英字幕付き音声動画の視聴による第二言語の 偶発的語彙学習

— 使用する音声動画マテリアルに関わる変数が持つ影響 —

研究者: 黒川 皐月 宮城県／東北大学大学院 学術研究員
(申請時: 東京都／東京外国語大学大学院 在籍)

《研究助言者: 西垣 知佳子》

概要

本研究は、英語を外国語として学習する日本人大学生148名を対象に、第二言語で動画を観る際にいかに字幕が偶発的語彙学習に役立つかを調べた。具体的には(1)英字幕の効果、(2)語彙カバー率の影響、(3)日本語字幕と英字幕の連続的な使用の効果である。参加者を4つの実験群と2つの追加グループに分け、実験一日目に3種類のターゲット語彙事前テストを行った。1週間後の二日目に指定の条件で内容理解を中心に動画を2回視聴してもらい事後テストを行った。これら語彙テストの正誤データに対して、一般化線形混合モデルのロジスティック回帰分析を行った。結果は、どのグループも動画視聴による語彙の学習が見られたが、英字幕、そして日本語字幕と英字幕の連続的な使用による効果は見られなかった。また、字幕の有無に関わらず、95%以上のカバー率がある学習者は有意により多くの語彙を習得できていることが分かった。第二言語での動画視聴に関する教育的示唆についても議論していく。

1 はじめに

近年は動画視聴に関連する第二言語習得研究の報告が数多くある(Wei & Fan, 2022)。ラジオを聴いたり、駅や空港のアナウンスを聞いたりするといった、視覚情報がない聴解活動以上に、動画を観るといった活動は昨今の現実世界にあふれている(Graham & Zhang, 2021)といえるかもしれない。この事実は、テレビだけでなくYouTubeやNetflixなどの動画視聴プラットフォームの人気具合から明らかだろう。このように現実世界に即した媒体や方法で外国語を学ぶことは意義深い。

それだけでなく、言語以外の視覚情報に頼らない文章の読解や聴解による学習に比べ、動画視聴による学習は質や量の面でより良くなる場合があるという主張もある(Mayer, 2005)。Richard E. Mayerと彼の共同研究者たちはCognitive Theory of Multimedia Learningを提唱、研究、そして発展させてきた。この理論の要点は、「人間は言葉だけの場合よりも、言葉と画像がともにあることでより良い学びをする(Mayer, 2005, p.1)」ことであり、この主張は人間の心がどのように働くかに関する3つの仮定をもとに理論的に体系化されたものである(Mayer, 2021)。このように、動画視聴による学習は理論の裏付けがあり、効果的かつ効率的であるとされているため、更なる研究は教育現場に対しても有益なものとなるだろう。

動画視聴を用いた第二言語習得の研究では、動画に字幕を加えることで、いかに学習が促進されるかといったことが中心的な話題となっている。字幕付き動画は学習者の内容に関する理解を促進するだけでなく、発音や文法、語彙、そして語用論的知識の獲得にまでも貢献するとされる(Wei & Fan, 2022によるレビュー論文を参照されたい)。その中でも本研究では、外国語環境の学習者が抱える重大な困難の1つが

語彙を学ぶこと (Webb & Nation, 2017), そして日本の文脈で行われた第二言語動画視聴による語彙習得研究が比較的少ないこと (i.e., Dizon & Thanyawatpokin, 2021; Raine, 2013; Rodgers, 2013) を考慮し、日本人大学生を対象とした字幕付きの第二言語動画視聴による英語語彙学習をテーマとする。字幕というと母語によるものと第二言語によるものが考えられる。英語で書かれた研究では、第一言語字幕と第二言語字幕はそれぞれ (L1) subtitles と captions と呼ばれることが多いが、本研究では第二言語字幕 (つまり英字幕) を中心的に扱いつつも、第一言語による字幕の是非についても検討するため、以降の議論では必要がある際、それぞれの字幕を明示する (より詳しい字幕の分類や用語については Montero Perez, 2022 を参照)。

2 先行研究

2.1 第二言語字幕付きの動画視聴による語彙習得の基礎研究

Montero Perez et al. (2013) は、第二言語字幕が語彙習得へ肯定的な影響を与えるのか調べた実証研究を使ってメタ分析を行った。メタ分析はテーマに沿って似た研究を集めて、Cohen's *d* や Hedges' *g* といった効果量を算出し、統計的に統合する分析のことである。これにより、ひとつひとつの研究を超えて、ある実験的介入に効果があるのかどうか、研究の傾向を確認することができる (Borenstein et al., 2021; Plonsky & Oswald, 2015)。Montero Perez et al. (2013) は、10個の実証研究を統合し、 $g = 0.87$ ($p < .001$) という、中～大程度の大きさの効果量 (Plonsky & Oswald, 2014) を導き出した。つまり、第二言語字幕付きで第二言語動画を視聴することは、字幕なしで同じ動画を視聴する場合よりも、多くの語彙を学ぶことができるという意味である。

しかし、Montero Perez et al. (2013) では考慮されなかった偶発的 (付随的) 学習の観点を組み込んだ近年の研究では、第二言語字幕の効果を支えるものとそうでないものが混在している。偶発的学習は、言語活動の際に学習者の注意が言語の内容に向いているときに、言語の形式面も併せて学習されることをいう (Hulstijn, 2001)。偶発的学習は、形式面に全く注意が向かなかつたり、意識がなかつたりする暗黙的学習 (Ellis, 2009) とは区別される。実験では、完全に言語形式に関する意図を排除することは難しいが、後に行われるテストの存在を実験参加者に知らせなかつたり、意図的に参加者の意識を言語の内容面に向けたりすることで、形式面に関して偶発的学習という構成概念を操作化する (Webb, 2020)。

偶発的学習の環境を作りだし、第二言語字幕の効果を支える研究には、例えば Teng (2022) が挙げられる。彼の研究では、英語を専攻する中国人大学生1年生82人が、BBCドキュメンタリーを英字幕付きで観るグループと字幕なしで観るグループに分けられた。動画視聴後の4つのターゲット語彙テストの結果を比べると、どの語彙テストにおいても字幕の有無が主効果として検出された。同様に Montero Perez et al. (2018) はドイツ語を母語とし、フランス語を学ぶ大学生を対象とした実験で、参加者に各2～4分ほどの異なる動画を3本観させた。ターゲット語彙が視聴した動画に出てきたかどうかを答える form recognition と呼ばれるタイプの語彙テスト (Schmitt, 2014) と各ターゲット語彙がどの動画で現れたかを答える clip association と呼ばれるテストにおいて、第二言語字幕の効果を示した。

反対に、第二言語字幕の効果を発見しなかった研究には、例えば Hao et al. (2022) が挙げられる。Hao et al. (2022) は、第二言語の熟達度が異なるサンプルそれぞれにおいて、実験群と統制群を用いて実験を行っているが、ここでの議論は熟達度の高いサンプルに絞る。英字幕付き動画を観るグループと字幕なしで観るグループの参加者は、それぞれの条件で4つの TED Talks シリーズの動画を視聴した。動画視聴後のターゲット語彙テストは与えられた定義に対応する英単語を選ぶ form recognition のテストである。統計分析の結果、英字幕付きグループは字幕無しグループに比べてテストの結果が有意に悪かった。また、Jelani and Boers (2018) の研究では、マレーシアの高校生を対象に英字幕付きの動画を視聴するグルー

プと字幕なしで動画を観るグループを分け、TED Talksシリーズから1本の動画を2回視聴させた。彼女らの研究で興味深い点は、ターゲット語彙の意味を書くテストで、全体的には確かに英字幕付きグループの方が英字幕なしグループよりも統計的に有意なレベルで得点が高かったが、それはターゲット語彙が書記言語として与えられた場合のみで、ターゲット語彙が音声のみで与えられた場合には得点に違いがなかったのである。

このほかにも偶発的語彙習得に対して第二言語字幕の効果を認める研究が多いものの(e.g., Kang, 2019; Teng, 2018; Wang & Pellicer-Sanchez, 2022)、字幕の効果が見られないことを示唆する研究には日本人英語学習者を対象としているものがある。Rodgers(2013)は博論研究として合計5つの研究を行い、そのうち4つは英字幕なしで第二言語の動画を視聴することに関する研究であり、最後の5つ目で英字幕付きの動画視聴に関する研究を実施した。参加者は日本人大学生で、英語の熟達度は初級から中級レベルとされている。10週に渡り、1つ目の研究の参加者は英字幕なしで、5つ目の研究の参加者は英字幕付きでアメリカのドラマの各エピソードを視聴した。語彙テストの結果は、英字幕の有無は偶発的語彙習得に影響がない、とのことだった。同じく未公開のRaine(2013)はTED Talkを使用し同様の結果を呈している。この原因として、RogersやRaineの研究で使用された動画が語彙や発話速度や訃りなどの音声の観点から、参加者レベルとはマッチしていなかった可能性がある。そのため、より適切なレベルの動画を使用して、日本人英語学習者が英字幕から偶発的語彙学習に関して恩恵を受けられるか、研究の余地がある。

2.2 字幕の効果的な使用に関する発展研究

前述のように、第二言語字幕が必ずしもいつも偶発的語彙習得に寄与するとは限らないことを反映し、近年は第二言語字幕をいかに効果的に扱うかといった観点を持つ研究も見られる。ターゲット語彙や重要な語彙のみを第二言語字幕として提示したり(Montero Perez, et al., 2014)、同様に語彙をハイライトして目立たせたり(Hsieh, 2020)することで、視聴者の意識をより語彙に向ける試みもある。ここではいかに学習者の内容理解を促進することで偶発的学習にも結び付けられるかを概観していく。

一つ目は動画の語彙カバー率である。語彙カバー率とはあるディスコースのまとまり(テキスト)において読み手あるいは聞き手が知っている語彙の割合のことであり(van Zeeland & Schmitt, 2013)、読解研究の分野で多くの研究がなされてきた。あるテキストの十分な内容理解を得るには、そのテキストに出てくる95%以上、できれば98%の語彙を知っている必要があるとされる(Schmitt et al., 2011)。とはいえ、第二言語動画視聴に関しては、動画の性質上、理解を促進する視覚イメージがあるため、98%以下の語彙カバー率でも十分な理解が得られるのではないかと考えられる。Durbahn et al.(2020)は、実験参加者一人一人について動画内の既知の単語量を算出し、それを動画の理解度と対照させた。その結果、理解度確認テストの得点が劇的に変化する語彙カバー率の閾値は無かったが、語彙カバー率と動画の理解度に中程度の相関($r = .39$)を発見した。下限であった語彙カバー率87%の実験参加者でも62%の理解度を示し、語彙カバー率が88%以降では7割近くの理解度になることから、第二言語動画視聴による十分な理解には、概ね90%以上の語彙カバー率を目指すのが良いと教育的示唆を述べている。しかし、実際にこの90%のカバー率で内容理解のみならず、偶発的に語彙が学習されるかについては検証が必要だ。

また、二つ目の動画の内容理解を促進する手立てとして、母語による字幕(これ以降第一言語字幕と呼ぶ)がある。そもそも第二言語字幕を付けるということは、視聴者がある程度の第二言語読解能力を有していることを想定し(Vanderplank, 2016)、動画理解には効率の良い言語処理がかかせない。それに対し、第一言語字幕は文字通り視聴者の第一言語であるため、言語処理に係る認知負荷は明らかに低いと考えられる。第一言語字幕を用いた研究では、確かに字幕がない場合よりも内容理解度テストの結果が良くなることが示されている(Birulés-Muntané & Soto-Faraco, 2016)。しかし、第一言語字幕による語彙

習得の効果は第二言語字幕の場合ほどは見られないことが多い(e.g., Frumuselu et al., 2015; Peters, 2019)。これは、第一言語による意味とそれに対応する第二言語の語彙形式を音声から対応させることが難しいためかもしれない。特に日本語と英語の関係のように、第一言語と第二言語の言語距離が遠い場合には、一般的に言語習得は第一言語と第二言語の言語距離が近い場合よりは難しいとされ(Wesche & Paribakht, 2010)、統語論的な違いや使用される語彙の隔たりのために、言語間の対応が見分けづらいのであろう。

そこで、効果的な第一言語字幕の使用法として、第一言語字幕付きで動画を視聴した後に第二言語字幕付きで同じ動画を観るといった連続的な字幕の使い方が考えられる。第一言語字幕付きで第二言語の動画を視聴する場合は、提示される書記言語が母語のため言語処理が浅くなってしまいう可能性が否めないが、それでもある程度の動画理解が心的表象として作成されたならば(Kintsch, 1998)、その知識をもとに2回目の動画視聴における理解が飛躍的に上がり、偶発的学習の土台となったり、あるいは2回目の動画視聴ではより言語形式面への注意が向いたりするかもしれない。Wi and Boers(2023)は、カナダの大学で学ぶ中国人英語学習者を対象に第一言語字幕と第二言語字幕の連続的な使用を調べた研究である。使用された動画はTED Talksであり、実験群は、1回目の視聴を第一言語字幕付きで、2回目を第二言語字幕付きで、3回目の視聴を字幕なしで視聴し、統制群は3回とも第二言語字幕付きで動画を視聴した。混合ロジスティック回帰分析の結果、ターゲット語彙の意味を書く語彙テストで、グループ間の違いは見られなかった。しかし、この結果を一般化してしまうのは早急である。というのも、Wi and Boers(2023)が認めるようにサンプルサイズがそもそも小さく(実験群 $n = 11$, 対照群 $n = 9$)、且つターゲット語彙が11個のみであったためだ。また、実験デザインの観点からも、彼らの実験参加者は合計3回の動画視聴ができたため、特に統制群における語彙の伸びの大きさは回数による影響が否めない。そのため、第一言語字幕と第二言語字幕の連続的な使用については引き続き研究する価値があるといえる。

3 研究設問

先行研究のセクションで概観したように、偶発的語彙学習に対する第二言語字幕付き動画視聴の影響を調べた基礎研究は日本では比較的稀有である。語彙カバー率の観点から適切なレベル(i.e., 概ね90%以上)の第二言語動画で日本人英語学習者が内容理解だけでなく、偶発的に語彙を学べるとすると、教育現場に応用しうる示唆を得ることができる。また、実際の語学学習を考えると、字幕のみを用いて動画を視聴したり、その逆に字幕を全く使わないで視聴したりする状況はあまりないだろう。字幕をどのように効果的に組み合わせ得るかという視点が第二言語動画視聴の応用研究に必要である。更に、使用する動画の語彙カバー率も学習教材の選定にあたって非常に重要な視点である。そのため、本研究では以下3つの研究設問を立てて、字幕効果を探っていく。

研究設問1

英語を外国語として学ぶ日本人大学生(及び大学院生)は、英字幕付きで英語の動画を視聴するほうが、字幕なしで同じ動画を視聴するよりも偶発的に学ばれる語彙量が多いか。

研究設問2

英語の動画を視聴する際、学ばれうる語彙量と動画の語彙カバー率との間に関連があるか。

研究設問3

英語を外国語として学ぶ日本人大学生(及び大学院生)は、英語の動画を1回目は日本語字幕付きで2回目は英字幕付きで視聴するほうが、同じ動画を1回目は字幕無しで2回目は英字幕付きで視聴するよりも偶発的に学ばれる語彙量が多いか。

4 本研究について

4.1 参加者

データ分析の対象となる本研究の参加者は、研究を紹介するフライヤーを通じて参加を募った合計148人の日本語を母語とする大学生、大学院生である。募集要項では参加条件を日本語が母語あるいは母語に準ずるレベルで、基本的に日常では日本語を使用すること、英語圏での海外経験が通算6か月未満であること、英語力がおおよそ2年以内の語学試験の結果でCEFRレベルB1からB2の範囲(英検:2級か準1級, TOEIC L&R: 550-935, TOEFL iBT: 42-94, IELTS: 4.0-6.5)であることを明示した。このレベル帯に絞った理由は、動画で十分な理解が得られることを担保するためと、おおよそ3000語レベルの語彙知識を有し、且つ英字幕を十分に読解で処理できる能力があるような参加者を募るためである。

4.2 使用動画とターゲット語彙

使用する動画は、理解しやすいように、動画内で使用される語彙に対応するイメージが多いことが予想されるドキュメンタリーかアニメーションから選ぶことを基本とし、且つ高頻度3000語レベルで90%以上の語彙カバー率があること(Durbahn et al., 2020)を条件とした。その結果、TED-Edシリーズから“The Evolution of the Book (https://www.youtube.com/watch?v=_YqYtdPUis4)”を選出した。TED-Edシリーズは科学的、学術的に興味深い話題をアニメーション形式で解説する動画だ。動画内容に関係のない最初と最後の部分を除くと、長さは3分59秒であり、総語数が618語のため、約155.1WPM(= Words Per Minute)の速さであった。このWPMでは速すぎる可能性を考慮し、ナレーションの発音が不明瞭にならない0.9倍速で動画を再生することに決定した。この結果、動画の速さは約139.6WPMとなった。また、動画で用いられている語彙カバー率の算出については、lextutorサイト上のCompleat Web VP v.2.6(Cobb, n.a.)を使用し、結果は以下の表1のとおりである。

■表1: 使用動画の語彙カバー率

頻度レベル	異なり語数 (%)	総語数 (%)	語彙カバー率 (%)
1k	205 (65.7)	485 (78.5)	78.5
2k	60 (19.2)	76 (12.3)	90.8
3k	23 (7.4)	26 (4.2)	95.0
4k	8 (2.6)	12 (1.9)	96.9
5k	5 (1.6)	5 (0.8)	97.7
6k	2 (0.6)	2 (0.3)	98.0
7k	2 (0.6)	4 (0.6)	98.6
8k	2 (0.6)	3 (0.5)	99.1
10k	1 (0.3)	1 (0.2)	99.3
12k	2 (0.6)	2 (0.3)	99.6
13k	1 (0.3)	1 (0.2)	99.8
リスト外	1 (0.3)	1 (0.2)	100.0
合計	312 (100)	618 (100)	≈100.0

注. 頻度レベル列におけるkは1000を意味する。

表1の頻度レベルの列でそれぞれ3kと2kの行をみると最頻3000語で動画内の語彙が95%カバーされ、2000語でも90.8%カバーされるため、2000から3000語レベルの語彙知識を持つ学習者にとっては、動画の十分な理解が得られると考えられる。このため、語彙レベルの観点から本動画は適切なレベルであると言える(Durbahn et al., 2020)。ターゲット語彙は表1を参考にしつつ、低頻度語や高頻度語であってもなるべく参加者が知らない可能性のある英単語を18語、品詞を限定することなく抽出した。このように選ばれたターゲット語彙は, alloy, binding, boom, carve, durable, dye, elite, flexibility, in bulk, mixture, mold, on-the-go, parchment, pop up(動詞), profound, soot, stack, walnutである。

4.3 テスト

今回使用するテストは、動画の理解を測る内容確認テスト、頻度3000語レベルまでの語彙知識量を測る語彙レベルテスト(VLT)、そして動画視聴によるターゲット語彙の学習効果を測る語彙テスト3種類である。語彙テストは、形式認識テスト、意味認識テスト、そして意味記述テストの3種類であるが、複数のテストを行う理由は、語彙習得には段階があり、1種類だけのテストでは学習者の語彙習得をきめの細かいレベルで測定することができない可能性があるためである(Webb, 2007)。

また、内容確認テストは1回目の動画視聴後にのみ行われたが、ターゲット語彙テスト3種類は事前テスト、事後テスト、そして遅延事後テストとして実施した。これらのターゲット語彙テストはすべて同じものであるが、テスト内の項目は参加者ごと、且つテストを受ける回ごとにランダム化されるようになっている。ターゲット語彙テストが難しくて参加者のモチベーションを下げることにないように、分析には含めない最頻2000語レベルまでのフィラー語も10単語それぞれのテストに含めた。

4.3.1 内容確認テスト

内容確認テストを用いる主たる理由は、参加者の意識を内容理解に向けさせ、動画理解に集中させるためである。実験の時間的制約があり、記述式の問題ではなく、全10問の多肢選択式問題をGoogle Formを用いて作成した。このテストで扱う内容は、ターゲット語彙テストにおいてヒントとはならないように、ターゲット語彙が現れる文とは関連がないように配慮した。また、内容確認テストは指示文や選択肢を含め、すべて日本語で作成された。なお、本実験での関心は内容理解というより語彙習得にあることに注意されたい。

4.3.2 語彙レベルテスト

本実験で用いる語彙レベルテストには、Webb et al.(2017)によるアップデートされた版の3000語レベルまでを使用する。各レベルは30問ずつあり、定義にあてはまる英単語を選ぶことが求められる。26語以上の正答数でそのレベルをマスターしていると考えられる。このテストを用いて本実験参加者が動画視聴するのに十分なレベルの語彙を有しているかを確認する。語彙レベルテストの結果、1000, 2000, 3000のすべての頻度レベルで26語以上正解している場合には、動画内の全単語の「95%以上」をカバーしているものとし、1000と2000のレベルのみ26語以上正解している場合は「90%以上(95%未満)」,そして前者2パターンいずれでもない場合は「90%未満」とみなすこととした。

4.3.3 形式認識テスト

形式認識テストは、提示された単語を見たこと、あるいは聞いたことがあるかを問うものである。意味に関する知識を持っている必要はないので、その意味で今回使用する3種類のターゲット語彙テストのうち一番簡単なレベルのテストであると言える。3種類のターゲット語彙テストの中で最初に実施された。見聞きしたことがあれば「はい(yes)」を選択し、そうでなければ「いいえ(no)」を選択するため、前者を1、後者を0とする二値データとして自動採点した。クロンバックの α は事前、事後、遅延事後テストでそれぞれ、 $\alpha = .65, .75, .84$ であった。以下は形式認識テストの例であるが、提示されている語は例題として出されたものである。

以下の語について、「見たこと」あるいは「聞いたこと」があれば Yes
なければ No をなるべく早く選んでください。
school

Yes

No

4.3.4 意味記述テスト

意味記述テストは、提示された単語の日本語訳、あるいは説明を書くものである。単語の形式をもとに、自力でメンタルレキシコン内の意味にアクセスし、取得する必要があるため、今回の語彙テストの中では一番難しいと考えられる (Laufer & Goldstein, 2004)。なお、本テストは後述の意味認識テストよりも先に行われたが、それは意味認識テストを先に行った場合、選択肢が意味記述テストのヒントになってしまう恐れがあるからである。意味記述テストの採点については、筆者及び、筆者同様に第二言語習得論や言語テスト理論を専攻する修士課程の学生の力を借りて採点した。採点基準は、ターゲット単語を動画内で使用される意味で回答できれば1、そうでなければ0として二値データになるよう採点した。例えば、今回の動画では pop up は動詞で「現れる」という意味で使用されているため、同じ意味の「出現する」であれば正解とし、「跳びあがる」は不正解とした。採点手順については、筆者と学生で採点基準や採点方法に関する摺り合わせの後に、全データの30%にあたる2940項目の回答を各々採点した。採点者間信頼性を示すカッパ係数は0.78となり、Plonsky and Derrick (2016) のメタ分析で示されるよりは幾分値が低いものの十分な程度と考えられるため、残りの回答データはすべて筆者が採点した。クロンバックの α は事前、事後、遅延事後テストでそれぞれ、 $\alpha = .51, .62, .67$ であった。以下は意味記述テストの例であり、提示されている語も本題として出されたものである。

以下の単語の意味を日本語で入力してください。
分からない、知らない場合は、何も回答せず、submitを押してください。

walnut

Submit

4.3.5 意味認識テスト

意味認識テストは、提示された単語の正しい意味を三択あるいは「知らない (I don't know)」から選ぶものである。提示された選択肢の中から該当する意味を選ぶため、語の形式と意味のつながりが強固である必要はない。そのため、今回の3種類の語彙テストの中では2番目の難しさであるが、前述の理由で最後に行われた。本テストも正しい意味が選べれば1、不正解あるいは「知らない」という選択肢を選んだ場合には0とする二値データとして自動採点した。クロンバックの α は事前、事後、遅延事後テストでそれぞれ、 $\alpha = .57, .68, .70$ であった。以下は意味認識テストの例であるが、提示されている語は例題として出されたものである。例題の指示文にあるように、推測の影響を出来る限り除くため、知らない単語が出

てきた場合、参加者は4つ目の選択肢である「知らない(I don't know)」を選ぶことが求められた。

以下の単語について適切な意味を選択肢(1)～(4)の中から一つ選んでください。
知らない、分からないものについては「正直に」(4)の“i don't know.”を選んでください。

school

(1) 食堂

(2) 会議室

(3) 学校

(4) i don't know.

注. 上記画像ではI don't know. とすべきところが、実験作成時のミスで小文字のiになってしまっている。

4.4 実験の流れ

実験の全工程は、2022年の9月から開始し翌年3月まで行われた。研究を紹介するフライヤーを通じて、実験参加希望者には Google Form でコンタクトに必要な個人情報や参加希望日時を記入してもらった。その後参加者に実験一日目の zoom リンク、あるいは zoom 録画を送り、実験全体の流れや実験参加は任意であること、いつでも参加辞退が可能であることを説明した。実験参加希望者には、実験参加同意書の記入と提出を求めた。実験一日目は、VLT や事前語彙テスト3種類など、Gorilla Experiment Builder (Anwyl-Irvine et al., 2020) を使用して作成した種々のテストを継続して受験してもらった。なお、実験参加者は申し込みの確定順に全6グループのうちいずれかのグループにランダムに割り振られていった。各グループの詳細については以下表2のとおりである。Exp1 と Con1 は1つ目の研究設問に答えるための実験群と統制群であり、Exp2 と Con2 は3つ目の研究設問に答えるための実験群と統制群である。なお、4グループ以外にも追加で2グループを含めている。英字幕付きで2回動画を観させ、2回目は意識を未知語に向けさせた Int グループを含めた理由については後述する。動画を一度のみ、しかもそれは字幕だけでなく音声もつけなかった Base グループを含めたのは、今回使用した動画がアニメーションだけでもどれだけ理解しやすいかを調べるためと、語彙テスト結果のベースラインにするためである。

■表2: 実験グループの詳細

グループ名	人数 (n)	1 回目の視聴	2 回目の視聴	補足
Exp1	25	英字幕付き	英字幕付き	研究設問1の実験群
Con1	25	字幕なし	字幕なし	研究設問1の統制群
Exp2	25	日本語字幕付き	英字幕付き	研究設問3の実験群
Con2	24	字幕なし	英字幕付き	研究設問3の統制群
Base	25	音声, 字幕なし	2回目の視聴なし	ベースライングループ
Int	24	英字幕付き	英字幕付き	2回目の視聴は意図的学習

実験一日目のおおよそ1週間後、参加者には再度 zoom に集まってもらう、あるいは特別な事情の際には実験の録画を共有するというかたちで、実験二日目を実施した。参加者には動画を観てもらふことを伝え、その際に動画視聴後に内容確認テストがあるので、内容理解に努めてもらうこと、それと同時に動画自体を楽しんでもらうよう伝えた。偶発的語彙学習環境を創り出すために、語彙テストがあることは伝えなかった。グループに応じた方法で動画を1回視聴してもらったのちに内容確認テストを実施した。その後、内容の再確認のため動画を再度観てもらふことを伝え、グループに応じた方法で動画を2回目の視聴を実施した。2回目の動画視聴後には、事前テストで使用したのと同じ語彙テストを事後テストとして実施した。更に1週間後の三日目に、参加者にアンケートと合わせて、語彙テストを遅延事後テストとしてメールで送信した。以上の実験工程は表3にまとめられている。

■表3: 実験の流れと内容

実験日	内容
一日目	①実験説明 ②VLTとターゲット語彙事前テスト
二日目	①動画視聴1回目 ②内容確認テスト ③動画視聴2回目 ④ターゲット語彙事後テスト
三日目	①アンケートとターゲット語彙遅延事後テスト

4.5 分析

分析には無料統計ソフトR(R core team, 2023)と関連パッケージを使用した。具体的には、前述の採点者間信頼性の計算に irr パッケージ (Gamer & Lemon, 2019)、内的一貫性指標クロンバックの α 算出に psych パッケージ (Revelle, 2023)、ルビーン検定に car パッケージ (Fox & Weisberg, 2019)、混合効果モデルの作成、検証に lme4 パッケージ (version 1.1-31; Bates et al., 2015)、データの処理統合に dplyr パッケージ (Wickham et al., 2023) である。

研究設問に答えるためのターゲット語彙テストのデータは事前テストと事後テストに限る。これは本実験では各テスト間に一週間の空きがあり、遅延事後テストまで取り組んでももらえないことが多々あったためである。分析対象となるデータは、参加者一人一人のターゲット語彙テストにつき、事前テストで正答した項目を事後テストの結果から除いたものとする。例えば、参加者Aさんが事前テストで単語Xを正解しているならば、同じAさんの事後テストの結果からは単語Xのデータを除くということである。こうすることで、もともと正解し習得されていたと考えられる単語を除き、学ばれる余地のある単語のみに絞った分析及び考察が可能になる。分析では、それぞれ研究設問の対象となるグループ (Exp1 vs. Con1 あるいは Exp2 vs. Con2) の各語彙テスト (前述のように、各参加者につき事前テストで正解だった単語を事後テストから除いたデータセット) に対して一般化線形混合モデルを構築する。目的変数である語彙テストデータは正解か不正解か、つまり1か0かの二値データであるため、二項分布となり、構築される一般化線形混合モデルはロジスティック回帰モデルとした (Brown, 2021)。

モデルに含めた固定効果 (本稿で興味のある変数) は、グループ間の違いと語彙カバー率、そしてそれらの交互作用である。これに対してランダム効果 (直接的には興味のない変数; 詳しくは Garcia, 2021 を参照) はターゲット語彙と参加者とした。モデルは Cunnings (2012) に見られるように、1つずつ固定効果を増やしてシンプルなモデルから複雑なものを構築していく流れをとり、カイ二乗検定の結果ならびに赤

池情報量基準(AIC)とベイズ情報量基準(BIC)の値などを総合的に確認することでベストフィットなモデルを選択した。なお、紙面上の都合から、結果で提示するのは最終的なモデルに限る。

5 結果と考察

5.1 記述統計量

5.1.1. 内容確認テスト

内容確認テストの記述統計量は表4のとおりである。グループに等分散性があるか確認するルビーン検定の結果、それぞれExp1とCon1のグループ間、Exp2とCon2のグループ間に等分散の仮定をしてよさそうである(それぞれ $p = .950, .437$ で対立仮説を棄却)。そのため、等分散を仮定した対応のないt検定を行い、両グループ間ともに内容確認テストの平均値の差に有意な違いがある結果となった(Exp1とCon1のグループ間は、 $t = 2.102, df = 48, p = .041$, Exp2とCon2のグループ間は、 $t = 3.140, df = 47, p = .003$)。そのため、英字幕、あるいは日本語字幕付きで第二言語動画を観るほうが、字幕がない場合に比べてより良い内容理解に貢献するといえる。

Baseグループの結果から、グループ平均点は他のどのグループよりも低いものの、今回使用したアニメーション動画は映像だけでも内容が理解しやすいものであったことがうかがえる。

■表4: 内容確認テストの記述統計量(満点10点)

グループ	<i>n</i>	<i>M</i>	<i>SD</i>
Exp1	25	7.72	1.25
Con1	25	6.92	1.38
Exp2	25	8.28	1.22
Con2	24	7.17	1.21
Base	25	6.56	1.24
Int	24	7.00	1.76

5.1.2. ターゲット語彙テスト

形式認識テスト、意味認識テスト、そして意味記述テストの得点の伸びに関する記述統計量は表5のとおりである。前述のとおり、表5のデータは、各参加者につき、事前テストで正答した項目を事後テストから除いたものであるため、平均値(*M*)が指すのは事前テストから事後テストでの得点の伸びである。どのグループ、どの種類のターゲット語彙テストに関しても得点の伸びが見受けられる。ただし、Baseグループは動画を音声も字幕もなしで視聴しているため、語彙を習得する余地のないはずであるが、実際にはどのテストにおいても得点の伸びが確認される。このうち形式認識テストにおいては、事前テストでターゲット語彙に触れているため、事後テストでもそれらが見聞きしたことのある単語として認識され、得点が上がったのだろう。意味記述テストの結果は他のグループと似た結果になっているため、事前テストから事後テストまでの一週間の期間に実験以外で学習が生じたかもしれない。Baseグループのテスト結果から、今回の学習の伸びは必ずしもすべて実験によるものではない可能性があることを注意点として挙げておく。

■表5: ターゲット語彙事前テストから事後テストの伸びに関する記述統計量

グループ	形式認識テスト			意味認識テスト			意味記述テスト		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Exp1	25	3.04	1.49	25	2.00	1.68	25	2.20	1.29
Con1	25	2.52	1.90	25	2.80	1.66	25	2.08	1.61
Exp2	24	3.21	1.72	25	2.56	2.35	25	1.92	1.55
Con2	24	3.00	1.82	24	2.79	1.79	24	1.42	1.25
Base	25	2.24	2.01	25	1.36	1.52	25	2.04	1.90
Int	23	3.61	2.08	23	2.43	1.95	24	1.96	1.73

注. 実験ソフトの問題でデータがとれていなかった参加者のテストデータは除いたため、参加者が少なくなっている場合がある。

5.2 研究設問1&2

Exp1とCon1グループのターゲット語彙テストデータに対して一般化線形混合モデルを当てはめた結果は表6のとおりである。採用されたモデルの固定効果はグループの違いとカバー率を含むが、これら2つの変数間の交互作用はデータ不足のためモデル化できなかった。

■表6: Exp1とCon1グループに対する一般化線形混合モデル分析の結果

形式認識テスト	固定効果	オッズ比	SE	95%信頼区間	<i>z</i>	<i>p</i>
	(Intercept)	0.59	0.30	[0.22, 1.62]	-1.02	.308
	グループ [Exp1]	2.64	1.05	[1.21, 5.77]	2.44	.015
	カバー率 [>90%]	1.61	0.84	[0.58, 4.46]	0.92	.355
	カバー率 [>95%]	2.51	1.33	[0.89, 7.08]	1.74	.082
	ランダム効果	分散	SD			
	参加者	0.80	0.89			
ターゲット語彙	0.72	0.85				
意味認識テスト	固定効果	オッズ比	SE	95%信頼区間	<i>z</i>	<i>p</i>
	(Intercept)	0.39	0.18	[0.16, 0.95]	-2.07	.039
	グループ [Exp1]	0.74	0.26	[0.37, 1.49]	-0.84	.403
	カバー率 [>90%]	1.06	0.52	[0.41, 2.77]	0.13	.898
	カバー率 [>95%]	1.92	0.90	[0.77, 4.81]	1.40	.163
	ランダム効果	分散	SD			
	参加者	0.77	0.88			
ターゲット語彙	0.55	0.74				
意味記述テスト	固定効果	オッズ比	SE	95%信頼区間	<i>z</i>	<i>p</i>
	(Intercept)	0.06	0.03	[0.02, 0.17]	-5.13	<.001
	グループ [Exp1]	1.33	0.40	[0.74, 2.40]	0.95	.344
	カバー率 [>90%]	1.26	0.55	[0.54, 2.95]	0.54	.587
	カバー率 [>95%]	3.69	1.55	[1.62, 8.40]	3.11	.002
	ランダム効果	分散	SD			
	参加者	0.40	0.63			
ターゲット語彙	2.56	1.60				

英字幕が有り無しというグループの違いが有意な影響を持ったのは形式認識テストのみにおいてであった。英字幕があるグループのほうが事後テストで約2.6倍ターゲット語彙の形式を認識できることが分かる($p = .015$)。また、カバー率については、95%以上のカバー率が意味記述テストにおいてのみ有意に肯定的な影響を持つことが分かる。つまり、動画内の単語の95%以上を知っている参加者は、約3.7倍ターゲット語彙の意味を筆記テストで正しく産出することができるということである($p = .002$)。まとめると、英字幕があることで結果が良くなるのは形式認識テストのみで、95%以上のカバー率があると、英字幕の有無に関わらず、意味記述テストの結果がより良いということである。

形式認識テストにおいて英字幕が有意に好影響を持ったのは、インプットを聴覚だけでなく視覚も合わせて受けることで、未知語が際立つのではないかと考えられる。しかし、意味認識テスト及び意味記述テストのモデルから、英字幕はターゲット語彙を未知語として際立たせるだけで、英字幕が無い場合よりもターゲット語彙の意味を推測しやすくさせることはなさそうである。

この結果は、偶発的語彙習得に対する第二言語字幕の効果を発見しなかった先行研究(e.g., Raine, 2013; Rodgers, 2013)を支持するものである。第二言語字幕の効果がなかった理由としてまず考えられるのは、本実験の動画が言語音声面で難しかったわけではないことだ。第二言語字幕は特にリスニングの苦手な学習者に対してセグメンテーション(明確な区切りのない音声言語を語彙単位で区切っていくこと; Rost, 2016)の手助けとなるとされるが(Charles & Trenkic, 2015), Ted-Edのナレーターは訛りなく非常に分かりやすい発音で話しており、話の速さも約137.1WPMであったため、英字幕のないグループの参加者にとってもセグメンテーションがしやすく、それによりリスニング面で大きな負担がなかったのかもしれない。つまり、本実験の対照群の結果は、仮に本実験の動画の発話速度が速かったり、訛りが強かったりするなど音声的に難しかった場合と比べて、より良いものであった可能性がある。しかし、この説明は内容確認テストの結果と照らし合わせるといささか相容れない。というのも、実際に英字幕グループは内容確認テストの結果が英字幕なしのグループに比べて有意なレベルで良いため、英字幕の恩恵が内容理解にあったことを示唆する。それはつまり、英字幕があったことにより英字幕グループの方がセグメンテーションがより良くできた可能性がある。そのため、動画の音声面の易しさにより英字幕が偶発的語彙学習を促進しなかったという説明は必ずしも適切でないかもしれない。

そこで2つ目に考えられる理由としては、英字幕の有無に関わらず、ターゲット語彙に向けられた注意の量、つまりはターゲット語彙の言語処理量が両グループで同じくらいであった可能性だ。本実験では、あらかじめ参加者に動画視聴の際は内容理解に努めて動画を楽しむよう指示を出しているため、その差は参加者ごとに違えど「英単語を学ぼう」という意識はあまりなかったはずである。すると、実験群と対照群のどちらにおいても、有限である注意(Mayer, 2021)がほとんど内容理解に使われ、ナレーションであれ英字幕であれ言語的側面(未知語の意味)に向けられる注意の総量がどちらのグループでも違いはほとんどなかったのかもしれない。

次に研究設問の2つ目となるカバー率と語彙習得の関係性について議論をしていく。本実験ではVLTの結果をもとに、各参加者が動画内に出てくる単語をどれだけ知っているかを、95%以上、90%以上(95%未満)、90%未満の3レベルで分けた。このカバー率の情報を統計モデルに組み込んだ結果、意味記述テストにおいて有意に肯定的な影響が確認された。これはつまり、動画内で出てくる単語の95%以上を知っている場合、未知語の意味を正しく推測し、それを自力で再生できる度合いが大きくなるということである。動画内で出てくる単語の多くを知っているということは必然的に未知語の割合が少ないため、限られた注意を限られた数の未知語に向け、イメージを含む周りの文脈から正しく未知語の意味を推測できることを示唆する。逆に言えば、カバー率が95%未満の場合は、多くの未知語が出てきて内容理解に多くの注意量を割いてしまうため、未知語の推測どころではなくなってしまっている可能性がある。今回のモデルでは残念ながらこのカバー率とグループの違いの交互作用をそもそも含むことができなかったため、95%以上のカバー率があり、加えて英字幕があればより多くの語彙量を習得できたかについては今後の課題である。

5.3 研究設問3

Exp2とCon2グループのターゲット語彙テストデータに対して一般化線形混合モデルを当てはめた結果は表7のとおりである。1回目の視聴で日本語字幕あり、なしというグループの違いはどの種類のターゲット語彙テストにおいても有意な結果とはならなかった。つまり、日本語字幕と英字幕を連続的に使用した場合と、1回目に字幕がなしで2回目に英字幕付きで動画を視聴した場合に、字幕の違いは偶発的語彙習得に影響を与えなかったことを意味する。また、カバー率については、研究設問1&2のモデルと同様に、95%以上のカバー率が意味記述テストにおいてのみ有意に肯定的な影響を持つことが分かる。動画内の単語の95%以上を知っている参加者は、約4.3倍ターゲット語彙の意味を筆記テストで正しく産出することができるということである($p = .001$)。グループの違いとカバー率の2つの変数間の交互作用はデータ不足のためモデル化できなかった。

■表7: Exp2とCon2グループに対する一般化線形混合モデル分析の結果

	固定効果	オッズ比	SE	95%信頼区間	z	p
形式認識テスト	(Intercept)	1.17	0.63	[0.41, 3.34]	0.29	.775
	グループ [Exp2]	0.78	0.30	[0.37, 1.66]	-0.64	.522
	カバー率 [>90%]	2.14	1.10	[0.79, 5.84]	1.49	.135
	カバー率 [>95%]	2.60	1.38	[0.92, 7.36]	1.80	.072
	ランダム効果	分散	SD			
	参加者	0.67	0.82			
	ターゲット語彙	0.69	0.83			
意味認識テスト	固定効果	オッズ比	SE	95%信頼区間	z	p
	(Intercept)	0.38	0.17	[0.16, 0.93]	-2.12	.034
	グループ [Exp2]	0.83	0.27	[0.44, 1.58]	-0.56	.575
	カバー率 [>90%]	0.98	0.44	[0.41, 2.36]	-0.04	.972
	カバー率 [>95%]	1.50	0.69	[0.61, 3.68]	0.88	.377
	ランダム効果	分散	SD			
	参加者	0.58	0.76			
ターゲット語彙	0.53	0.73				
意味記述テスト	固定効果	オッズ比	SE	95%信頼区間	z	p
	(Intercept)	0.04	0.02	[0.01, 0.12]	-5.76	< .001
	グループ [Exp2]	1.37	0.39	[0.78, 2.39]	1.10	.272
	カバー率 [>90%]	1.81	0.79	[0.77, 4.27]	1.35	.176
	カバー率 [>95%]	4.32	1.93	[1.80, 10.37]	3.28	.001
	ランダム効果	分散	SD			
	参加者	0.14	0.37			
ターゲット語彙	2.22	1.49				

Exp2とCon2のグループの結果も、Exp1とCon1グループの結果同様、偶発的語彙習得に影響を与えるのは、字幕の有無というより第二言語での動画視聴そのものようである。この結果は同じく第一言語と第二言語の字幕の連続的な使用について研究したWi and Boers(2023)と似た結果である。しかし、本実験でExp2とCon2で字幕効果がなかった理由としては、Wi and Boers(2023)で述べられているようなサンプルサイズの問題ではなく、研究設問1で述べたような注意量が関係しているかもしれない。

あるいは、Exp2とCon2のテスト結果に違いがないのは、日本語字幕が内容理解には好影響を持つが、

英語字幕と併用されたとしても偶発的語彙学習には影響がない可能性もあり得る。日本語字幕と英語字幕の連続使用を2つ目の研究設問として調査したのは、日本語字幕が内容理解を促進し、英語字幕付き動画を2回目の視聴をしたときに、未知語の形式と意味を結びつけやすくなるとの予想のためである。実際に実験群のほうが統計的に有意なレベルで内容理解が優れているが、動画視聴による偶発的語彙学習に内容理解の度合いはあまり関係せず、偶発的語彙学習には、後に触れるカバー率のような別の要素が大きな影響を持つのかもしれない。

最後にもう一つの可能性を挙げる。本実験の記述統計量を確認すると(表5参照)、Exp1, Exp2, Con1, Con2のすべてのグループが似通った結果であるのが分かる。このことは、本実験で偶発的語彙習得の量にグループ間の違いが見つけれなかったもう一つの可能性を指し示すものである。それは視聴回数である。視聴回数に絡めながら本実験の流れを振り替えると、どのグループも指定の条件で1回目の視聴を終えた後には内容確認テストを受けている。内容確認テストの結果は、それぞれの実験で字幕付きのグループのほうが統計的により良いものであり、それは1回目の視聴時点で字幕の条件に応じて参加者に何らかの質的な違いがあったことを示唆するものである。しかし、ターゲット語彙テストについては1回目の動画視聴終了時点で実施したのではなく、もう一度参加者に指定の条件で動画を視聴してもらった後に実施した。2回目の視聴チャンスがあったことで、2回目に字幕がなかったグループ(Con1とCon2)の参加者は2回目の視聴でより集中して動画を視聴し、反対に1回目に字幕があったグループの参加者は既に十分な内容理解が得られていると考え、2回目の視聴がおろそかになってしまった可能性が挙げられる。結果的に実験群と対照群の学習が似通ったのかもしれない。

カバー率については、2つ目の研究設問と同様に、動画内で出てくる単語の95%以上を知っている場合、意味記述テストの結果がより良くなることが確認された。結果が一つ目の研究設問と似ているため、音声面ですれほど難しくない動画を視聴することによる偶発的語彙学習で大切なことは、字幕の有無というより、動画内で出てくる語彙を95%以上知っていることかもしれない。

6 まとめと課題

本研究では、第二言語動画を字幕付きで観ることが偶発的語彙習得を促進するかといったテーマで研究を進めてきた。最後に各研究設問をまとめつつ、教育的意義や今後の研究の課題をまとめていく。

まず1つ目の研究設問では、日本人大学生が、英字幕付きで英語の動画を視聴する際に、字幕なしで同じ動画を視聴する場合よりも偶発的に学ばれる語彙量が多いかを問うた。本実験の結果は予想に反して、偶発的に学ばれる語彙量は英字幕の有無に関係しなかった。英字幕の有無が影響を与えなかった理由としては、語彙に向けられる注意量が実験群でも対照群でも同じだった可能性や、使用された動画の音声的な難しさの観点からも議論した。ただし、本実験の結果は、英字幕の効果を全否定するものではない。実際に本実験では実験群の内容理解は有意に対照群よりも良かったし、英字幕の効果は偶発的語彙習得だけでなく、発音や語用論的知識など多角的に研究されてきている(Wei & Fan, 2022)。そのため、授業内あるいは自主学習において第二言語で動画を視聴する場合には、細部に至る内容理解を得るために第二言語字幕を付けることを積極的に選択肢に入れてみるのも良いだろう。しかし、英字幕を付けると英文を読むだけになり、動画の音声面に注意がいなくなってしまうのではないかと、という危惧もある。確かに、初級レベルの学習者によってはその可能性が否めないが、中級レベル以上の学習者であれば第二言語字幕付きの動画を観ている際、リスニングにも注意が向いていることを示す研究もある(Charles & Trenkic, 2015)。そして動画を選ぶ際に重要なことは、面白そうな動画をやみくもに選ぶのではなく、内容理解に関しては動画内で出てくる語彙の90%以上(Durbahn et al., 2020)を、そして偶発的な語彙の習得に関しては95%以上を学習者が知っているような動画を選択することであると言える。

また、第二言語字幕付きの動画を使って偶発的に単語を学ばせるのではなく、意図的に学ばせるのはど

うかといった考えもあるだろう。本実験で含めた6つ目のIntグループは、2回目の視聴で未知語があればそれに注意を向けるよう指示されたグループであるが、記述統計量を見る限り(5.1.2参照)、結果は形式認識テストを除いて偶発的学習グループ群の結果とあまり違いがなさそうである。テストで測定された語彙量に違いはなさそうだが、偶発的語彙学習の良さは明示的に教えるのが困難な単語の意味が学習されたり、文脈と合わせた使用法や意味が獲得されたりすることだとされている(Schmitt, 2010)。もし教室内外での語彙指導及び学習が意図的学習に偏るものであれば、字幕を組み合わせた第二言語動画視聴による偶発的学習の比重を増やすのも良いかもしれない。もちろん、偶発的学習と意図的学習は相互排他的なものではなく、どちらの学習タイプもうまく使われていく必要がある。

3つ目の研究設問では、日本語字幕と英字幕の連続的な使用を、1回目は字幕が無い状態から2回目は英字幕付きで視聴する場合と対照させて、偶発的に学ばれる語彙量に違いがあるかを検証した。残念ながらこちらの場合も前者の実験群と後者の対照群では語彙習得量に違いがなく、カバー率が良ければ語彙習得も良くなることが分かった。日本語字幕と英字幕の連続的な使用の効果がなかった理由は、1つ目の研究設問同様、字幕の有無に関わらず、言語面に向けられた注意量が同じであることや、使用した動画が音的に難しくなかったことが考えられる。さらに1つ目の研究設問のグループの記述統計量も踏まえて、3つ目の研究設問グループ間で偶発的語彙習得の量に違いがなかったもう一つの理由として、視聴回数と回数に応じた集中度合いの要因も考えられた。今回研究設問の対象となった4グループ(Exp1, Exp2, Con1, Con2)はすべて動画を2回視聴している。1回目の視聴時点では、日本語であれ英語であれ字幕のあるグループのほうが内容理解度が有意に良かったので、偶発的語彙習得の量もまた多かつたかもしれない。しかし、1回目に字幕があったグループの参加者は、最初の視聴で既に十分な理解を得られているため、2回目の視聴時は動画視聴にあまり集中せず、逆に1回目に字幕のなかったグループの参加者は内容理解度が低く、2回目の視聴も集中して取り組んだ可能性が考えられた。そのため、結果的にすべてのグループで似たような語彙習得量になったのかもしれない。繰り返し学習することが好影響を与えるのは語彙習得も例外ではなく、語彙に何回も触れるほど習得の可能性も高まる(Uchihara et al., 2019)。そのため、クラス内や自主学習の環境では第二言語で動画を観るときには、それが本実験で使用したような短いマテリアルである場合、繰り返し視聴することが良いだろう。その際、英字幕や日本語字幕をうまく活用していくことに語彙習得にとどまらない意義があると考え。ただし、視聴回数が増えるにしたがって、学習者の集中力やモチベーションが低下しないように、動画を立て続けに視聴するというよりは、視聴回数に応じて注目すべきポイントを変えたり、言語活動を間に含めたりするなどの工夫が必要だと考える。もちろん、字幕の効果的な使用法に関する研究は今後更に必要であるのは言うまでもない。今回、統計モデルに含むことのできなかつた字幕の有無とカバー率の交互作用の問題は興味深い。つまり、動画内に出てくる単語の95%以上を知っている場合に更に英字幕があることで偶発的語彙学習が英字幕のない場合より良くなるのか、研究の価値がある。更に、学習者の視聴回数に応じたふるまいの違いについては視線計測機などの使用が見込まれる(Teng, 2021)。ただし、ヘッドマウント型の視線計測機は参加者になるべく顔を動かさないことを求める厳密な実験となるため、参加者の現実世界における振る舞いまでは考慮に入れることができないかもしれない。参加者へのインタビューやアンケートを通じて、動画視聴に関する率直な意見を聞くのもいいだろう。この点において、実際の教育現場の先生方によるアクションリサーチ(Burns, 2011)にも期待したい。

教室内外で使用できる技術は字幕に限らず、学習に有効とされる様々なコンピューター支援言語学習の道具がある(Chun, 2011)。第二言語視聴と字幕に関する本研究の結果が、現場の指導者の方々の指導レパートリーの一つになったり、研究者が更なる研究を考案したりすることに寄与できることを願う。

謝辞

本研究の機会をくださった公益財団法人 日本英語検定協会のご担当者様、選考委員会の先生方に、厚くお礼を申し上げます。中でも、研究助言者の西垣知佳子先生(千葉大学)、私が修士課程に在

籍していたときの指導教官であった根岸雅史先生(東京外国語大学),そして私の師匠である内原卓海先生(東北大学)の3人の先生方には非常に貴重で有益なご助言,ご指導を賜りましたこと,感謝のしようがありません。また,本実験の参加者を集めるのに協力して下さった各先生方にもこの場をお借りしてお礼申し上げます。実験に参加して下さった参加者の皆様にも多大なお礼を忘れません。

引用文献

- Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior research methods*, 52(1), 388-407. <https://doi.org/10.3758/s13428-019-01237-x>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models Using lme4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Birulés-Muntané, J., & Soto-Faraco, S. (2016). Watching subtitled films can help learning foreign languages. *PLoS ONE*, 11(6). Article e0158409. <https://doi.org/10.1371/journal.pone.0158409>
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021). *Introduction to meta-analysis* (2nd ed). John Wiley & Sons.
- Brown, V. A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science*, 4(1), 1-19. <https://doi.org/10.1177/2515245920960351>
- Burns, A. (2011). Action research in the field of second language teaching and learning. In E. Hinkel (Ed), *Handbook of research in second language teaching and learning volume 2*. Routledge.
- Charles, T., & Trenkic, D. (2015). Speech segmentation in a second language: The role of bimodal input. In Y. Gambier, A. Caimi & C. Mariotti (Eds), *Subtitles and language learning: Principles, strategies and practical experiences* (pp. 173-198). Peter Lang.
- Clenton, J., & Booth, P. (2021). *Vocabulary and the four skills: Pedagogy, practice, and implications for teaching vocabulary*. Routledge.
- Chacón-Beltán, R., Abello-Contesse, C., & del Mar Torreblanca-López, M. (2010). Insights into non-native vocabulary teaching and learning. *Multilingual Matters*.
- Chun, D. M. (2011). Computer-assisted language learning. In E. Hinkel (Ed), *Handbook of research in second language teaching and learning volume 2*. Routledge.
- Cobb, T. (n.d.). Compleat Web VP (Version 2.6) [Computer Program]. <https://www.lexutor.ca/vp/comp/>
- Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research*, 28(3), 369-382. <https://doi.org/10.1177/0267658312443651>
- Dizon, G., & Thanyawatpokin, B. (2021). Language learning with Netflix: Exploring the effects of dual subtitles on vocabulary learning and listening comprehension. *Computer Assisted Language Learning Electronic Journal*, 22(3), 52-65.
- Durbahn, M., Rodgers, M., & Peters, E. (2020). The relationship between vocabulary and viewing comprehension. *System*, 88, 1-13. <https://doi.org/10.1016/j.system.2019.102166>
- Ellis, R. (2009). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 3-26). *Multilingual Matters*.
- Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J., & Reinders, H. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching*. *Multilingual Matters*.
- Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (3rd ed). Sage, Thousand Oaks CA.
- Frumuselu, A. D., De Maeyer, S., Donche, V., & Plana, M. D. M. G. C. (2015). Television series inside the EFL classroom: Bridging the gap between teaching and learning informal language through subtitles. *Linguistics and Education*, 32, 107-117. <https://doi.org/10.1016/j.linged.2015.10.001>
- Gamer, M., & Lemon, J. (2019). irr: Various Coefficients of Interrater Reliability and Agreement. R package version 0.84.1, <<https://CRAN.R-project.org/package=irr>>.
- Garcia, G. D. (2021). *Data visualization and analysis in second language research*. Routledge.
- Graham, S., & Zhang, P. (2021). Vocabulary and listening: Current research, tools, and practices. In J. Clenton & P. Booth (Eds), *Vocabulary and the four skills: Pedagogy, practice, and implications for teaching vocabulary*. Routledge.
- Hao, T., Sheng, H., Ardasheva, Y., & Wang, Z. (2022). Effects of dual subtitles on Chinese students' English listening comprehension and vocabulary learning. *The Asia-Pacific Education Researcher*, 31(5), 529-540. <https://doi.org/10.1007/s40299-021-00601-w>
- Hinkel, E. (2011). *Handbook of research in second language teaching and learning volume 2*. Routledge.
- Hsieh, Y. (2020). Effects of video captioning on EFL vocabulary learning and listening comprehension. *Computer Assisted Language Learning*, 33(5-6), 567-589. <https://doi.org/10.1080/09588221.2019.1577898>
- Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In C. P. Robinson (Ed), *Cognition and second language instruction* (pp. 258-276). Cambridge University Press.
- Jelani, N. A. M., & Boers, F. (2018). Examining incidental vocabulary acquisition from captioned video: Does test modality matter? *ITL-International Journal of Applied Linguistics*, 169(1), 169-190. <https://doi.org/10.1075/itl.00011.jel>

引用文献

- Kang, E. (2019). Enhancing Korean EFL learners' vocabulary learning and listening comprehension through video captions. *STEM Journal*, 20(2), 91-108. doi: 10.16875/stem.2019.20.2.91
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, 54(3), 399-436. <https://doi.org/10.1111/j.0023-8333.2004.00260.x>
- Mayer, R. E. (2005). Introduction to multimedia learning. In R. E. Mayer (Ed), *The Cambridge handbook of multimedia learning* (pp. 1-18). Cambridge University Press.
- Mayer, R. E. (2021). *Multimedia learning* (3rd ed). Cambridge University Press.
- Montero Perez, M. (2022). Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching*, 55(2), 163-192. doi:10.1017/S0261444821000501
- Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology*, 18(1), 118-141. <http://ilt.msu.edu/issues/february2014/monteroperezetal.pdf>
- Montero Perez, M., Peters, E., & Desmet, P. (2018). Vocabulary learning through viewing video: The effect of two enhancement techniques. *Computer Assisted Language Learning*, 31(1-2), 1-26. <https://doi.org/10.1080/09588221.2017.1375960>
- Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, 41(3), 720-739. <http://dx.doi.org/10.1016/j.system.2013.07.013>
- Peters, E. (2019). The effect of imagery and on-screen text on foreign language vocabulary learning from audiovisual input. *TESOL quarterly*, 53(4), 1008-1032. <https://doi.org/10.1002/tesq.531>
- Plonsky, L. (2015). *Advancing quantitative methods in second language research*. Routledge.
- Plonsky, L., & Derrick, D. J. (2016). A meta-analysis of reliability coefficients in second language research. *The Modern Language Journal*, 100(2), 538-553. <https://doi.org/10.1111/modl.12335>
- 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>
- Plonsky, L., & Oswald, F. L. (2015). Meta-analyzing second language research. In L. Plonsky (Ed). *Advancing quantitative methods in second language research* (pp. 106-128). Routledge.
- R Core Team (2022). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <<https://www.R-project.org/>>.
- Raine, P. (2013, January). Incidental increase in depth of vocabulary knowledge through the viewing of subtitled, authentic videos [Proceedings]. JALT 2012 Conference. https://www.researchgate.net/publication/315658011_Incidental_Increase_in_Depth_of_Vocabulary_Knowledge_Through_the_Viewing_of_Subtitled_Authentic_Videos
- Revelle, W. (2023). psych: Procedures for Psychological, Psychometric, and Personality Research. Northwestern University, Evanston, Illinois. R package version 2.3.6, <<https://CRAN.R-project.org/package=psych>>.
- Robinson, P. (2001). *Cognition and second language instruction*. Cambridge University Press.
- Rodgers, M. P. (2013). *English language learning through viewing television: An investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions* [Doctoral dissertation, Victoria University of Wellington]. <http://researcharchive.vuw.ac.nz/handle/10063/2870>
- Rost, M. (2016). *Teaching and researching listening* (3rd Ed). Routledge.
- Schmitt, N. (2010). Key issues in teaching and learning vocabulary. In R. Chacón-Beltán, C. Abello-Contesse, M. del Mar Torreblanca-López (Eds), *Insights into non-native vocabulary teaching and learning*. Multilingual Matters.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913-951. <https://doi.org/10.1111/lang.12077>
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26-43. <https://doi.org/10.1111/j.1540-4781.2011.01146.x>
- Ted-Ed. (2016, April). The evolution of the book [Video]. YouTube. <https://www.youtube.com/watch?v=YqYtdPUis4>
- Teng, F. (2018). Incidental vocabulary learning for primary school students: the effects of L2 caption type and word exposure frequency. *The Australian Educational Researcher*, 46, 113-136. <https://doi.org/10.1007/s13384-018-0279-6>
- Teng, M. F. (2021). *Language learning through captioned videos: Incidental vocabulary acquisition*. Routledge.
- Teng, M. F. (2022). Incidental L2 vocabulary learning from viewing captioned videos: Effects of learner-related factors. *System*, 105, 1-12. <https://doi.org/10.1016/j.system.2022.102736>
- Uchiyama, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning*, 69(3), 559-599. <https://doi.org/10.1111/lang.12343>
- Vanderplank, R. (2016). *Captioned media in foreign language learning and teaching: Subtitles for the deaf and hard-of-hearing as tools for language learning*. Springer.
- Van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L2 and L2 listening comprehension: The same or different from reading comprehension? *Applied linguistics*, 34(4), 457-479. <https://doi.org/10.1093/applin/ams074>

引用文献

- Wang, A., & Pellicer - Sánchez, A. (2022). Incidental vocabulary learning from bilingual subtitled viewing: An eye - tracking study. *Language Learning*, 72(3), 765-805. <https://doi.org/10.1111/lang.12495>
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46-65. <https://doi.org/10.1093/applin/aml048>
- Webb, S. (2020). Incidental vocabulary learning. In S. Webb (Ed), *The Routledge handbook of vocabulary studies* (pp. 225-239). Routledge.
- Webb, S. (2020). *The Routledge handbook of vocabulary studies*. Routledge.
- Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.
- Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL-International Journal of Applied Linguistics*, 168(1), 33-69. <https://doi.org/10.1075/itl.168.1.02web>
- Wei, R., & Fan, L. (2022). On-screen texts in audiovisual input for L2 vocabulary learning: A review. *Frontiers in Psychology*, 13. doi: 10.3389/fpsyg.2022.904523
- Wesche, M. & Paribakht, T. (2010). *Lexical Inferencing in a First and Second Language: Cross-linguistic Dimensions*. Multilingual Matters. <https://doi.org/10.21832/9781847692245>
- Wi, I., & Boers, F. (2023). Sequential Use of L1 and L2 Captions: Exploring the Benefits for Vocabulary Acquisition. *TESOL Quarterly*. <https://doi.org/10.1002/tesq.3243>
- Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). dplyr: A Grammar of Data Manipulation_. R package version 1.1.2, <<https://CRAN.R-project.org/package=dplyr>>.