

発話の流暢性における複数語ユニットの有用性: 複数語ユニットの知識と使用の観点から

研究者:東京都/早稲田大学大学院 在籍 瀧澤 嵩太郎

《研究助言者:小泉 利恵》

概要

本研究では、「複数語ユニットの知識」と「発話における複数語ユニットの使用」が、発話の流暢性の諸側面とどのような関連性を持つかを調査した。発表語彙としての複数語ユニットの知識を測定するため、今回新たにテストを作成し、既存の発表単一語彙と受容複数語ユニットの知識を測るテストとの関連性を調べた。日本人大学生63名に発話課題として意見を自由に述べるタスクを実施し、3つの発話流暢性の側面(speed, breakdown, and repair fluency)を目的変数として、3種類の語彙テストと、発話をn-gram分析したものを説明変数として予測をした。分析の結果、(1)発表単一語彙と発表複数語ユニットは無声ポーズとの相関を示したが、発表単一語彙の方が予測力は強かった;(2)複数語ユニット使用では低頻度語同士の結びつきの強いn-gram(mutual information)がspeed fluencyとrepair fluencyを予測した;(3)発表単一語彙の知識がbreakdown fluencyを、複数語ユニットの知識と使用がspeedとrepair fluencyを予測した。本研究によって、複数語ユニットと発話流暢性の密接な関連が示された。

1 はじめに

複数語ユニット(Multi-word unit)と発話の流暢性の関連性が指摘されてから久しい(Pawley

& Syder, 1983)。言語(英語)の実に50%は複数語ユニットから成ると言われており(Erman & Warren, 2000)、コロケーション(例: make a mistake)やイディオム(例:raining cats and dogs)などがよく知られている。複数語ユニットは単一語彙と同様に処理されると考えられており(Wray, 2002)、単一語彙と同程度の認知負荷でより多くの発話が可能となるため、流暢性と密接に関係しているとされている(Kormos, 2006)。この複数語ユニットと発話流暢性の関係性は、「発話内での複数語ユニット使用」の側面と「複数語ユニットの知識」の側面に分けて考えることが出来る(Koizumi, 2013)。つまり、実際の発話の中で複数語ユニットを使用すると流暢になるという見方と、複数語ユニットの知識をたくさん持っていることで流暢に話せるという見方である。多くの先行研究は、「発話内での複数語ユニットの使用」と発話流暢性の関係を調査しているが、「複数語ユニットの知識」と発話流暢性の関連を調査している研究は少ない。本研究では、発表語彙としての複数語ユニットの知識を測るテストを、先行研究の複数語ユニットのリストをもとに作成した(Simpson-Vlach & Ellis, 2010)。これにより、発話における複数語ユニットの使用と発表語彙としての複数語ユニットの知識が発話の流暢性をどのように予測するかを調査した。

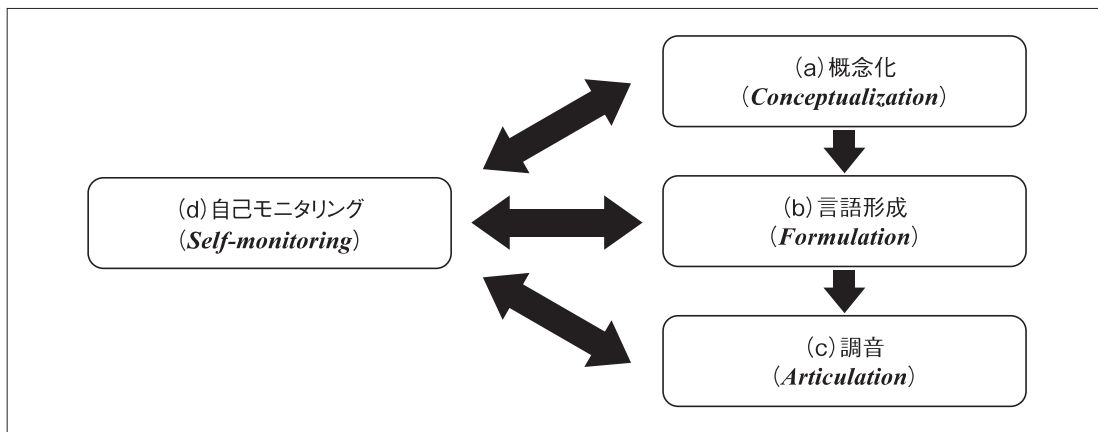
2 先行研究

2.1 発話流暢性の多面性と 発話生成モデル

発話が流暢であるということの定義は長らく流暢性研究で議論的になってきたが、Segalowitz (2010) によれば流暢性は大きく分けて3つの構成概念からなる。1つは発話流暢性 (Utterance fluency) であり、客観的に話す速さなどを数値化出来るものを指す。2つ目は認知流暢性 (Cognitive fluency) であり、発話流暢性を可能にする認知的な効率性 (作業記憶や語彙処理など) を指す。そして3つ目は知覚流暢性 (Perceived fluency) であり、聞き手が話し手の流暢性を評価したものを指す。発話流暢性の測定方法は様々なものが提案されているが (Kormos, 2006)、近年は発話流暢性の多面性を考慮し、次の3つの下位項目ごとに指標を算出する方法がとられている (Tavakoli & Skehan, 2005)。1つ目の speed fluency は話す速度を表し、2つ目の breakdown fluency はポーズの頻度と長さを表

す。ポーズには無声ポーズと有声ポーズがあり、無声ポーズは発声は一切無い沈黙を指す。一方で、有声ポーズは uhm, ahh, hmm など意味を持たない発声を指し、沈黙が長く続かないように間を埋める機能がある (Kormos, 2006)。そして3つ目の repair fluency は同じ発語の繰り返し (例: we can…we can use it) や発語の自発的訂正 (例: if you use…if we watch) などの頻度を表す。ポーズの頻度と長さに関しては、発話生成モデルの観点から (Levelt, 1989)、節の内部におけるポーズと節の外部 (節と節の間) におけるポーズに分けることが出来る (Bosker et al., 2013; Kahng, 2014)。

発話生成モデルは第一言語獲得がモデルのもとになっているが (Levelt, 1989)、第二言語習得にも応用されている (Kormos, 2006; Segalowitz, 2010)。発話生成の構成要素は大きく分けて (a) 概念化、(b) 言語形成、(c) 調音、(d) 自己モニタリングの4つであり、発話生成のプロセスは (a) → (b) → (c) の順番に行われる。ただし習熟度が高まり言語知識が自動化している場合はほぼ同時並行的に行われるとされる (図1を参照)。



■ 図1: 発話生成モデル (Kormos, 2006; Segalowitz, 2010を参考に筆者が作成)

(a) 概念化では、伝える内容や伝達の手段・順番などを決定し、言語化される前の概念 (preverbal message) が形成される。(b) 言語形成では、(a) 概念化で特定された概念に対して、語彙、統語、形態素、音韻のレベルで符号化を行い、当該言語特有の言語形式に変換をする。そして、符号化された概念は (c) 調音にて発声器官から音として産

出される。(d) 自己モニタリングは (a) から (c) のプロセスと同時に働き、それぞれの段階で処理が上手くいっているかを監視しフィードバックをする役割を持つ。

発話の流暢性の3つの側面 (speed, breakdown, and repair fluency) は、発話生成モデルの (a) ~ (d) の段階に紐づけて考えることが出来る。例え

ば repair fluency(同じ発話の繰り返しや発話の自発的訂正)は、間違いに自ら気づき修正をすることによって発生すると考えられるため、(d)の自己モニタリングと関係が深いとされている(Kormos, 1999)。第二言語流暢性研究で特に研究が進んでいるのが、breakdown fluency(無声・有声ポーズ)と発話生成モデルの関連である。ポーズの発生原因は(1)言いたいことが決まらない・分からない場合、と(2)語彙文法などの言語知識の欠如やそれに伴う処理速度の低下によるものの2つに大きく分けることが出来る(Kormos, 2006)。(1)は、発話生成モデルの(a)概念化の段階が上手く処理出来ていないことが原因と考えられ、(2)は発話生成モデルの(b)言語形成の段階において、言語知識の欠如により、言いたいことに対して言語の符号化が上手く処理出来ないことに起因するとされている(Lambert et al., 2017)。第二言語流暢性研究によれば、母語話者と第二言語学習者を比較した際、発話内における節の外部のポーズ数には差がない一方で、節の内部でのポーズ数は第二言語学習者の方が圧倒的に多いことが分かっている(Kahng, 2014, 2018)。このことは、語彙や文法などの言語知識の量や質の違いによるものであるとされている。つまり、母語話者は膨大な言語知識を持ち、かつその処理を無意識化で行っているが、第二言語学習者は母語話者に比べれば限られた言語知識しかなく、かつ言語知識に意識的にアクセスしなければならない場合が多い。以上のことから、節の内部で発生するポーズは、語彙や文法などの言語知識の量やその処理に起因するものであると考えられている(Skehan et al., 2016)。

2.2 発話における複数語ユニット使用と発話流暢性

2.2.1 複数語ユニット

複数語ユニットとは、Wray(2002: p.9)によれば、語の連鎖であり、文法規則によって生成され分析されることがなく、連鎖ごと記憶に保持され引き出されるものである。この定義には諸説あるが、連鎖ごと記憶に保持されている可能性については実証研究が増えてきている(Conklin & Schmitt, 2012)。複数語ユニット(例:in the *middle* of the)と1語のみ変更を加え

た非複数語ユニット(例:in the *front* of the)を提示し、認識速度を測定すると複数語ユニットの方が速く認識・処理されることが分かっている(例:Tremblay et al., 2011)。一方、連鎖ごと記憶から引き出されるかどうかについては実証がほとんどされていないのが現状である(Myles & Cordier, 2017)。しかし、発話生成モデルの観点からも、複数語ユニットは発話の流暢性を高める可能性を秘めている(Kormos, 2006)。

複数語ユニットは定義や同定方法が様々な考察されており(Perkins & Wray, 2000)、母語話者の直感による同定や複数の基準に照らし合わせて同定する方法などは信頼性・妥当性が低いという批判がある(Myles & Cordier, 2017)。一方、自然言語処理やコーパス技術によるn-gramの自動算出と語同士の結びつきの強さの指標を組み合わせた解釈方法は安定性があり(Kyle & Crossley, 2015)、作文能力や発話能力と密接な関係があることが明らかになっている(Eguchi & Kyle, 2020; Siyanova-Chanturia & Spina, 2019)。

N-gramは隣り合う語同士を抽出する方法で、例えばThis is a pen.という文を分析する場合、bi-gram(2語)ではThis is / is a / a pen, tri-gram(3語)ではThis is a / is a penが抽出される。抽出されたn-gramはCorpus of Contemporary American English (COCA, Davies, 2008)などの大規模コーパスを参照して、どのくらい洗練されているかを語彙の洗練性の指標によって数値化する。洗練性の指標にはn-gramの頻度を表すfrequencyの他に、語同士の結びつきの強さを表すt-scoreとmutual information(相互情報量)がある。T-scoreはお互いの語が高頻度のn-gram(例:the + man)で高くなる一方で、mutual informationはお互いの語が低頻度のn-gramで高くなり(例:okey + dokey)、学習者の習熟度や発話の分かりやすさとの関連が報告されている(Durrant & Schmitt, 2009; Gablasova et al., 2017; Saito, 2020)。

2.2.2. N-gram使用と発話流暢性

複数語ユニットと発話流暢性の関係についての調査は近年増えてきているが、n-gram分析を使用した調査はTavakoli and Uchihara (2019)

のみである。Tavakoli and Uchihara (2019) はイギリスの大学入学志願者56名に対してスピーキングテスト (Test of English for Educational Purposes) を行い、発話流暢性と n-gram (bi-gram, tri-gram) の頻度と結びつきの強さに関連があるかを調査した。結果は、高頻度の tri-gram と調音速度 (articulation rate) に正の相関が見られ ($r = .396$)、頻出する tri-gram を多く使っている人は話す速度が速くなる傾向があることが分かった。また、mutual information と節外の無声ポーズ数に負の相関が見られ、低頻度語同士のより洗練された n-gram を使用することで、発話の概念化に無声ポーズの時間をより割くようになる傾向が明らかになった。

2.3 語彙知識と発話流暢性

語彙知識には発表語彙と受容語彙があり、発表語彙は受容語彙に比べてより言語の産出 (話す、書く) に関連が深いとされている (Fitzpatrick & Clenton, 2017)。発話運用能力の中でも発話の流暢性は、言語の知識量や言語の処理の効率性に直結しているとされ (Segalowitz, 2016)、発表語彙の知識量との関連性が調査されてきた。Uchihara and Saito (2019) は39名の日本人大学生に対して発表語彙のテスト (Lex30: Meara & Fitzpatrick, 2000) を行い、絵描写タスクでスピーキングテストを行った。5名の英語母語話者に録音した発話を聞かせて0から1000のスケールで流暢性を判断させた。相関分析の結果流暢性と発表語彙テストのスコアは $r = .342$ で有意な相関を示した。また、Uchihara et al. (2021) は日本人大学生40名に同じく発表語彙のテスト (Lex30) を行い、発話課題として絵描写タスクを行い、流暢性の指標として調音速度、有声ポーズ数、無声ポーズ数を算出した。結果は、発表語彙テストのスコアが調音速度と無声ポーズ数に対して有意な相関関係を示した (それぞれ $r = .48, -.43$)。これらの研究は単一語彙のみをテストとして使用している一方、De Jong et al. (2013) と Koizumi and In' nami (2013) は複数語彙ユニットも含めてテストをしている。

De Jong et al. (2013) はオランダ語学習者179名に発表語彙のテストとしてオランダ語の発表

単一語彙に加え、発表コロケーション (2語) のテストを実施し、2つのテストを「語彙知識」として合計した上で発話流暢性指標との関係を調査した。分析の結果、語彙知識は無声ポーズ数と有声ポーズ数を有意に予測した。Koizumi and In' nami (2013) は日本人高校生224名に語彙テスト (発表コロケーション (2語) 含む) を実施し、発話流暢性指標 (speed fluency and repair fluency) との関連を調査した。分析の結果、コロケーションは speed fluency の17% を予測し (語彙知識全体で32%)、repair fluency の7% を予測した (語彙知識全体で17%)。

3 研究の目的

本研究は3つの研究目的からなる。1つ目は、発表複数語ユニット知識と発話流暢性の諸側面の関連性を調査することである。De Jong et al. (2013) と Koizumi and In' nami (2013) は発話流暢性との関連で2語のコロケーションをテストで使用したが、Tavakoli and Uchihara (2019) が指摘するように、発話流暢性との関連では4語や5語の複数語ユニットも考慮して調査をする必要がある。よって2語以上の複数語ユニットの知識を測るテストが必要であったが、既存のテストでそのような知識を測るものが存在しない。したがって、本研究では Simpson-Vlach and Ellis (2010) の複数語ユニットのリストを参考に、発表複数語ユニットの知識を測るテストを作成した (4.3語彙知識測定を参照)。発表複数語ユニット独自の発話流暢性に対する予測力を検証するため、「複数語 vs 単一語」「発表 vs 受容」を比較対象として、発表単一語彙と受容複数語ユニットを回帰分析に含めることで影響力を統制した。

2つ目は、発話で使用される複数語ユニットと発話流暢性の諸側面の関連性を調査することである。先行研究では、発話データを n-gram で分析し、発話流暢性との関連を調査したものは Tavakoli and Uchihara (2019) のみである。発話データにおける複数語ユニットを自動で同定する n-gram 分析は、流暢性との関連で今後さらなる検証が必要である。本研究は、Tavakoli and

Uchihara (2019) よりも多くのサンプル数を集めた上に、日本という外国語環境で調査を行っており、新たな実証として分野への貢献が期待される。

3つ目は、発表複数語ユニットの知識と発話での複数語ユニット使用のどちらが発話流暢性の諸側面とより関連があるかを調査することである。知識として持っている語彙力と実際の発話で使用できる語彙力のどちらが発話流暢性にとって重要なかは、研究としてだけでなく学習にも示唆を与える。しかし、筆者が知る限りでは、知識と使用の影響を同時に見た研究は存在しない。発表複数語ユニットの知識の独自の影響を調査するため、発表単一語彙と受容複数語ユニットも分析に含める。

研究課題1

発表複数語ユニットの知識は発話流暢性のどの側面に対してどのような関連を持つか？また、発表単一語彙と受容複数語ユニットの発話流暢性に対する影響力は、発表複数語ユニットの知識と比べた場合どのようなものか？

研究課題2

発話における複数語ユニットの使用は、発話流暢性のどの側面に対してどのような関連を持つか？

研究課題3

発表複数語ユニットの知識と発話における複数語ユニットの使用は、発表単一語彙と受容複数語ユニットの発話流暢性に対する影響力を考慮した場合に、どちらがより発話流暢性の諸側面に対して関連性を持つか？

4 方法

4.1 参加者

都内の私立大学に通う日本人大学生63名(学部学科は様々)が本実験に参加し、63名全員が実験内のタスクを遂行した。学年は1年生から5年生以上まで様々で、18名が英語圏に1ヵ月以上滞在歴があった(最大5年間)。また、5名が大学入学以前に日本国内において英語によるイマージョン教育(注1)を経験していた。参加者の習熟度は、過去2年以内に受験した英語資格試験(英検、TOEIC Reading & Listening、TOEFL ITP、TOEFL iBT、IELTS)の結果を質問紙にて収集し、Common European Framework of Reference (CEFR, Council of Europe, 2001) に従い、6段階(A1~C2)のレベルに変換を行った。なおスコアの変換にはETS (Educational Testing Service) の公式サイト(注2)と文部科学省が公表している「各資格・検定試験とCEFRとの照合表」(注3)を参照した。63名のうち質問紙に回答した46名は概ねB1からB2レベルであった。なお、サンプルサイズはG*Power (Erdfeulder et al., 2009) を使用し、重回帰分析において説明変数の数が5個での検定力が0.80になるように計算した結果を基に決定した。

4.2 発話課題

発話能力を調査する先行研究の多くは絵描写タスクを使用しているが(Derwing et al., 2004)、現実世界では絵を描写する機会が実際あまりなく、意見を自由に述べるタスクの方がより現実的であるという指摘がある(Suzuki & Kormos, 2019)。よって本研究で使用するタスクは、賛成か反対かを示した上で意見を自由に述べるタスクを使用した(トピックは次頁を参照)。

注1 : 使用言語が英語で科目の内容を学ぶ教授法の総称。イマージョン教育、内容言語統合型学習、インターナショナルスクール、国際バカロレアが該当。なお英語で行われる英語科目は含まないものとする。

注2 : <https://www.ets.org/> (2021年5月18日)

注3 : https://www.mext.go.jp/b_menu/houdou/30/03/_icsFiles/afiefieldfile/2019/01/15/1402610_1.pdf (2021年5月18日)

(トピック)

Do you agree with the following statement?:

“Smartphones have more negative than positive effects on our everyday lives”

Please explain why you think so and include specific details or experiences supporting your opinion.

4.3 語彙知識測定

4.3.1. Productive Vocabulary Levels Test (以降PVLТ)

学習者の「発表単一語彙知識」の広さを測定するために、Laufer and Nation(1999)のVersion Bのうち2000語、3000語、5000語レベルを実施した。それぞれのレベルごとに18問ずつあり、短

い英文の中に単語を回答する箇所が空欄で埋め込まれている。参加者は答えとなる単語の最初の数文字がヒントとして与えられた(以下の例題を参照)。採点基準は先行研究(De Jong et al., 2012)に倣い、単語自体を知っているかどうかのみに焦点をあて、文法活用(時制、三単現のs、単数複数形など)は考慮しなかった。点数は正解の場合は1点、不正解の場合は0点とした。

(例題)

1. It is the de ____ that counts, not the thought. (答え) deed
2. Plants receive water from the soil through their ro _____. (答え) roots / root

4.3.2. Phrasal Vocabulary Size Test (以降PVST)

学習者の「受容複数語ユニット知識」の広さを測定するために使用した(Martinez & Schmitt,

2012)。1000語レベルから5000語レベルまでそれぞれ8問ずつあり、例文中で太字になっている複数語ユニットの意味を表すものを4つの選択肢から選んでもらった(以下例題を参照)。1問1点で50点満点とした。

(例題)

1. GO ON: It will go on.
a. ○sleep
b. ○repeat
c. ○be fast
d. ○continue
(答え)d. continue
2. LEAD TO: No one knows what it will lead to.
a. ○want
b. ○have inside
c. ○cause in the future
d. ○find
(答え)c. cause in the future

4.3.3. Productive Academic Formulas Test (以降PAFT)

学習者の「発表複数語ユニット知識」の広さを測定するため、Simpson-Vlach and Ellis(2010)によるAcademic Formulas List(AFL)をもとに筆者自身がテストを作成した。AFLはBritish National CorpusやMichigan Corpus of Academic Spoken Englishなどの学術的話し言葉・書き言葉コーパスをもとに作成されており、Spoken, Written, Coreの3つのセクションに分かれている。項目が機能(意見陳述など)ごとに類別されている点に加え、教師が評価した指標によって「教える価値」が高い順に項目がリスト化されており、学習者や教師にとって有益なリストとなっている。

今回使用したPAFTは、様々な選定基準を設けた上で10段階の選定プロセスを経ることにより、AFLの607項目から35項目を厳選した(テスト作成手続きは資料1を参照)。テストの信頼性はクロンバックの α 係数で調べ、PVLТ, PVSTとの相関係数を確認することで既存のテストとの関連性を確認し、テストの内容の妥当性を確保した。18名の日本人大学生に対して予備調査を行った結果、クロンバックの α 係数は $\alpha = .846$ であり、PVLТ, PVSTとも強い相関を示した(それぞれ $r = .685, .720$)。本調査では63名の回答を再度信頼性分析にかけ、予備調査と同様にクロンバックの α 係数とPVLТ, PVSTとの相関を確認した。技術的な問題で使用不可となった1項目を分析から除外した結果、クロンバックの α 係数は $\alpha = .880$ であり、PVLТ, PVSTと強い相関を示した(それぞれ $\rho = .678, .687$: PAFTとPVLТが正規分布していないためスピアマンの相関係数を使用)。テストのさらなる信頼性確保のため、34項目を項目分析にかけ、項目を削除した場合の α 係数と項目間の相関を考慮したうえでさらに3項目を削除した。その結果、クロンバックの α 係数は $\alpha = .884$ とさらに高い信頼性を示し、PVLТ, PVSTとの相関もさらに強くなった(それぞれ $\rho = .681, .698$)。信頼性分析は以上とし、合計31項目で点数化した。

なお、項目を削除することによって構成概念の代表性を損なってしまう危険性については、削除した3項目はAFL(Simpson-Vlach & Ellis, 2010)の3つの機能(referential expressions, stance expressions, discourse organizing functions)に属しているが、残存項目でこの3機能は十分カバーされているため、問題ないと判断した。

PVLТと同様、例文の中に回答箇所が空欄で埋め込まれており、複数語のそれぞれの語頭にヒントとして数文字を与えた(テストは資料2と3を参照)。例文作成にあたり、*Collins Dictionary, Cambridge Dictionary, Oxford*

*Learner's Dictionary, Longman Dictionary of Contemporary English*を参照した。また、例文内で使用される語彙のレベルは、General Service List(West, 1953)とBNC/COCA headword lists(Nation, 2017)の2000語レベル以下になるように統制した。テストの項目のレベルの適切さ、ヒントとして与えている語頭の数字の適切さ、例文の適切さについては、第二言語習得論に精通している3名の大学院生(うち1名は米国在住の日英バイリンガル)に確認をしてもらい、問題ないことがわかった。

採点基準として、複数語ユニットの特性上語形の変化が許されない箇所(○on the other hand → × on the other hands)は厳しく採点した。それ以外はPVLТと同様、文法活用の誤り(○come up with → ○coming up with)は許容範囲内とし、正解は1点、不正解は0点とした。

4.4 発話における複数語ユニット使用測定

本研究では、複数語ユニットをn-gramと定義し、bi-gram(2語)とtri-gram(3語)を分析対象とする。以下はbi-gram、tri-gramそれぞれをスコア化する4つの異なる指標である。指標の算出はTools for Automatic Analysis of Lexical Sophistication (TAALES, Kyle & Crossley, 2015)を使用して行い、全ての指標の参照先のコーパスは、日本の英語学習環境がアメリカ英語に影響を受けていることを鑑み、COCA(Davies, 2008)に設定した(Saito, 2020)。また、本研究のデータが発話データであることからCOCAのspoken subsectionを設定した(Tavakoli & Uchihara, 2019)。

4.4.1. Logarithmized frequency

複数語ユニットがどのくらいの頻度で使われるかを表す指標で、発話データにおけるbi-gram/tri-gramのCOCAでの頻度数を合計することで算出した。頻度数はZipfian effectに

(式)

$$T\text{-score} = (\text{複数語ユニットの実際の頻度数} - \text{語同士の偶然共起する頻度数}) \div \sqrt{\text{実際の頻度数}}$$

より非正規分布化してしまうので、ログ変換 (logarithmized) を行ったものを算出した。

4.4.2. Range

複数語ユニットがどのくらい幅広く使われるかを表す指標で、発話データにおけるbi-gram/tri-gramが使われるCOCA内ドキュメント数を合計することで算出した。

4.4.3. T-score

複数語ユニットの語同士の結びつきの強さを表す指標で、発話データにおけるbi-gram/tri-gramのCOCAでの実際の頻度数から、語同士が偶然共起する頻度数を引き、実際の頻度数の平方根で割ることで算出した(前述式を参照)。それ

ぞれの語に高頻度語が使われているほどスコアが高くなりやすいという性質を持つ(Gablasova et al., 2017)。

4.4.4. Mutual information

複数語ユニットの語同士の結びつきの強さを表す指標で、発話データにおけるbi-gram/tri-gramのCOCAでの実際の頻度数を語同士が偶然共起する頻度数で割り、それを2を底とする対数でとることで算出した(以下式を参照)。それぞれの語に低頻度語が使われているほどスコアが高くなりやすく、それぞれの語がお互いの語以外の語との結びつきが弱いという性質を持つ(Gablasova et al., 2017)。

(式)

$$\text{Mutual information} = \log_2(\text{複数語ユニットの実際の頻度数} \div \text{語同士の偶然共起する頻度数})$$

4.5 発話流暢性

指標を算出するにあたり、音響分析ソフトのPraat(Boersma & Weenink, 2020)を使用し、De Jong et al.(2020)によるscriptを実行し自動で無声ポーズを特定した。ポーズの定義は先行研究に倣い(Bosker et al., 2013), 0.25秒以上とした。発話データは全て手動で文字に起こし、無声ポーズの位置を細かく特定するため、テキストファイル上でAnalysis of Speech Unit (AS-unit: Foster et al., 2000)ごとにアノテーションを行った。AS-unitの評定者間信頼性については、流暢性研究に精通している大学院生一人に発話データ全体の20%を分析してもらい、69%の一致度が得られた。その後一致度が100%になるように話し合いを行った上で、残りの80%を筆者が分析した。Praat上にて流暢性の指標を手動でアノテーションした後、R(R Core Team, 2020)によって全ての流暢性指標を自動で算出した。なお、指標の標準化のために使用する総音節数は、有声ポーズと繰り返し・自己訂正・擬発音を除いたpruned transcriptionにて算出した。以下は発話流暢性の3つの側面(speed,

breakdown, and repair fluency)ごとの指標の算出方法である(指標と算出方法のまとめは表1を参照)。

4.5.1. Speed fluency(1指標)

調音速度は、総発話時間から無声ポーズの時間を引いた発声時間によって総音節(シラブル)数を割ることで算出した。

4.5.2. Breakdown fluency(5指標)

大きく分けて無声ポーズと有声ポーズに分けることが出来る。無声ポーズの指標は近年の流暢性研究の動向を踏まえて(Bosker et al., 2013; Kahng, 2014), ポーズの頻度, ポーズの長さ, ポーズの位置の3つを踏まえて算出した。なお, ポーズの位置の算出には節(clause)を特定する必要があり, Foster et al.(2000)に倣って節を定義した。

節内無声ポーズ割合は, 「節内」の無声ポーズの「数」を表す指標で, 総節内無声ポーズ数を総音節数で割ることで算出した。

節外無声ポーズ割合は, 「節外」の無声ポーズの「数」を表す指標で, 総節内無声ポーズ数を総

表1: 発話流暢性指標と算出方法のまとめ

	発話流暢性指標	算出方法
<i>Speed fluency</i>		
	調音速度	総音節数 ÷ 総発声時間
<i>Breakdown fluency</i>		
頻度	節内無声ポーズ割合	総節内無声ポーズ数 ÷ 総音節数
	節外無声ポーズ割合	総節外無声ポーズ数 ÷ 総音節数
	有声ポーズ割合	総有声ポーズ数 ÷ 総音節数
時間	節内無声ポーズ時間	総節内無声ポーズ時間 ÷ 総節内無声ポーズ数
	節外無声ポーズ時間	総節外無声ポーズ時間 ÷ 総節外無声ポーズ数
<i>Repair fluency</i>		
	繰返し割合	総繰返し数 ÷ 総音節数
	自己訂正割合	総訂正数 ÷ 総音節数
	擬発言割合	総擬発言数 ÷ 総音節数

注. 調音速度以外の指標は数値が大きくなるほど非流暢であることを示す

音節数で割ることで算出した。

節内無声ポーズ時間は、「節内」の無声ポーズの「長さ」を表す指標で、総節内無声ポーズ時間(秒)を総節内無声ポーズ数で割ることで算出した。

節外無声ポーズ時間は、「節外」の無声ポーズの「長さ」を表す指標で、総節外無声ポーズ時間(秒)を総節外無声ポーズ数で割ることで算出した。

有声ポーズ割合は、有声ポーズ(例:hmm, ah)の数を表す指標で、総有声ポーズ数を総音節数で割ることで算出した。連続して有声ポーズをした場合はそれぞれを数に数えた(例:hmm…ahh…ehhで3つ)。また、有声ポーズか否かを判断する基準は、その発語が発話の内容にとって意味のあるものかどうか、とした。例えば、言葉に詰まった時の「sorry」や、間を埋めるために使われる「well」などは話している内容(今回は携帯電話について)と直接関係がないため有声ポーズとして数えた。

4.5.3. Repair fluency(3指標)

繰返し割合は、同じ発話を繰り返した数を表す指標で、総繰返し数を総音節数で割ることで算出した。

自己訂正割合は、言い間違いを自ら訂正した数を表す指標で、総訂正数を総音節数で割ることで算出した。

擬発言割合は、節の始まり時点で言い間違い

た発話を自ら撤回し新たな発話を始めた数を表す指標で、総擬発言数を総音節数で割ることで算出した。

4.6 手順

本研究はコロナ禍により実験を対面で実施することが出来なかったため、すべてZoomを使用してオンラインで行った。スピーキングの録音はPraat(Boersma & Weenink, 2020)を使用するため、参加者には事前にPraatを自身のPCにダウンロードしておくように指示をした。また、3つの語彙テストと質問紙はGoogle Formsを使用して作成をし、回答してもらった。PVLTとPVSTはもともとのフォーマットの通りになるように注意して作成した。実験は、「発話課題→5分休憩→PVLT→PAFT→5分休憩→PVST→質問紙」の順番に行った。

参加者はまず実験の流れを確認した後、Praatの操作方法を含めて発話課題の流れを筆者と一緒に確認した。参加者には、発話課題の説明・注意書き・トピックが書かれたWord fileをZoom上でスクリーンシェアしたのを見てもらった。筆者が書かれている説明を英語で読み上げ、参加者に一緒に確認をしてもらった。説明が終了した後、参加者は2分間の準備時間が与えられ、時間は筆者がiPhoneのタイマーをセットし、画

面上に表示をした。準備が終了した後、Praatの録音開始ボタンを押してもらい、スクリーンシェアされたトピックが書かれているWord fileを見ながらスピーキングを行った。話す時間の制限は特に指定しなかった(発話時間と発話語数の記述統計量は資料4を参照)。スピーキングが終了した後、Praatの録音停止ボタンを押してもらい参加者自身のPCに録音した音声ファイルをダウンロードしてもらった。その後Zoom上のChat機能を使用し、ダウンロードした音声ファイルを送信してもらった。

参加者には、本番を行う前に練習用の発話課題を使用してオンラインでのスピーキングに慣れてもらった。練習用には自己紹介のトピックを使用し、30秒の準備時間の後、制限時間なしでスピーキングを行った。練習では、本番と同じように手続きを行うことでその後スムーズに本番を行うことが出来た。練習の時点で、参加者側の環境(騒音がないか、気が散らないかなど)を確認した。また、練習で録音した音声ファイルを聞いて音質を確認することで、本番での録音方法(直接PCに話しかけるか、マイク付きのイヤホンセットするか、PCの代わりにiPhoneなどの携帯端末を使用して録音するかなど)に不安がないようにした。

4.7 分析

語彙知識、複数語ユニットの使用、そして発話流暢性の諸側面の関係性をモデル化するため、まず変数間の相関関係を確認した。その後、語彙知識と複数語ユニットの使用を説明変数に、発話流暢性の諸側面をそれぞれ目的変数としてステップワイズ法による重回帰分析を繰り返し行った。語彙知識については、PAFT独自の影響を見るため、PVLTとPVSTも説明変数に組み込んだ。発話における複数語ユニットの使用は全部で8変数あるため(bi-gramとtri-gramのfrequency, range, t-score, mutual information)、分析結果の解釈を容易にするためにTavakoli and Uchihara (2019)に倣って主成分分析を行い、変数をより少ない成分に要約した。重回帰分析を行うにあたり、Larson-Hall(2010)に基づき、残差の正規性、残差の等分散性、多重共線性を確認し、全て

前提を満たしていることを確認した。

5 結果

5.1 予備結果

資料5~7に語彙知識、複数語ユニットの使用(変数集約前)、発話流暢性の記述統計量を示す。複数語ユニット使用の指標であるbi-gramを除いたほとんどの変数が正規分布に従っていないことが分かる。発話流暢性指標のうち特にbreakdown fluencyとrepair fluencyは正規分布しないことが報告されている(Lambert et al., in press)。また、n-gramのうちtri-gramは、分析対象から外される項目が多いために分布が偏ったと考えられる(表8のn/aを参照)。

5.2 N-gramの主成分分析結果

N-gramの8変数は、固有値(1に設定)とスクリープロットの結果、3つの因子に要約するのが適切であったため、3因子の因子負荷量を0.6以上に設定しプロマックス回転を行って主成分分析を実施した。前提条件の確認として、サンプルサイズの妥当性を検証するためKaiser-Meyer-Olkinの標本妥当性の測度を実施し、KMO = .073の十分な結果を得た(.07以上が良いとされる: Field, 2009)。また観測変数間の相関を調べるためBartlettの球面性検定を行い、結果が有意であったことから($p < .0001$)今回の分析の妥当性が確認できた(Field, 2009)(分析結果は表2を参照)。分析の結果、3因子が特定され、累積因子寄与率は89.4%であった。また因子間の相関も見られた。1つ目の因子はtri-gram t-score, range, frequencyが高い因子負荷量を示しているため「高頻度3語」という因子名をつけた。2つ目の因子はbi-gram frequency, range, t-scoreが高い因子負荷量を示しているため「高頻度2語」という因子名をつけた。3つ目の因子はbi-gram mutual information, tri-gram mutual informationが高い因子負荷量を示しているため「相互情報量因子」という因子名をつけた。

表2: N-gram指標に対する主成分分析結果と因子相関行列

N-gram measures	第1因子 (高頻度3語)	第2因子 (高頻度2語)	第3因子 (相互情報量因子)
Tri-gram t-score	0.953	0.068	-0.044
Tri-gram range	0.947	0.090	-0.066
Tri-gram frequency	0.832	0.257	-0.060
Bi-gram frequency	0.041	0.927	0.003
Bi-gram range	0.136	0.897	0.021
Bi-gram t-score	0.156	0.850	0.133
Bi-gram mutual information	-0.308	0.275	0.917
Tri-gram mutual information	0.441	-0.342	0.638
因子相関行列			
高頻度3語	1.000		
高頻度2語	0.506	1.000	
相互情報量因子	0.157	0.329	1.000

5.3 語彙知識、複数語ユニット使用、 発話流暢性の相関分析結果

表3に語彙知識と発話流暢性の相関分析の結果、および複数語ユニット使用の指標と発話流暢性の相関分析の結果を示す(全変数の相関行列

表は資料8を参照)。語彙知識は概ねbreakdown fluencyのうち無声ポーズとの間に有意な小～中程度の相関関係が見られる。特にPAFTは節内と節外無声ポーズの数と長さすべてと小～中程度($r_s=.25$ は小, $.40$ は中程度:Plonsky & Oswald,

表3: 語彙知識・複数語ユニット使用と流暢性指標のスピアマン相関分析結果

	Speed fluency	Breakdown fluency				有声ポーズ	Repair fluency		
		無声ポーズ					繰返し	自己訂正	擬発言
		頻度		時間					
調音速度	節内	節外	節内	節外					
語彙知識									
PVLT	.151 [-.101, .384]	-.228 [-.450, .021]	-.360* [-.558, -.124]	-.322* [-.527, -.080]	-.520*** [-.680, -.312]	.005 [-.243, .253]	-.047 [-.292, .203]	-.130 [-.366, .121]	-.040 [-.285, .210]
PVST	.253* [.005, .471]	-.308* [-.516, -.065]	-.234 [-.455, .015]	-.239 [-.460, .009]	-.102 [-.341, .149]	-.055 [-.298, .217]	-.208 [-.433, .042]	-.312* [-.520, -.070]	-.053 [-.297, .197]
PAFT	.202 [-.048, .429]	-.276* [-.490, -.030]	-.261* [-.478, -.014]	-.295* [-.506, -.051]	-.321* [-.527, -.080]	-.010 [-.257, .238]	-.188 [-.416, .063]	-.226 [-.449, .023]	-.007 [-.254, .241]
複数語ユニット使用									
高頻度2語	.286* [.041, .498]	-.246 [-.465, .002]	-.024 [-.270, .225]	-.233 [-.455, .015]	-.071 [-.313, .180]	-.208 [-.433, .042]	-.298* [-.508, -.054]	-.207 [-.433, .042]	-.255* [-.473, .008]
高頻度3語	.307* [.065, .516]	-.102 [-.341, .149]	-.031 [-.277, .219]	-.083 [-.324, .283]	.038 [-.212, .283]	-.205 [-.431, .045]	-.246 [-.465, .002]	-.175 [-.405, .076]	-.401* [-.590, -.170]
相互情報量	.273* [.027, .488]	-.241 [-.461, .008]	-.040 [-.285, .210]	-.063 [-.306, .188]	.088 [-.164, .328]	-.324** [-.529, -.083]	-.225 [-.448, .024]	-.170 [-.401, .081]	-.246 [-.465, .002]

注. * $p < .05$, ** $p < .01$, *** $p < .001$: []は95%信頼区間を表す

2014)の相関関係を示した。複数語ユニット使用は、語彙知識とは対照的にbreakdown fluencyとの相関はなく、speed fluencyにすべての指標が有意な小～中程度の相関関係を示した。また、repair fluencyとは高頻度2語と高頻度3語が小～中程度の相関関係を示した。

5.4 Speed fluencyを予測する重回帰モデル

発話流暢性のspeed fluencyを語彙知識と複数語ユニット使用で予測した(表4)。調音速度に対

して相互情報量因子が8.7%を予測し、PVSTが追加の9.5%を予測、あわせて18.2%を予測した。PAFTは説明変数には入らなかったが、PVSTが予測していることから、発表知識よりも受容知識、単一語彙知識よりも複数語ユニットの知識が発話の速さを予測することが分かった(研究課題1)。複数語ユニット使用については相互情報量因子が重要な説明変数であった(研究課題2)。結果の重回帰モデルから、発話の速さに対して知識(PVST)と使用(相互情報量因子)の両側面から予測をしたことが明らかになった(研究課題3)。

■表4: Speed fluencyを予測する回帰モデル

目的変数	予測変数	R^2	Adjusted R^2	ΔR^2	β	ΔF	P
調音速度	相互情報量	0.102	0.087	0.087	0.32	6.90	0.011
	相互情報量 + PVST	0.208	0.182	0.095	0.33	8.10	0.006

注. ΔR^2 はadjusted R^2 による

5.5 Breakdown fluencyを予測する重回帰モデル

発話流暢性のbreakdown fluencyを語彙知識と複数語ユニット使用で予測した(表5)。(1)節内無声ポーズ割合に対して相互情報量因子が12.5%を予測し、PVLTが追加の12.6%を予測、併せて25.1%を予測した。(2)節外無声ポーズ割合に対してPVLTが13.1%を予測し、相互情報量因子が追加の8.4%を予測、あわせて21.5%を予測した。(3)節内無声ポーズ時間に対してはPVLTのみが7.3%を予測した。(4)節外無声ポーズ時間に対してはPVLTが22.8%, PVSTが追加で4.1%を予測し、あわせて26.9%を予測した。最

後に(5)有声ポーズ割合に対しては相互情報量因子のみが13.7%を予測した。

Breakdown fluencyのすべての指標に対してPAFTは予測をしなかった。また、(5)有声ポーズ割合は予測しなかったものの、PVLTが概ね重要な説明変数であった(研究課題1)。複数語ユニット使用では、高頻度語同士よりも低頻度語同士のn-gram(相互情報量因子)が(1), (2), (5)において重要な説明変数であり、無声ポーズと有声ポーズの「数(頻度)」を予測した(研究課題2)。Breakdown fluencyに対しては、おもに知識(おもにPVLT)と使用(おもに相互情報量)が予測をしたことが明らかとなった(研究課題3)。

■表5: Breakdown fluencyを予測する回帰モデル

目的変数	予測変数	R^2	Adjusted R^2	ΔR^2	β	ΔF	P
(1) 節内無声ポーズ割合	相互情報量	0.139	0.125	0.125	-0.37	9.83	0.003
	相互情報量+PVLT	0.275	0.251	0.126	-0.37	11.26	0.001
(2) 節外無声ポーズ割合	PVLT	0.145	0.131	0.131	-0.38	10.36	0.002
	PVLT+相互情報量	0.241	0.215	0.084	-0.31	7.56	0.008
(3) 節内無声ポーズ時間	PVLT	0.088	0.073	0.073	-0.30	5.85	0.019
(4) 節外無声ポーズ時間	PVLT	0.241	0.228	0.228	-0.49	19.32	<.001
	PVLT+PVST	0.293	0.269	0.041	0.29	4.45	0.039
(5) 有声ポーズ割合	相互情報量	0.151	0.137	0.137	-0.39	10.85	0.002

注. ΔR^2 はadjusted R^2 による

5.6 Repair fluencyを予測する重回帰モデル

発話流暢性のrepair fluencyを語彙知識と複数語ユニット使用で予測した(表6)。繰返し割合に対して相互情報量因子が11.6%を、PAFTが追加の8.6%を予測し、あわせて20.2%を予測した。自己訂正割合に対してはPVSTが8.6%を予測し、擬発言割合に対しては高頻度3語が6.9%を予測した。PAFTは繰返しの回数に対して予測を示した。また、複数語ユニット(PAFT, PVST)が説明変数に入っていることから、repair fluencyに

とって単一語彙よりも複数語ユニットの方が重要な要素であることが分かった(研究課題1)。複数語ユニット使用は、低頻度語同士の洗練されたn-gram(相互情報量因子)を多く使えると同じ発話を繰返す回数が減り($\beta = -0.36$)、高頻度のtri-gramを多く使うことで間違っただ話を始めて訂正する回数が減る($\beta = -0.29$)ことが分かった(研究課題2)。Repair fluencyに対しては知識と使用の両側面から予測をしたことが明らかになった(研究課題3)。

■表6: Repair fluencyを予測する回帰モデル

目的変数	予測変数	R^2	Adjusted R^2	ΔR^2	β	ΔF	P
繰返し割合	相互情報量	0.130	0.116	0.116	-0.36	9.15	0.004
	相互情報量+PAFT	0.228	0.202	0.086	-0.32	7.54	0.008
自己訂正割合	PVST	0.101	0.086	0.086	-0.32	6.86	0.011
擬発言割合	高頻度3語	0.084	0.069	0.069	-0.29	5.57	0.021

注. ΔR^2 はadjusted R^2 による

6 考察

本研究では、複数語ユニットの知識と発話での複数語ユニットの使用が発話流暢性の諸側面をどのように予測するかを調査することが目的であった。発表複数語ユニット(2語以上)の知識を測定するテストが存在しないため(De Jong et al., 2013; Koizumi & In'nami, 2013), 今回PAFTとして新たに作成をした。PAFT独自の予測を検証するため、PVLとPVSTの予測力を同時に検証した。本研究が取り組んだ課題は(1) PAFTは発話流暢性のどの側面を予測するか。またPVL, PVSTとの関連性はどのようなものか; (2) 複数語ユニットの使用(n-gram)は発話流暢性のどの側面を予測するか;そして、(3) PAFTと複数語ユニット使用(n-gram)は、どちらがより発話流暢性の諸側面に対して関連があるか。またPVL, PVSTとの関連性はどのようなものか、の3つである。

6.1 課題1

発話流暢性の3つの側面(speed, breakdown, and repair fluency)のうち、PAFTはbreakdown fluencyの節内と節外無声ポーズの頻度と長さ両方に有意な相関を示し(表3), repair fluencyの繰返しの頻度を予測した(表6)。先行研究では、発表語彙知識はspeed fluency(De Jong et al., 2013; Koizumi & In'nami, 2013; Uchihara et al., 2021), breakdown fluency(De Jong et al., 2013; Uchihara et al., 2021), repair fluency(De Jong et al., 2013; Koizumi & In'nami, 2013)と関連があり、今回調査した2語以上からなるPAFTもおおむね同じような結果となった。

今回の分析で明らかになったことが全部で3点ある。1点目は、発表単一語彙知識を測定するPVLと2語以上の発表複数語知識を測定するPAFTが、発話流暢性に対して似たようなパターンの結果を示したことである。表3の相関分析において、PVLとPAFTが共通してbreakdown fluencyのうち節内無声ポーズ割合を除くすべての指標と小~中程度で相関した。一方、「受容知識」を測定するPVSTはbreakdown fluency

のうち節内無声ポーズ割合のみと相関している。このことから、PVLТとPAFTは共通してbreakdown fluencyと関連が深い「発表語彙」として捉え、解釈できる可能性が示唆された。発表語彙知識の方が受容語彙知識よりもスピーキングにおける無声ポーズと関連が深いという結果は、語彙研究並びに流暢性研究に大きな示唆を与えうる。これまでに1つの研究内において発表語彙知識および受容語彙知識を発話流暢性との関連で調査した研究は筆者が知る限りKoizumi and In'nami (2013)のみであるが、Koizumi and In'nami (2013)はbreakdown fluencyを目的変数として調査していない。Foster (2020)が述べているように、発表語彙知識は即興的発話をする際に引き出されると考えられるため、処理速度が鍵となる流暢性との関連が疑われてきた。実際に、De Jong et al. (2013)では発表単一語知識と発表複数語知識を測るテストの合計点が、無声および有声ポーズ数と相関している。今回の結果によって、発表語彙知識は単一語であれ複数語であれ、節内節外のポーズの長さ、および節外のポーズ数と関連があることが明らかとなった。

2点目は、複数語の知識を測定するPVSTとPAFTが節内の無声ポーズ数と相関している点である。節内の無声ポーズ数は聞き手が流暢性を評価する際にもっとも重要な指標である(Suzuki et al., 2021)。先行研究では、複数語の知識を測定するテスト(PVST)を単体で実施し、発話流暢性と関連を調査した研究はKahng (2020)のみであり、PVSTと節内の無声ポーズ数の相関が報告されている($r = -.32, p < .05$)。今回の結果では、単一語彙知識を測るPVLТは節内の無声ポーズ数とは相関していないことから、複数語ユニットの知識を持っていることが節内の無声ポーズ数にとって重要である可能性が確認できた。節内の無声ポーズは発話生成モデルの言語形成の処理段階(図1参照)と密接に関連しており(Kahng, 2014, 2018)、複数語ユニットの知識を多く持っていることで、語の連鎖全体で処理されることにより(Wray, 2002)、複雑な符号化処理の負荷を減少させることが出来たと考えられる(Kormos, 2006)。

3点目は、発表語彙としてのPVLТ・PAFT

とspeed fluencyとの関連である。先行研究では、発表単一語彙知識(Koizumi & In'nami, 2013; Uchihara et al., 2021)、発表複数語知識(Koizumi & In'nami, 2013)、発表単一語と発表複数語知識の合計点(De Jong et al., 2013)でspeed fluencyとの関連が見られるが、今回のPVLТとPAFTはどちらもspeed fluencyの指標である調音速度とは相関しなかった。考えられる要因の1つにテスト形式がある。PVLТは発表語彙知識のテストとして早くから脚光を浴びてきたが(Laufer & Nation, 1999)、あらかじめ答えさせる語彙を指定している点で「統制された」テストであり、その妥当性が疑問視されている(Meara & Fitzpatrick, 2000)。一方、Uchihara et al. (2021)が使用しているLex30は、語彙連想課題を用いることであらかじめ答えさせる語彙を指定しない「自由な」テストとして注目されている(Meara & Fitzpatrick, 2000)。今回作成したPAFTはPVLТの形式を踏襲しており、その点でどちらも調音速度と関連しなかったと考えられる。他に考えられる要因として、speed fluencyの指標の算出方法による違いがある。「統制された」発表語彙テストを使用したKoizumi and In'nami (2013)はspeed fluencyの指標として1分間の総発語数を使用しているが、ポーズを含んでいるため、調音速度のような純粋なspeed fluencyの指標にはなっていない。またDe Jong et al. (2013)は、音節の平均発声時間を算出しており、調音速度とは逆の指標になっている。今後は語彙知識と流暢性の関係性を算出方法も含めて体系的に検証する必要があるだろう。

6.2 課題2

学習者の言語使用から複数語ユニットを同定する方法論については結論が出ていないが(Siyanova-Chanturia & Pellicer-Sánchez, 2019)、今回の研究ではn-gram分析を行い、頻度や結びつきの強さを表す指標を使用した。発話データに対してn-gram分析を行い、発話の流暢性との関連性を調査した研究はTavakoli and Uchihara (2019)のみだが、今回の研究結果とは一部類似しているもののいくつか異なる点も明

らかになった。

本研究は、Tavakoli and Uchihara(2019)同様、n-gramの数ある変数を主成分分析によって集約したが、抽出された因子と因子数が非常に似ている点は注目に値する。Mutual informationによって測定されるn-gramはt-scoreやfrequencyと比べて特異である点や、t-scoreとfrequencyが同じ因子になり、どちらも似たようなn-gramを測定している点が興味深い。この結果はSaito(2020)やEguchi and Kyle(2020)の因子分析結果とも類似している。

発話流暢性との関連で類似している点としては、高頻度3語が調音速度と相関している点が挙げられる。ただ、今回の研究では、高頻度の3語だけではなく、高頻度2語と相互情報量因子も調音速度と有意に相関しており(表3)、発話内で使用される複数語ユニットと発話のスピードには関連があると言える。特に、重回帰分析の結果では相互情報量因子のみが調音速度を予測できていることから(表4)、相互情報量因子で測られる2語と3語のn-gramは発話のスピードを予測する重要な指標であることが分かる。

重回帰分析の結果を見ると、発話流暢性の他の側面(breakdown, repair fluency)に対しても、相互情報量因子が重要な説明変数であることが分かった。Breakdown fluency(表5)では無声ポーズと有声ポーズの頻度を予測しており、repair fluency(表6)では発話の繰返しに対して予測をした。先行研究から、相互情報量は複数語ユニットの洗練性を示す指標(frequencyやt-scoreなど)の中でも特に様々な目的変数を説明できる指標であることが分かっている。Durrant and Schmitt(2009)は学習者と母語話者の作文データを比べ、相互情報量が二者を区別する要素であることを示した。また、Saito(2020)では発話の分かりやすさ(comprehensibility)に対して相互情報量が強く予測をしており、Eguchi and Kyle(2020)ではわずかではあるが相互情報量が全体的な(holistic)スピーキング能力を予測している。

研究課題2の以上の結果を受けて、具体的に相互情報量によって測られるn-gramにはどのようなものがあるのかを提示することで、教育的示唆も生まれると考えられる。そこで、本研究で参

加者が使用したn-gramのうち、相互情報量の平均数値が一番高かった参加者と一番低かった参加者を抜粋し、具体的なn-gramを提示したい(表7はbi-gram, 表8はtri-gram)。相互情報量は、数値が3を超えると重要なコロケーションであるとみなされる(Durrant & Schmitt, 2009; Paquot, 2018)。

分析の結果、bi-gram(表7)では、平均スコアの高い参加者は相互情報量の数値が3を超える頻度が11(発話全体で21%)であるのに対し、平均スコアの低い参加者は6(発話全体で6%)であった。Tri-gram(表8)では、平均スコアの高い参加者は相互情報量の数値が3を超える頻度が23(発話全体で17%)であるが、平均スコアの低い参加者はわずか2(発話全体で3%)にとどまった。

6.3 課題3

語彙知識(PAFT, PVLT, PVST)と複数語ユニット使用(n-gram)のどちらがより発話流暢性と関連があるかを調査したのは本研究が筆者が知る限り初めてである。重回帰分析の結果、speed fluency(表4)とrepair fluency(表6)に対しては「知識」と「使用」の両側面から予測がされ、breakdown fluencyに対しては、単一語彙知識のPVLTが全体的に大きく予測をしたため、「知識」が重要な側面であると言えるが、複数語ユニット使用の指標である相互情報量もポーズの頻度を予測したため、「使用」の側面も重要であると言える。

Breakdown fluencyに関して、相関分析では、無声ポーズの頻度に対して語彙知識が主に相関関係を示しており、相互情報量は強い相関を示していなかったが、重回帰分析では相互情報量が予測変数に入った。理由として考えられることは、語彙知識、語彙使用、流暢性の3者関係の理論的つながりである。発話流暢性に対しては、語彙知識が予測をするパターンと語彙使用が予測をするパターンがそれぞれ考えられるが、語彙知識→語彙使用→発話流暢性という理論的關係性も十分考えられる(Koizumi, 2013)。つまり、知識として持っている語彙をまずは発話で使うことができ、発話で使うことが出来た語彙(使用)が今度は流暢性と関連する、ということである。

表7: bi-gram MIのスコアごとの具体例

MI	MIスコアが高い参加者 (Mean = 2.21)			MIスコアが低い参加者 (Mean = 1.07)		
	頻度	%	実際の使用例	頻度	%	実際の使用例
≥5	4	7.55%	<i>each other, everyday life</i>	2	1.98%	<i>wake up</i>
≥4 & <5	2	3.77%	<i>agree with, positive effect</i>	2	1.98%	<i>agree with, last year</i>
≥3 & <4	5	9.43%	<i>I think, communicate with</i>	2	1.98%	<i>I think, up till</i>
≥2 & <3	9	16.98%	<i>when we, effect on</i>	14	13.86%	<i>result of, addict to</i>
≥1 & <2	9	16.98%	<i>because I, but we</i>	17	16.83%	<i>stay up, job in</i>
<1	11	20.75%	<i>situation if, other when</i>	39	38.61%	<i>cause we, in past</i>
n/a	12	22.64%	<i>talk each, other easily</i>	21	20.79%	<i>negative affection, enrich we</i>

注. n/aはCOCA (Corpus of Contemporary American English) に事例が見つからず分析不可能であることを示す; %は頻度を総発語数で割ったものを表す; MIは相互情報量を指す

表8: tri-gram MIのスコアごとの具体例

MI	MIスコアが高い参加者 (Mean = 3.06)			MIスコアが低い参加者 (Mean = 1.39)		
	頻度	%	実際の使用例	頻度	%	実際の使用例
≥5	10	7.52%	<i>first of all, be more likely</i>	0	0.00%	-
≥4 & <5	4	3.01%	<i>I agree with, to depend on</i>	0	0.00%	-
≥3 & <4	9	6.77%	<i>to spend time, less likely to</i>	2	3.23%	<i>effect on we, and so on</i>
≥2 & <3	2	1.50%	<i>there be two, be the reason</i>	2	3.23%	<i>we have to, it be because</i>
≥1 & <2	8	6.02%	<i>effect on we, deal with problem</i>	3	4.84%	<i>so when we, have to use</i>
<1	7	5.26%	<i>agree with this, we be more</i>	0	0.00%	-
n/a	74	55.64%	<i>eyesight will get, of all by</i>	45	72.58%	<i>bluelight inject we, physical effect on</i>

注. n/aはCOCA (Corpus of Contemporary American English) に事例が見つからず分析不可能であることを示す; %は頻度を総発語数で割ったものを表す; MIは相互情報量を指す

重回帰分析では語彙知識と語彙(複数語ユニット)使用をまとめて予測変数としたために、語彙使用の指標である相互情報量が語彙知識より優先された可能性が高い。加えて、語彙知識のうち複数語知識のPVSTとPAFTが相互情報量に取って代われ、PVLТは残っていることから、同じ複数語として、相互情報量がPVSTとPAFTより優先された可能性が考えられる。

Speed fluencyとrepair fluencyに関して、今回は単一語彙使用の調査は出来ていないが、発話流暢性にとって複数語ユニットの知識を持っていること及び発話で使用できることが重要であることが確認できた。語彙知識のサイズが大きいことで、発話生成の処理において語彙処理

以外の概念化などの他のリソースにより多くの注意を向けることができ、流暢性に繋がると考えられているが(Koizumi & In'nami, 2013)、今回の複数語ユニットの知識でも同じことが言えるであろう。特に、単一語彙知識が予測変数に入っていないことから、単一語彙よりも複数語ユニット知識のサイズの方が流暢性に関わりが深い可能性がある。また、発話で複数語ユニットを使用することで発話スピードが増すという点では、単一語彙と同程度の認知負荷でより多くの発語数を稼げることが要因と考えられる(Wray, 2002)。また、複数語ユニットが別名定型表現であることから分かるように、決まった言い回しはバリエーションが少ないために正

確性を伴う(‘Zones of safety’: Boers et al., 2006)。よって、複数語ユニットを使用することでその分間違いを犯す確率が低くなり、繰り返しや訂正する頻度(repair fluency)が減ると考えられる。

7 今後の課題と結論

本研究では、学習者の発表複数語ユニットの知識と発話における複数語ユニットの使用が発話流暢性の諸側面に対してどのように関連を示すかを調査した。課題1では、PAFTはPVLととも無声ポーズの多くと相関し、発話流暢性に対する発表語彙の重要性が改めて認識されることとなった。また、PAFTはPVSTとともに節内の無声ポーズ数と相関し、複数語ユニットの知識を持つことで、発話生成における言語形成の処理を効率的に出来る示唆も得られた。発表複数語知識としてのPAFTが、発表単一語彙知識を測るPVLと似たパターンの結果を示したことは大変興味深いが、今回の研究は「受容単一語彙」のテストを含んでいない為、「単一語彙」であることが重要なのか、「発表語彙」であることが重要なのかは定かではない。今後、複数語ユニットの発話における役割を調査するうえでも、「受容単一語彙」のテストを含めた研究が必要になるであろう。また、探索的因子分析や構造方程式モデリングなどを使用し、PVLとPAFTが1つの潜在因子にまとめられるかどうかも見いていく必要があるだろう。また、PAFTはPVLと同様に紙上(オンライン上)でテストを行う形式であったが、発表語彙知識の研究分野発展のためには、発話データと比較をする際はテストを音声で行う必要がある(Foster, 2020; Uchihara & Saito, 2019)。特に、発話の流暢性との関連では、回答の反応時間(reaction time)を計測できるテストを使用することで、流暢性との関連性を妥当性を持って検証することが可能になるはずである(De Jong et al., 2012, 2013)。

本研究の課題2では、発話における複数語ユニットの使用をn-gramで検証した。高頻度2語、高頻度3語、相互情報量因子すべてが調音速度と

相関を示し、複数語ユニットと発話のスピードの関係性を改めて認識することが出来た。また、相互情報量因子は流暢性の予測に大きく貢献をし(表4~6)、相互情報量と発話能力の関係を示した先行研究との整合性を確かめることが出来た。しかし、従来のn-gram分析には1つ欠点がある。それは、n-gramの分析は発話データをすべて文字に起こしたものを分析にかけると、語と語の間に入るポーズを考慮していない点である(Myles & Cordier, 2017)。例えば、「first of all」というtri-gramを「first … hmm … first … of … all」といった具合にポーズをしながら発話をした場合、「first of all」は学習者にとって複数語ユニットである可能性が低く、個々の単語のパーツをその場で組み合わせただけにすぎない。Myles and Cordier(2017)は0.2秒以上をポーズとして、ポーズ間の発話のみを複数語ユニットの分析対象とするべきであると提唱している。今後の課題として、発話における複数語ユニットを同定する場合、無声・有声ポーズをきちんと考慮した方法で分析することが必要であろう。

課題3では、発話流暢性の諸側面に対する語彙の知識と複数語ユニットの使用の関連性を調査した。Breakdown fluencyは単一語彙知識(PVL)が指標の多くを予測したものの、複数語ユニット使用(相互情報量)も無声ポーズ数を予測した。一方、speed fluencyとrepair fluencyは複数語ユニットの知識と使用が予測をした。Breakdown fluencyに関しては、複数語ユニット使用が複数語ユニット知識より優先されたことで、知識→使用→流暢性の理論的つながりが示唆された。よって、今後このつながりを検証するため、分析で予測関係のパスをある程度自由に引くことが出来るパス解析や、課題1でも触れたように変数同士のパスだけでなく潜在変数を考慮できる構造方程式モデリングを使用して分析をすることが考えられる。また、今後の課題として、音声面を考慮した発表語彙テストを実施するほか(音声による受容語彙テスト:McLean et al., 2015)、語彙以外の文法や発音の知識についても「知識と使用」という観点から流暢性との関係性を調査していくことが挙げられる。

最後に、本研究の方法論に関する限界点を述べ、今後の課題としたい。まず発話課題であるが、

1つのトピックのみ実施したため、トピックの影響を考慮できなかった。今後は2つ以上のトピックを使用し、データを平均化するなどしてトピックの影響を考慮したい。また、統計分析に使用した重回帰分析であるが、ステップワイズ法を用いており、分析結果の再現性が低くなっている恐れがある。階層的重回帰分析を使用することで理論に基づいた変数の投入ができ、再現性も高まると考えられる。また、発話課題の指示とそれに伴う発話時間のばらつきについて述べ締めくくりとしたい。発話課題において、被験者は2分間の準備時間を与えられた後、制限時間なしで発話を行うよう指示をされた。資料4の発話時間の記述統計を見ると、24.9秒から346.3秒までの幅がある上に、標準偏差が57.6秒と大きい値を示している。制限時間を設けなかったことが大きなばらつきを生む結果になったと考えられる。制限時間が設定されていなければ、発話を時間内に収める必要がなくなるので、認知的負荷が減少する。そのことによって発話の流暢性に影響が出ている可能性がある。話す内容が決まっている絵描写タスクなどとは違い、今回使用した自由に話す意見型のタスクではこうした発話課題の指示と発話時間が発話のパフォーマンスに影響があることを指摘しておきたい。

8 教育的示唆

「どうすればペラペラと英語を話せるようになるのか」という疑問は教師だけではなく、英語を学習するあらゆる人が持つものである。この疑問に対し、本研究は語彙の知識と発話での語彙使用の両側面から示唆のあるものであった。語彙の知識の側面から得られた示唆は、「発表語彙」知識はポーズの頻度・長さに関連が深いという点である。ただ語彙を知っているというだけではなく、口からすぐに出てくる状態を作ることによって、発話の途中で止まることが少なくなり、止まっている時間が短くなるということである。本研究で使用した発表単一語テストは、例文が与えられている中に目標語彙が埋め込まれていることから、発表語彙力を増強するにあたっては語

彙リストなどを使った学習法よりもコミュニケーション活動などを通して語彙を学習することが効果的であろう。

また、発話での語彙使用の側面から得られた示唆として、お互いの結びつきが強い語同士のペア(例:wake up, last year)は発声の速さと無声ポーズの数と関連を持つという点が挙げられる。近年のメタ分析調査により、聞き手が発話の流暢性を判断する際には節内の無声ポーズが重要であることが明らかになっており(Suzuki et al., 2021)、これらの複数語ユニットを発話で使いこなすことで、発話自体が流暢になるだけでなく、他者が発話を聞いた時に流暢と判断されやすくなる、ということである。実際の発話で複数語ユニットを使いこなすには、日頃から英語で話す機会をたくさん持ち、発話の中で使ってみる積極的な姿勢が必要であろう。教室では、意味のやり取りを中心にしたコミュニケーション活動を増やし、目標語彙を意識しながら発話をするように学習者を促すことが効果的であると考えられる。

謝辞

まずはこのように研究をする機会を与えてくださった公益財団法人 日本英語検定協会関係者の皆様、ならびに選考委員の皆様にご場を借りて感謝申し上げます。助言担当についてくださった小泉利恵先生には、研究初期の段階で有益なアドバイスをしていただき、その後スムーズに実験を行うことが出来ました。ランカスター大学博士課程の鈴木駿吾さんからは研究に関する多大なフィードバック、助言、励ましをいただきました。早稲田大学修士課程の原田美涼さんと野々村大樹さんにはresearch assistantとして分析を手伝っていただきました。この場を借りて厚く御礼申し上げます。

参考文献(*は引用文献)

- * Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a Lexical Approach to the test. *Language Teaching Research*, 10(3), 245-261. <https://doi.org/10.1191/1362168806lr195oa>
- * Boersma, P., & Weenink, D. (2020). PRAAT: Doing phonetics by computer (version 6.1.44) [Computer program]. www.praat.org
- * Bosker, H. R., Pinget, A. F., Quené, H., Sanders, T., & De Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, 30(2), 159-175. <https://doi.org/10.1177/0265532212455394>
- * Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, 32(2012), 45-61. <https://doi.org/10.1017/S0267190512000074>
- * Davies, M. (2008). The Corpus of contemporary American English (COCA). <https://www.english-corpora.org/coca/>
- * De Jong, N. H., Pacilly, J., & Heeren, W. (2020). *Praat scripts to measure speed fluency and breakdown fluency in speech automatically*. <https://osf.io/7xy53/>
- * De Jong, N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2012). Facets of speaking proficiency. *Studies in Second Language Acquisition*, 34(1), 5-34. <https://doi.org/10.1017/S0272263111000489>
- * De Jong, N. H., Steinel, M. P., Florijn, A., Schoonen, R., & Hulstijn, J. H. (2013). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics*, 34(5), 893-916. <https://doi.org/10.1017/S0142716412000069>
- * Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, 54(4), 655-679. <https://doi.org/10.1111/j.1467-9922.2004.00282.x>
- * Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *IRAL - International Review of Applied Linguistics in Language Teaching*, 47(2), 157-177. <https://doi.org/10.1515/iral.2009.007>
- * Eguchi, M., & Kyle, K. (2020). Continuing to explore the multidimensional nature of lexical sophistication: The case of Oral Proficiency. *The Modern Language Journal*, 104(2), 381-400. <https://doi.org/10.1111/modl.12637>
- * Erdfelder, E., FAul, F., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149-1160. <https://doi.org/10.3758/BRM.41.4.1149>
- * Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, 20(1), 29-62. <https://doi.org/10.1515/text.1.2000.20.1.29>
- * Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.
- * Field, A. (2009). *Discovering statistics using SPSS* (Third). SAGE Publications.
- * Fitzpatrick, T., & Clenton, J. (2010). The challenge of validation: Assessing the performance of a test of productive vocabulary. *Language Testing*, 27(4), 537-554. <https://doi.org/10.1177/0265532209354771>
- * Fitzpatrick, T., & Clenton, J. (2017). Making Sense of Learner Performance on Tests of Productive Vocabulary Knowledge. *TESOL Quarterly*, 51(4), 844-867. <https://doi.org/10.1002/tesq.356>
- * Foster, P. (2020). Oral fluency in a second language: A research agenda for the next ten years. *Language Teaching*, 53(4), 446-461. <https://doi.org/10.1017/CB09781107415324.004>
- * Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21(3), 354-375. <https://doi.org/10.1093/applin/21.3.354>
- * Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, 67(S1), 155-179. <https://doi.org/10.1111/lang.12225>
- * Kahng, J. (2014). Exploring Utterance and Cognitive Fluency of L1 and L2 English Speakers: Temporal Measures and Stimulated Recall. *Language Learning*, 64(4), 809-854. <https://doi.org/10.1111/lang.12084>
- * Kahng, J. (2018). The effect of pause location on perceived fluency. *Applied Psycholinguistics*, 39(3), 569-591. <https://doi.org/10.1017/S0142716417000534>
- * Kahng, J. (2020). Explaining second language utterance fluency: Contribution of cognitive fluency and first language utterance fluency. *Applied Psycholinguistics*, 41(2), 457-480. <https://doi.org/10.1017/S0142716420000065>
- * Koizumi, R. (2013). Vocabulary and Speaking. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 1-7) [online edition]. John Wiley & Sons, Inc. <https://doi.org/10.1002/9781405198431.wbeal1431>
- * Koizumi, R., & In'nami, Y. (2013). Vocabulary knowledge and speaking proficiency among second language learners from novice to intermediate levels. *Journal of Language Teaching and Research*, 4(5), 900-913. <https://doi.org/10.4304/jltr.4.5.900-913>
- * Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning*, 49(2), 303-342. <https://doi.org/10.1111/0023-8333.00090>
- * Kormos, J. (2006). *Speech production and second language acquisition*. Lawrence Erlbaum.
- * Kyle, K., & Crossley, S. A. (2015). Automatically assessing lexical sophistication: Indices, tools, findings, and application. *TESOL Quarterly*, 49(4), 757-786. <https://doi.org/10.1002/tesq.194>
- * Lambert, C., Aubrey, S., & Leeming, P. (2020). Task preparation and second language speech production. *TESOL Quarterly*, 55(2), 331-365. <https://doi.org/10.1002/tesq.598>
- * Lambert, C., Kormos, J., & Minn, D. (2017). Task repetition and second language speech processing. *Studies in*

参考文献 (*は引用文献)

- Second Language Acquisition*, 39(1), 167-196. <https://doi.org/10.1017/S0272263116000085>
- * Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. Routledge.
- * Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16(1), 36-55. <https://doi.org/10.1191/026553299672614616>
- * Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.
- * Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, 33(3), 299-320. <https://doi.org/10.1093/applin/ams010>
- * McLean, S., Kramer, B., & Beglar, D. (2015). The creation and validation of a listening vocabulary levels test. *Language Teaching Research*, 19(6), 741-760. <https://doi.org/10.1177/1362168814567889>
- * Meara, P., & Fitzpatrick, T. (2000). Lex30: An improved method of assessing productive vocabulary in an L2. *System*, 28(1), 19-30. [https://doi.org/10.1016/S0346-251X\(99\)00058-5](https://doi.org/10.1016/S0346-251X(99)00058-5)
- * Myles, F., & Cordier, C. (2017). Formulaic sequence(FS) cannot be an umbrella term in SLA: Focusing on psycholinguistic FSs and their identification. *Studies in Second Language Acquisition*, 39(1), 3-28. <https://doi.org/10.1017/S027226311600036X>
- * Nation, P. (2017). The BNC/COCA Level 6 word family lists (Version 1.0.0). <http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>
- * Paquot, M. (2018). Phraseological competence: A missing component in university entrance language tests? Insights from a study of EFL learners' use of statistical collocations. *Language Assessment Quarterly*, 15(1), 29-43. <https://doi.org/10.1080/15434303.2017.1405421>
- * Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and Communication* (pp. 191-226). Longman. <https://doi.org/10.4324/9781315836027-12>
- * Perkins, M., & Wray, A. (2000). The functions of formulaic language. *Language and Communication*, 20, 1-28.
- * Plonsky, L., & Oswald, F. L. (2014). How big is "big" ? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878-912. <https://doi.org/10.1111/lang.12079>
- * R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.r-project.org/>
- * Saito, K. (2020). Multi- or single-word units? The role of collocation use in comprehensible and contextually appropriate second language speech. *Language Learning*, 70(2), 548-588. <https://doi.org/10.1111/lang.12387>
- * Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.
- * Segalowitz, N. (2016). Second language fluency and its underlying cognitive and social determinants. *IRAL - International Review of Applied Linguistics in Language Teaching*, 54(2), 79-95. <https://doi.org/10.1515/iral-2016-9991>
- * Simpson-vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, 31(4), 487-512. <https://doi.org/10.1093/applin/amp058>
- * Siyanova-Chanturia, A., & Pellicer-Sánchez, A. (2019). Formulaic language: Setting the scene. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 1-16). Routledge.
- * Siyanova-Chanturia, A., & Spina, S. (2019). Multi-word expressions in second language writing: A large-scale longitudinal learner corpus study. *Language Learning*, 70(2), 420-463. <https://doi.org/10.1111/lang.12383>
- * Skehan, P., Foster, P., & Shum, S. (2016). Ladders and snakes in second language fluency. *IRAL - International Review of Applied Linguistics in Language Teaching*, 54(2), 97-111. <https://doi.org/10.1515/iral-2016-9992>
- * Suzuki, S., & Kormos, J. (2019). Linguistic dimensions of comprehensibility and perceived fluency: An investigation of complexity, accuracy, and fluency in second language argumentative speech. *Studies in Second Language Acquisition*, 42, 143-167. <https://doi.org/10.1017/S0272263119000421>
- * Suzuki, S., Kormos, J., & Uchihara, T. (2021). The relationship between utterance and perceived fluency: A meta-analysis of correlational studies. *The Modern Language Journal*, 105(2), 435-463. <https://doi.org/10.1111/modl.12706>
- * Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239-277). John Benjamins.
- * Tavakoli, P., & Uchihara, T. (2019). To what extent are multiword sequences associated with oral fluency? *Language Learning*, 70(2), 506-547. <https://doi.org/10.1111/lang.12384>
- * Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning*, 61(2), 569-613. <https://doi.org/10.1111/j.1467-9922.2010.00622.x>
- * Uchihara, T., & Saito, K. (2019). Exploring the relationship between productive vocabulary knowledge and second language oral ability. *Language Learning Journal*, 47(1), 64-75. <https://doi.org/10.1080/09571736.2016.1191527>
- * Uchihara, T., Saito, K., & Clenton, J. (2021). Re-examining the relationship between productive vocabulary and second language oral ability. In Clenton, J., & Booth, P. (Eds.), *Vocabulary and the four skills: Pedagogy, practice, and implications for teaching vocabulary* (pp. 146-165). Routledge.
- * West, M. (1953). *A general service list of English words*. Longman.
- * Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

資料1: PAFT 作成手続き

Academic Formulas List (Simpson & Ellis, 2010)	Spoken	Written	Core
	<i>n</i> = 200	<i>n</i> = 200	<i>n</i> = 207
項目削除基準	項目削除数		
① Formulas Teaching Worth (FTW = 教師による評価)が 1.0 以下のもの	-123	-144	-172
② 固有名詞 (例: University of Michigan)	-3	-1	
③ 疑問文と否定文	-4	-3	
④ 過度な口語調 (例: blah blah blah)	-9		
⑤ 冠詞が最後に来ているもの (例: look at the)	-6	-8	-5
⑥ be 動詞から始まる / で終わるもの (例: is likely to be)	-1	-2	-1
⑦ Partial duplicates (Rogers, 2021) は FTW が高いものを優先する (例: in such a way: <u>1.11</u> vs in such a way that: <u>2.32</u> vs such a way that: <u>1.22</u>)	-22	-13	-10
⑧ 文法規則だけで言えるもの (例: might be able to)、and/so が含まれるもの、 統語的に完成していないもの (例: there is no)	-12	-4	-2
⑨ その他	-6	-12	-1
⑩ テストの例文が作りづらいもの	-1	-2	-3
最終リスト			
	<i>n</i> = 12	<i>n</i> = 10	<i>n</i> = 13
信頼性分析・項目分析適用後			
削除項目: 14, 17, 29, 32	<i>n</i> = 10	<i>n</i> = 8	<i>n</i> = 13

資料2: テストに使用したPAFT35項目(うち4項目は削除項目)

1. Several observations can be made wi re t the bar graph on the page.
2. I can't tell wh o n the teacher likes me.
3. Height increases as a fu o aging in children.
4. I t o health conditions, the U.S. is worse than Japan.
5. She was treated differently o th ba o gender.
6. He passed away. I oth w, he is no longer in this world.
7. They have a house in France a w a in Spain.
8. Mike and I arrived at the station a t sa ti.
9. She died a a re o a car accident.
10. He walked around i or t stay warm.
11. In t ca o Mary, the story is truth.
12. Students discuss ideas, as op t just copying from books.
13. Young drivers are m li t have accidents than older drivers.
14. I k o understood your question. (削除項目)
15. Her high GPA has not t d wi the amount of studying she does.
16. Let me tell you a li b ab myself.
17. Do you s w I'm saying? (削除項目)
18. It tu o that I was wrong.
19. Starting with a small goal is t best wa t motivate yourself.
20. We've been asked to c u wi some new ideas.
21. If you lo a the painting closely, you can see it's very nice.
22. L m just say one thing.
23. I'm going to Kyoto a th en o the month.
24. It do ma if you're a few minutes late. We'll wait for you.
25. T sa th has happened again and again in history.
26. Tom likes math. O t o h, John likes English.
27. It i imp t read the instructions carefully before you start.
28. He explained i s a w that many people couldn't understand him.
29. This could g r t questions such as: How are shadows made? (削除項目)
30. Everyone has h or h own choices.
31. Ov a p o ten years, he stole a million dollars from the company.
32. He had included a la nu o interesting stories in his speech. (削除項目)
33. The newspaper failed to t i ac an important problem.
34. His poor exam score is d t t fa t he is always playing games every night.
35. Those students are from a wi ra o backgrounds.

資料3: PAFTの答えとAFL (Simpson & Ellis, 2010)でのSpoken, Written, Coreの種別

1. Several observations can be made **with respect to** the bar graph on the page. (Core)
2. I can't tell **whether or not** the teacher likes me. (Core)
3. Height increases as a **function of** aging in children. (Core)
4. **In terms of** health conditions, the U.S. is worse than Japan. (Core)
5. She was treated differently **on the basis of** gender. (Core)
6. He passed away. **In other words**, he is no longer in this world. (Core)
7. They have a house in France **as well as** in Spain. (Core)
8. Mike and I arrived at the station **at the same time**. (Core)
9. She died **as a result of** a car accident. (Core)
10. He walked around **in order to** stay warm. (Core)
11. In **the case of** Mary, the story is truth. (Core)
12. Students discuss ideas, **as opposed to** just copying from books. (Core)
13. Young drivers are **more likely to** have accidents than older drivers. (Core)
14. I **kind of** understood your question. (Spoken)
15. Her high GPA has **nothing to do with** the amount of studying she does. (Spoken)
16. Let me tell you a **little bit about** myself. (Spoken)
17. Do you **see what** I'm saying? (Spoken)
18. It **turns out** that I was wrong. (Spoken)
19. Starting with a small goal is **the** best **way to** motivate yourself. (Spoken)
20. We've been asked to **come up with** some new ideas. (Spoken)
21. If you **look at** the painting closely, you can see it's very nice. (Spoken)
22. **Let me** just say one thing. (Spoken)
23. I'm going to Kyoto **at the end of** the month. (Spoken)
24. It **doesn't matter** if you're a few minutes late. We'll wait for you. (Spoken)
25. **The same thing** has happened again and again in history. (Spoken)
26. Tom likes math. **On the other hand**, John likes English. (Written)
27. It **is important to** read the instructions carefully before you start. (Written)
28. He explained **in such a way** that many people couldn't understand him. (Written)
29. This could **give rise to** questions such as: How are shadows made? (Written)
30. Everyone has **his or her** own choices. (Written)
31. **Over a period of** ten years, he stole a million dollars from the company. (Written)
32. He had included a **large number of** interesting stories in his speech. (Written)
33. The newspaper failed to **take into account** an important problem. (Written)
34. His poor exam score is **due to the fact that** he is always playing games every night. (Written)
35. Those students are from a **wide range of** backgrounds. (Written)

資料4: 発話時間・発話語数の記述統計

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>Skewness</i>	<i>Kurtosis</i>	<i>p-value of Shapiro-Wilk</i>
発話時間	104.4	57.6	24.9	346.3	1.50	3.60	< .001
発話語数	105.4	54.5	28.0	306.0	1.53	3.18	< .001

資料5: 発話流暢性指標の記述統計

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>Skewness</i>	<i>Kurtosis</i>	<i>p-value of Shapiro-Wilk</i>
<i>Speed fluency</i>							
調音速度	2.70	0.68	1.19	4.01	-0.30	-0.48	0.42
<i>Breakdown fluency</i>							
節内無声ポーズ割合	0.23	0.13	0.05	0.67	1.26	2.13	< .001
節外無声ポーズ割合	0.08	0.03	0.02	0.25	2.35	10.56	< .001
節内無声ポーズ時間(秒)	0.85	0.30	0.45	1.73	1.01	0.61	< .001
節外無声ポーズ時間(秒)	0.89	0.35	0.37	1.96	0.90	0.59	0.004
有声ポーズ割合	0.11	0.11	0.00	0.47	1.63	2.88	< .001
<i>Repair fluency</i>							
繰返し割合	0.10	0.08	0.00	0.43	1.74	4.23	< .001
自己訂正割合	0.03	0.03	0.00	0.10	0.79	-0.16	< .001
擬発言割合	0.01	0.01	0.00	0.06	1.85	4.90	< .001

注. それぞれ取りうる値は0以上

資料6: 語彙知識の記述統計

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>CV</i>	<i>Skewness</i>	<i>Kurtosis</i>	<i>p-value of Shapiro-Wilk</i>	<i>Cronbach's alpha</i>
PVLT	22.41	6.87	11	41	0.31	0.68	0.07	0.028	0.825
PVST	33.79	6.12	20	44	0.18	-0.44	-0.49	0.059	0.780
PAFT	11.40	6.40	1	26	0.56	0.43	-0.72	0.035	0.884

注. PVLT(発表単一語)は54点満点, PVST(受容複数語)は50点満点, PAFT(発表複合語)は31点満点; CVはcoefficient of variation(変動係数)で, SDをMで割ること算出

資料7: 複数語ユニット使用(n-gram 指標)の記述統計

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>Skewness</i>	<i>Kurtosis</i>	<i>p-value of Shapiro-Wilk</i>
<i>Bi-gram</i>							
Frequency (logarithmized)	1.67	0.18	1.27	2.28	0.45	1.08	0.356
Range	0.24	0.05	0.14	0.42	0.57	0.82	0.123
Mutual information	1.61	0.20	1.07	2.22	0.41	1.20	0.294
T-score	85.59	26.17	26.08	176.12	0.59	1.63	0.048
<i>Tri-gram</i>							
Frequency (logarithmized)	1.10	0.23	0.63	1.97	0.87	2.02	0.024
Range	0.09	0.05	0.01	0.24	1.10	1.61	0.002
Mutual information	2.37	0.29	1.40	3.06	-0.37	1.74	0.127
T-score	43.52	20.35	14.04	120.41	1.45	3.00	< .001

注. それぞれ取りうる値は0以上

資料8: 語彙知識、複数語ユニット使用、発話流暢性指標のスパイマン相関行列

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. PVLT	—													
2. PVST	.592***	—												
	[.404, .733]													
3. PAFT	.681***	.698***	—											
	[.522, .795]	[.545, .807]												
4. 読音速度	0.151	.233*	0.202	—										
	[-.101, .384]	[.005, .471]	[-.048, .429]											
5. 節内無声ボーズ割合	-0.228	-.308*	-.276*	-.717***	—									
	[-.450, .021]	[-.516, -.065]	[-.490, -.030]	[-.819, -.571]										
6. 節外無声ボーズ割合	-.360**	-0.234	-.261*	-.363**	.556***	—								
	[-.558, -.124]	[-.455, .015]	[-.478, -.014]	[-.560, -.127]	[.357, .706]									
7. 節内無声ボーズ時間 (秒)	-.322*	-0.239	-.295*	-.315*	.587***	.419***	—							
	[-.527, -.080]	[-.460, .009]	[-.506, -.051]	[-.522, -.073]	[.398, .729]	[.191, .604]								
8. 節外無声ボーズ時間 (秒)	-.520***	-0.102	-.321*	-0.006	.280*	.411***	.612***	—						
	[-.680, -.312]	[-.341, .149]	[-.527, -.080]	[-.254, .242]	[.494, .035]	[.181, .598]	[.429, .746]							
9. 有声ボーズ割合	0.005	-0.355	-0.01	-.712***	-.471***	0.198	-0.097	-0.2	—					
	[-.243, .253]	[-.298, .196]	[-.257, .238]	[-.816, -.564]	[.253, .644]	[-.052, .425]	[-.337, .155]	[-.427, .050]						
10. 繰返し割合	-0.047	-0.208	-0.188	-.771***	.607***	.256*	.329***	-0.024	.532***	—				
	[-.292, .203]	[-.433, .042]	[-.416, .063]	[-.855, -.647]	[.423, .743]	[.008, .473]	[.089, .533]	[-.270, .225]	[.328, .689]					
11. 自己訂正割合	-0.13	-.312*	-0.226	-.482***	.318*	0.149	0.111	-0.104	.331**	.465***	—			
	[-.336, .121]	[-.520, -.070]	[-.449, .023]	[-.652, -.266]	[.076, .524]	[-.102, .383]	[-.141, .349]	[-.343, .147]	[.090, .535]	[.245, .639]				
12. 綴音割合	-0.04	-0.363	-0.007	-.358**	0.244	0.157	.280*	0.124	0.156	.358**	0.18	—		
	[-.285, .210]	[-.297, .197]	[-.254, .241]	[-.556, -.120]	[-.004, .464]	[-.094, .390]	[.034, .493]	[-.127, .361]	[-.096, .388]	[.121, .557]	[-.071, .410]			
13. 高頻度二語	0.191	0.211	0.067	.286*	-0.246	-0.024	-0.233	-0.071	-0.208	-.298*	-0.207	-0.255*	—	
	[-.059, .419]	[-.038, .436]	[-.183, .310]	[.041, .498]	[-.465, .002]	[-.270, .225]	[-.455, .015]	[-.313, .180]	[-.483, .042]	[-.508, -.054]	[-.433, .042]	[-.473, -.008]		
14. 高頻度三語	0.03	0.041	-0.118	.307*	-0.102	-0.031	-0.083	0.038	-0.205	-0.246	-0.175	-.401**	.638***	—
	[-.219, .276]	[-.209, .286]	[-.356, .133]	[.065, .516]	[-.341, .149]	[-.277, .219]	[-.324, .168]	[-.212, .283]	[-.461, .045]	[-.465, .002]	[-.405, .076]	[-.590, -.170]	[.463, .765]	
15. 相互情報量因子	-0.165	-0.392	-0.236	.273*	-0.241	-0.04	-0.063	0.088	-.324**	-0.225	-0.17	-0.246	0.221	.352**
	[-.396, .086]	[-.332, .160]	[-.457, .012]	[.027, .488]	[-.461, .008]	[-.285, .101]	[-.305, .188]	[-.164, .328]	[-.529, -.083]	[-.448, .024]	[-.401, .081]	[-.465, .002]	[-.14, .552]	

注: * p < .05, ** p < .01, *** p < .001; []は95%信頼区間を表す