フレーズに最初から番号をつけたものであり、Larry PageとIM01, IM02では、ポーズの位置が異なるため発話フレーズ番号は異なる。

発話時間長の中央値は、Larry Page＝1.56 秒、IM01＝1.65 秒、IM02＝0.86 秒であった。図２に示すように、発話フレーズ頻度数は、Larry PageとIM01は約20フレーズ、IM02は約40フレーズであり約２倍のポーズで発話を区切っている。３秒以上の長い発話時間長の発話テキストを調べた。下記はLarry Pageの発話文（３秒以上）である。IM01は、(1)、(2)の発話フレーズを含む文で、IM02は、(2)の発話フレーズを含む文で、３秒以上の発話時間長であった。ポーズを挿入せず、３秒以上の大きなまとまりで読み上げることで、聴衆は聴きとりやすくなり、その結果、説得力が増すのかもしれない。

(1) I spent the middle of that night scribbling out the details and convincing myself

(2) **he nodded knowingly, fully aware it would take much longer**

(3) **Amazingly, I had no thought of building a search engine.**

３秒未満の発話フレーズにおいては、その区切り方が大切なのかもしれない。下記に例として、各発話者のポーズの位置を●で記して列挙する。Larry Pageの発話文では、ポーズは挿入されていない。一方、IM01では動詞の直後、IM02では、主語の部分である名詞句において、２回ポーズが挿入されている。今後、構文解析と照らし合わせて、パブリック・スピーチにおける適切なポーズの位置も検討していきたいと考えている。

〈Larry Page〉

The optimism of youth is often underrated!

〈IM01〉

The optimism of youth is ● often underrated!

〈IM02〉

The optimism ● of ● youth ● is often underrated!

4．今後の展望

今後の予定として、アイルランド国立大学（ゴールウェイ校）での英語話者の大学生による英語スピーチ音声や、九州大学での日本人英語学習者の英語スピーチ音声を収録してデータベース化する。独自に作成するデータベースのスピーチ音声において、文全体の時間構造（例えばポーズの時間長、発話フレーズ時間、発話速度）、語彙の頻度や難易度、適切な意味チャンクなどを検討する予定である。音声データの定量的な指標と応用言語学の理論を用いたテキスト分析を統合していく予定である。

謝　辞

カワイサウンド技術・音楽振興財団に研究者一同、感謝の意を表します。

参考文献

Fuyuno, M., Yamashita, Y., Kawase, Y., & Nakajima, Y. (2014). Analyzing speech pauses and facial movement patterns in multimodal public-speaking data of EFL learners. *Learner Corpus Studies in Asia and the World, 2*, 237 – 251

Yamashita, Y & Fuyuno, M. (2015). The characteristics of a good public speech performance in terms of phonological and non-verbal gesture aspects, *ATEM Journal, 20*, 33 – 44