



大規模言語モデルを利用した英語の移動動詞選択要因の探索的研究

森下, 裕三

(Citation)

統計数理研究所共同研究レポート, 469:27-40

(Issue Date)

2024-03-25

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/0100487704>

(URL)

<https://hdl.handle.net/20.500.14094/0100487704>



大規模言語モデルを利用した英語の移動動詞選択要因の探索的研究

森下 裕三(桃山学院大学)

An Exploratory Study into the Selection Factors of Motion Verbs in English Using a Large Language Model

MORISHITA, Yuzo (St. Andrew's University)

Abstract

In this study, I examined whether LLMs (large language models) can be used to classify verbs by meaning. Furthermore, I clarified what factors an LLM focuses on to distinguish such meanings by investigating an internal mechanism called “attention.” Although some of the results differed from the claims made in previous research, they suggested that LLMs could be used as exploratory tools for further research in linguistics.

キーワード

英語、語の意味、大規模言語モデル、移動表現

1. はじめに

語の意味にかかわる研究は、多くの言語学者によって議論が続けられ、アプローチの方法も多様性を見せている。たとえば、語の意味を概念として捉えるものもあれば、語の意味を文脈として捉えるものもある。これらは、範列関係 (**paradigmatic relation**) に注目するものと連辞関係 (**syntagmatic**) に注目するものとして捉えることもできるかもしれない (cf. Geeraerts 2010: 165–178)。本節では、従来の研究では内省に頼って研究が進められてきた語の意味について、文脈に依拠した手法を応用しながら分析するための方法論を示したい。

語の意味を概念として捉えるということは、語そのものの意味に注目することで、ある語が他の語とどのように異なるのかを明らかにしようというものである。たとえば、英語の移動動詞には **come** や **walk** など 100 以上の動詞がある。詳細は 2 節で述べるが、**come** は直示性という経路を語彙化しており、**walk** は移動の様態を語彙化している (Talmy 1985, 2000)。他にも **stingy** と **thrifty** という語を比べると、どちらの形容詞も「お金をあまり使わない」といった性質をあらわすという点では共通している。しかし、**stingy** には否定的な評価がとれない、**thrifty** には肯定的な評価がともなうと考えられている (Fillmore 1982: 125)。これらの語が語彙化している要素や、これらの語がもつ評価は周囲にどのような語が生起していようが基本的に変わらない。

語を文脈から捉えるということは、注目している語だけでなく、その語の周囲に生起している語を観察することによって、対象とする語の性質を明らかにしようというものである。この考えは、分布仮

説 (distributional hypothesis) と呼ばれ、Firth (1957) による主張に端を発している。この系譜にある研究としては、Louw (1993) や Stubbs (1995) らによる意味的韻律 (semantic prosody) のように現代の言語研究において広く知られているものもある。たとえば、cause という語は次のように動詞や名詞としてさまざまな語と共起する。

- (1) a. East German restriction which caused today's trouble.
b. ... a certificate showing the cause of death. (Stubbs 1995: 31)

たしかに cause という語はさまざまな語と共起するが、多くの場合、trouble や death といった否定的な意味合いの語と共起することが確認されている。つまり、cause という語の周囲にどのような語が生じやすいのかを考慮することで、cause という語の意味にアプローチしようというのである。

本稿では、この分布意味論という考え方をさらに推し進め、大規模言語モデルを利用した手法を採用する。このような仕組みを取り入れることによって、語の意味の違いを言語学の外側から観察し、語の意味にとって重要な要素を探索的に調査できるのではないだろうか。本研究では、具体的な例として移動動詞のうち、(2a) および (2b) にあるような直示性を語彙化したものと、(3a) および (3b) にあるような移動の様態を語彙化したものとの区別を扱う。

- (2) a. And the people came by the thousands and thousands. (COCA, 2017)
b. He took the elevator to the street and went straight to O'Flanagan's Bar. (COCA, 1994)
- (3) a. [...] the planes that were going to fly from Great Britain to here. (COCA, 2006)
b. Pencil Whiskers swaggered over to her [...]. (COCA, 2008)

英語の移動動詞には、(2) や (3) に示した自動詞のもの以外にも John entered the building. という例にも見られるように他動詞のものもある。この他動詞の enter を自動詞の come や climbed と置き換えることはできない。一方、(2) や (3) に示した例では、次のように動詞を置き換えられることも少なくない。

- (4) a. He took the elevator to the street and swaggered straight to O'Flanagan's Bar.
b. Pencil Whiskers went over to her [...].

つまり、直示的移動動詞と移動様態動詞の違いは、統語的な違いではなく純粋に意味的な違いである。分布仮説が正しければ、大規模言語モデルによって、どういう時には直示的な動詞が選択され、どういう時には様態を語彙化した動詞が選択されるのかがある程度まで正確に予測できるはず

である。また、このような予測のために、大規模言語モデルがどのような情報を手がかりとしているのかを確認することができれば、これらの動詞の違いが、どのような仕組みに由来するのかを明らかにすることにも繋がり、語の意味の研究をさらに進めることができるのではないだろうか。本研究では、2018年にGoogleによって発表されたBERTと呼ばれる大規模言語モデルを利用して、以上のような問いに対して探索的な試みをおこなう。

BERTとは、注目している語だけでなく、その周囲に生起する語も考慮した分析が可能な大規模言語モデルである。後述する注意機構（attention）と呼ばれる手法によって注目する語から離れた位置に生起する語の情報も考慮に入れることができ、さまざまなタスクにおいて高い性能を示している。そのため、言語研究への応用にも耐えうると考えられる。

本研究の構成は次のとおりである。2節では、移動表現についての先行研究のうち、本研究とかわりの深い部分について確認し、さらに、語の選択にかかわる先行研究、および大規模言語モデルの概略を示す。3節では、本研究における、データの収集方法や、実験の内容を確認する。その後、4節では、実験結果からどのような示唆を得ることができたのかについて、言語学との関係を議論する。最後に5節では、改めて言語学的な視点に立ち返り、言語の研究にとって大規模言語モデルがどのような意義をもつのかについて述べる。

2. 先行研究

本節では、直示的移動動詞と移動様態動詞の違いについて、本研究とかわりの深い部分のみを確認していくことにする。さらに、従来の言語研究で広く利用されてきた統計的な手法と比較しながら、深層学習を言語研究に応用するための仕組みのうち、特にBERTや注意機構と呼ばれるものについても概観する。

2.1 移動表現の類型論と英語の移動動詞

言語学において類型論は以前から議論が続けられており、代表的なものにGreenberg (1963)による基本語順の類型論がある。また、言語を類型化しようとする場合、意味に注目した類型化も可能である。意味に注目した類型論として、多くの研究者によって研究が進められているものに移動表現の類型論がある。これは、Talmy (1985, 1991, 2000)によって提唱され、移動物 (FIGURE)、移動 (MOTION)、経路 (PATH)、基準物 (GROUND) という移動事象に必須の要素のうち、どの概念をどの要素によって語彙化するのかという点に着目した類型である。移動事象にとって重要な経路 (PATH) が語彙化される要素によって、世界の諸言語を衛星枠付け言語 (satellite-framed language) と動詞枠付け言語 (verb-framed language) に分類するという類型論である。衛星 (satellite) と呼ばれる不変化詞や前置詞によって語彙化されることが多いことから、ロマンス諸語以外の、英語を含む多くのヨーロッパの言語は衛星枠付け言語とされている。

この衛星枠付け言語のひとつである英語の移動表現では、次のような例が問題となる。

- (5) a. John came into the restaurant.
b. Mary went out of the building.

これらの例では、into や out of といった前置詞だけでなく、動詞である come や go によっても直示性という経路が語彙化されている。また、これらの表現では、前置詞句が動詞に後続しているということから、統語的には (6) に示す例と同じだということになる。

- (6) a. John ran into the restaurant.
b. Mary walked out of the building.

英語の話者は、ある人物がレストランへ入っていく場面を叙述する時には、come か run のどちらかの動詞を選ばなければならない。つまり、(5a) のように直示的移動動詞を使うか、それとも (6a) のように移動様態動詞のどちらを使うのかという選択に迫られるのである。そして、この謎を解き明かすことが本研究の課題でもある。

2.2 二項ロジスティック回帰分析から大規模言語モデルを利用した分析へ

さきほど述べたように、言語研究において、語や構文など 2 つの異なる選択肢が考えられる場合に、話者がどちらを選択するのかが研究課題となることがある。たとえば、次の文では give という動詞が異なる構文に現れている。

- (7) a. Susan gave the children toys.
b. Susan gave toys to the children. (Bresnan et al. 2007: 4)

(7a) の例では、give は二重目的語構文と呼ばれる構文に生起しており、動詞に the children と toys という名詞句が後続している。一方、(7b) の例では、give は与格構文と呼ばれる構文に生起しており、動詞に toys という名詞句と to the children という前置詞句が後続している。これらの構文に意味の違いがあるということは以前より多くの研究者によって主張されてきた (e.g. Green 1974)。しかし、(7a) と (7b) のどちらの文も文法的な誤りがあるわけではない。なんらかの理由によって、話者はどちらかの構文を選択しているのである。

どのような理由から、話者がどちらの語や構文を選択しているのかを明らかにするために、特に社会言語学における研究では二項ロジスティック回帰 (binomial logistic regression) と呼ばれる統計的手法が利用されることが多かった (e.g. Szendrői 2006)。しかし、二項ロジスティック回帰分析では、先行研究などに基づいて、分析に先立って説明変数を決め、それらの説明変数からどちらの語や構文が選択されるのかを決めるのに最適なモデルを作ることが求められる。

本研究でも、話者がどのような理由から、直示的移動動詞と移動様態動詞のうちどちらの

タイプの動詞を選択するのかを明らかにすることを目的としている。しかし、移動表現の研究では、これまで動詞の意味そのものが注目され研究が進められてきた。そのため、なぜ一方の動詞が選択されるのかということ自体を問われることが少なく、適切な説明変数を設定すること自体が難しい。実際には、上述の移動物 (FIGURE)、移動 (MOTION)、経路 (PATH)、基準物 (GROUND) を説明変数として想定することもできるが、それ以外の要因がかかわっていないのかどうかを検討する必要もあるだろう。このような理由から、本研究では、あらかじめ説明変数を決める必要のある二項ロジスティック回帰分析ではなく、これらの動詞の選択にかかわる妥当な説明変数を探索するために、大規模言語モデルを活用することにした。

2.3 大規模言語モデルの仕組みと活用方法

本研究では、すでに述べたように、大規模言語モデルを利用した探索的な手法によって、英語の移動動詞にかかわる話者の選択の仕組みを模索することを目的としている。本節では、大規模言語モデルがどういうものなのか、そして、本研究で採用した BERT と呼ばれる大規模言語モデルと、マスク付き言語タスク (Masked-language Task) と呼ばれる手法について簡単に説明しておきたい。

大規模言語モデルは、近年になって急速に発展を遂げているもので、大量のデータを使って深層学習により訓練した上で、多くのパラメータをもつ言語モデルである。本研究で利用する BERT (Bidirectional Encoder Representations from Transformers) の場合、BookCorpus と呼ばれるコーパス、および英語版の Wikipedia を組み合わせた 33 億語を学習データとして使い、3 億 4,000 万のパラメータを持つと言われている (Devlin et al. 2018)。前後の文脈を頼りにして、任意の語を予測するマスク付き言語タスクと呼ばれる方法などによって訓練されている。なお、このモデルを使って、任意の文の一部を予測させるというマスク付き言語タスクを利用することもできる。実際に、このタスクを試した例を以下で示したい。

- (8) a. So when Selig came to town, it wasn't for ice fishing or wild-rice soup.
(COCA, 2001)
- b. Then again, the Peter I remember would've thrown me down on this desk ten seconds after I walked in.
(COCA, 2008)
- c. Now Lenore woke up out of her dumb forever-sleep and ran into the room with her nightie on and her hair looking like she just came out of some tornado.
(COCA, 2000)

ここでは、BERT が文脈を頼りに正しく下線部の came、walked および ran を予測できるかどうかを確認するために、上の (8) の例から次のような文を作った。

- (9) a. So when Selig [MASK] to town, it wasn't for ice fishing or wild-rice soup.
- b. Then again, the Peter I remember would've thrown me down on this desk ten seconds after I [MASK] in.
- c. Now Lenore woke up out of her dumb forever-sleep and [MASK] into the room with her nightie on and her hair looking like she just came out of some tornado.

これらの例に対して、マスク付き言語タスクをおこなった結果は次に示すとおりである。

- (10) a. So when Selig came to town, it wasn't for ice fishing or wild-rice soup.
- b. Then again, the Peter I remember would've thrown me down on this desk ten seconds after I walked in.
- c. Now Lenore woke up out of her dumb forever-sleep and came into the room with her nightie on and her hair looking like she just came out of some tornado.

(10a) の例にある *came* と、(10b) の例にある *walked* は正確に予測できているが、(10c) の *ran* については誤った予測をしている。

本研究では、このような BERT による予測だけでなく、後述する注意機構と呼ばれる仕組みも確認することで、より詳細な分析を試みる。Clark et al. (2019) は、冠詞と名詞句、あるいは動詞と直接目的語といった依存関係が、かなりの程度で注意機構による注意の向けられ方と一致することを示している。マスク付き言語タスクの精度を確認することで、どの程度まで大規模言語モデルが正確に動詞の意味の違いを予測できるのかを確認し、その上で、英語の移動表現における直示的移動動詞と様態移動動詞の違いにどのような要因が関与しているのか把握したい。

認知言語学においては、萌芽的ではあるものの既に BERT を活用した研究は進められており、構文文法や語の多義性といったトピックで成果が報告されている (Madabushi et al. 2020, Fonteyn 2021)。

3. 実験手法と結果

ここでは、前節までに説明したマスク付き言語タスクを利用することで、任意の移動動詞の前後に生起する要素の情報のみから、どの程度の精度で経路動詞と直示的移動動詞の違いを判別できるのかを知るための実験をおこなう。

本研究では、統語的には同じ振る舞いを示す直示的移動動詞と移動様態動詞について、BERT という大規模言語モデルがどの程度まで正確な予測ができるかどうかを検証することを目的としている。なお、今回の実験では、事前学習済みの *bert-base-uncased* というモデルを使用し、実験は次のような手順でおこなった。

データの抽出のために、2020 年の 10 月にダウンロードした Davis (2008-) による COCA (The Corpus of Contemporary American English) のテキスト版を利用した。このコーパスから、

Levin (1993: 265–266) が掲載しているすべての移動様態動詞、および直示的移動動詞である *come* と *go* の活用形が含まれる文を Python を利用した正規表現によって無作為に 2,000 件抽出した。

次に、無作為に抽出された 2,000 件の用例から、今回の実験に不適切な次のような用例をすべて取り除いた。i) 物理的移動をあらわさないと考えられるもの、ii) 著作権の都合で @ @ @ のようにマスクされている部分を含むもの、iii) そして、前後の文脈がなければ、物理的な移動をあらわしているかどうか判断できないものである。なお、本研究では、厳密には物理的移動をあらわさないものの、探索的研究ということから、疑問文や否定文のものもデータに含めることにした。

不要な用例を取り除き、残った物理的移動をあらわす用例に対して、BertForMaskedLM というモデルを利用して、マスク付き言語タスクをおこなった。なお、本研究では、このタスクによる正解と不正解の区別は、直示的移動動詞か移動様態動詞かによって判断された。つまり、元の用例に含まれる動詞が *went* であれば、マスク付き言語タスクによる予測結果が *went* となる場合だけでなく *came* であっても正解とした。一方、元の用例に含まれる動詞が *walked* で、マスク付き言語タスクによる予測結果が *climbed* であっても移動様態動詞のため正解となるが、ここで予測結果が *went* であれば直示的移動動詞であるため不正解とした。

用例によっては、次の (11b) の結果に見られるように、マスク付き言語タスクによる予測結果が物理的な動詞ですらないということもあった。

- (11) a. I would have come sooner if I'd known you were sick. (COCA, 2002)
b. I would have died sooner if I'd known you were sick.

このような予測結果を示した用例は、今回の実験ではすべて取り除いた。また、直示的移動動詞と移動様態動詞の分布は均一ではなく、直示的移動動詞が生起しているものが 60%程度に対して、移動様態動詞が生起しているものが 40%程度となっており偏りがみられた。

以上のような手順にしたがって収集した 100 例を本研究では分析の対象とした。この 100 例に対するマスク付き言語テストの正答率は以下のとおりである。

表 1: マスク付き言語タスクの結果

	正解	不正解	合計
直示的移動動詞だと予測	55	10	65
移動様態動詞だと予測	27	8	35
合計	82	18	100

また、正解率を 50%とした場合と本結果を比較して、二項検定を実施したところ有意差も確認された ($p < .05$)。

4. 実験結果についての考察

コーパスから無作為に抽出された物理的移動が生起する用例において、大規模言語モデルを利用したマスク付き言語タスクによる動詞の予測は、精度の高いものであることが確認できた。しかし、先に述べたとおり、BERT が直示的移動動詞か移動様態動詞かを判別する上で、どのような要因が関与しているのかを探索的に調査することを本研究では目的としている。本節では、前節でみた結果の概要を踏まえて、本実験で利用した用例の文の長さ、および注意機構に注目した分析をおこなう。

4.1 文の長さと正解率

本実験で使用した COCA から無作為に抽出した用例には、10 語未満の短いものもあれば 30 語を超える長いものまである。直示的移動動詞である `come` と `go` は、コーパス調査でも会話や小説におけるセリフなどで使われることが多く、最も一般的な移動様態動詞と比べても大きな違いが見られる (Biber et al. 1999: 367)。書きことばと比べると、会話やセリフでは、文の長さも短くなることが予測される。そのため、本実験で利用した 100 の用例についても、各文に生起する語数を文の長さと定義し、それぞれの動詞が生起する文の長さの平均値を計算した。その結果、文の長さで生起する動詞の種類には次のような関連が見られた。

表 2: 文の長さと生起する動詞の種類の関連性

	文の長さの平均値
直示的移動動詞が生起	12.21
移動様態動詞が生起	16.76

このような事実を踏まえ、マスク付き言語タスクにおいて、より短い文では直示的移動動詞だと予測し、より長い文では移動様態動詞だと予測する傾向にあるかどうかを調査した。BERT によって、直示的移動動詞だと予測された文の長さの平均値、および移動様態動詞だと予測された文の長さの平均値を先ほどと同じように計算した。結果は次の表 3 および図 1 に示す通りである。

表 3: マスク付き言語タスクの予測結果

	文の長さの平均値
直示的移動動詞だと予測	12.56
移動様態動詞が生起	16.37

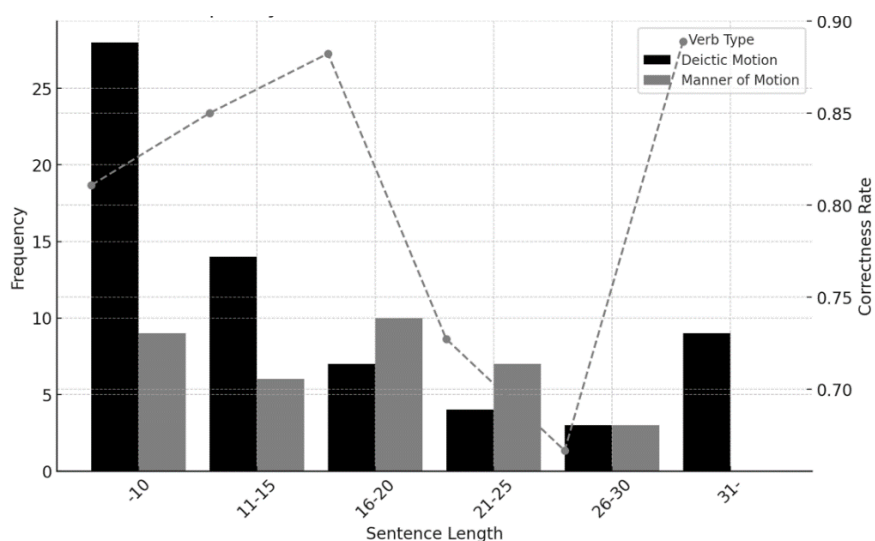


図 1: マスク付き言語タスクにおける直示的移動動詞と移動様態動詞の予測

図の折れ線グラフは、それぞれの文の長さにおけるマスク付き言語タスクの正解率を示している。平均値だけを示した表 3 の結果からはわかりにくい、短い文においては、マスク付き言語タスクによって直示的移動動詞だと予測されやすいことが図 1 からわかる。本研究では、分析の対象としたデータの量が少ないということも考えられるが、文の長さとは正解率の間には相関が見られそうにない。しかし、短い文でも長い文でも極端に正解率が低下しているわけでもない。これは、前述のとおり BERT が長い文であっても適切に処理できる仕組みをもっているからだと考えられる。

4.2 注意機構の分析

前節までで、大規模言語モデルによる予測の精度がたしかなものであることを確認したが、本節では注意機構にも目を向けることで、より詳細な仕組みに迫りたい。先行研究としてあげた Clark et al. (2019) による研究でも、注意機構の詳細な調査によって統語的な情報の分析が可能であることが示されていた。本研究では、どの程度まで意味的な情報の分析に注意機構が有効なのかを模索したい。大規模言語モデルが動詞の選択をする際に、周囲の要素のうち、どの要素に注意を向けているのかが明らかになれば、今後の言語研究にも新たな視点を加えることができるのではないだろうか。ただし、注意機構は仕組みが複雑であるため、本節では、注意機構を可視化しつつ、本研究で明らかになった点への言及にとどめたい。

まず、直示的移動動詞と移動様態動詞とで特徴的な違いが見られた例として、否定文にこれらの動詞が生じた場合について確認してみたい。なお、これ以降に示す例はすべて COCA から抽出した用例ではなく、本研究の目的にあわせて COCA の用例を修正して作例したものである。

- (12) a. He didn't go to the restaurant.
 b. He didn't walk to the restaurant.

(12a) および (12b) の文を BERT が処理する場合、注意機構では動詞がどの要素に注意を向けているのかを bertviz という Python のパッケージによって可視化したものが以下の図 2 および図 3 である。

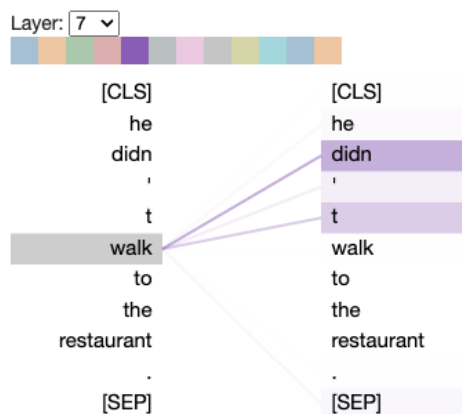
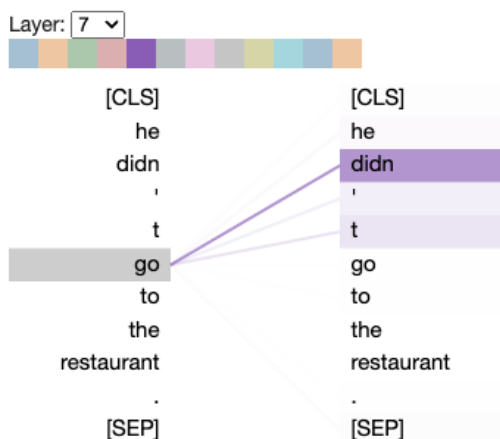


図 2: 否定文に go が生起する文の注意機構 図 3: 否定文に walk が生起する文の注意機構

それぞれの図において、グレーで示された動詞がより強く注意を向けている要素が、より濃い色で示されている。walk が生起している例の方と比べて、go が生起している例の方がより didn't という否定をあらわす助動詞への注意が向けられている。これらの図では、それぞれの語以外に [CLS] と [SEP] という項目が見られるが、[CLS] は各文の先頭に置かれる特殊トークンと呼ばれるもので、本研究の分析では使うことがないため、詳細な説明は避ける。また、[SEP] も [CLS] と同じく特殊トークンと呼ばれるものだが、こちらも本研究の分析ではそれほど重要な意味を持たないため後ほど簡単に触れておくにとどめたい。

他にも、直示的移動動詞と移動様態動詞とで違いが確認されたのは、直示性をあらわす there や here が移動の経路として現れた場合である。

(13) a. He went there.

b. He skipped there.

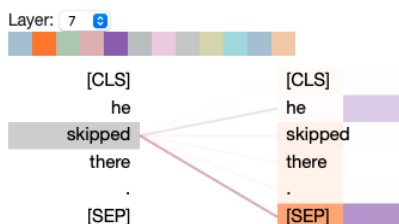
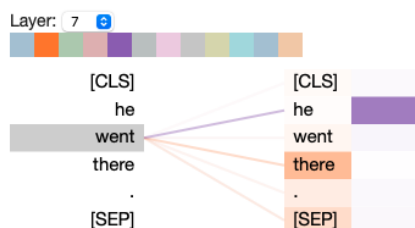


図 4: go と there が共起する文の注意機構 図 5: skip と there が共起する文の注意機構

移動様態動詞である **skip** が生起している図 5 では、[SEP] と呼ばれる文と文の境界部分をあらわす特殊トークンに注意が向けられており、周囲の語はそれほど参照されていない。

最後に、直示的移動動詞と移動様態動詞とで違いが見られなかったもののうち、興味深いものを確認しておきたい。

- (14) a. We went across the street to the field.
 b. We walked across the street to the field.

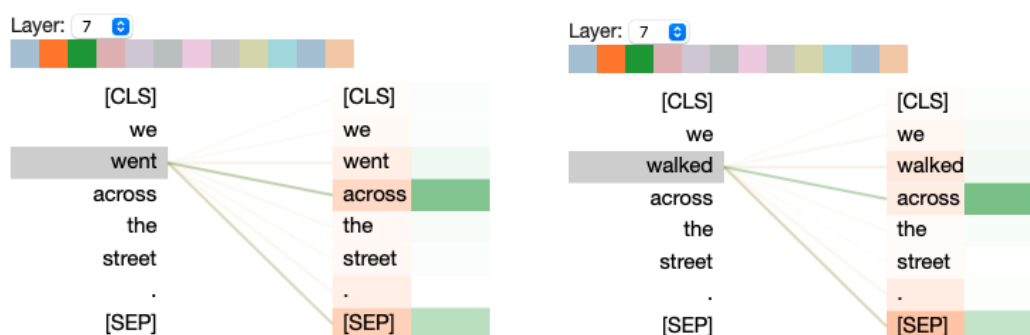


図 6: go と across が共起する文の注意機構 図 7: walk と across が共起する文の注意機構

一般的に、移動様態動詞は通過点をあらわす across のような経路表現と共起しやすいということが知られている (e.g. Stefanowtisch and Rohde 2004)。そのため、(14a) と (14b) のような例では、動詞から across への注意の向けられ方には違いが見られると予測していた。しかし、実際にはそのような差は確認できなかった。

以上のことから、直示的移動動詞と移動様態動詞の選択には、否定文かどうか、there や here といった直示的な経路表現があるかどうか、重要であるという可能性が示された。否定文では、移動したかどうか重要なのであって、移動の様態を指定する必要がないことから、移動の様態を語彙化していない come や go が生起しやすいというのは、妥当な結果だと考えられる。一方で、経路をあらわす前置詞や不変化詞がそれほど動詞と強い関係にないというのは想定外の結果であった。しかし、このような結果こそ、探索的な研究の成果であり、今後、さらに詳細な分析が求められる部分であろう。また、主語である名詞句の性質についても、どのような名詞句が重要なのか、そして、there や here との関係についても、さらに分析を進めていくべきであろう。

5. まとめと今後の課題

本研究は、大規模言語モデルがどの程度まで、語の意味の研究に資する可能性をもっているのかを検証しながら、移動動詞の意味の違いにかかわる要素を探索的に調査することを目的としていた。

今後は、今回の探索的な調査から得られた結果を含め、従来の先行研究でも議論されてきた点なども説明変数として考慮に入れ、動詞の選択にとってより重要な変数を見定めていく必要がある。移動表現の類型論での議論において重要な要素だと考えられてきた移動物 (FIGURE)、移動 (MOTION)、経路 (PATH)、基準物 (GROUND) という変数と、今回の調査によって明らかになった *there* と *here* の有無や、否定文であるかどうかなどの変数を加えて、二項ロジスティック回帰分析などを利用することが考えられる。

さらに、BERT の *attention* に目を向けることで、言語研究の新たな可能性を示すことができたのではないかと考えている。これまでの言語研究では、仮説を検証する方法として、i) 内省、ii) コーパスの利用、iii) 心理実験という方法が広く採用されてきた。しかし、本研究では、コーパスから採取した用例の一部を変更することで、どのように *attention* が変化するのかを示した。

謝辞

本研究は科学研究費プロジェクト「英語の移動表現における経路概念の機能的分析」(20K13069) による研究成果の一部である。

引用文献

- Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan E. (1999). *The Longman Grammar of Spoken and Written English*: Person Education Limited.
- Bresnan, J., Cueni, A., Nikitina, T. & Baayen, R. H. (2007). Predicting the dative alternation. In G. Boume, I. Kraemer & J. Zwarts (Eds.), *Cognitive foundations of interpretation*. (pp.69–94). Royal Netherlands Academy of Science.
- Clark, K., Khandelwai, U., Levy, O. & Manning, C. D. (2019). What does BERT look at? An analysis of BERT's attention. <https://arxiv.org/abs/1906.04341>
- Davis, M. (2008–). The Corpus of Contemporary American English (COCA). <https://www.english-corpora.org/coca/>
- Devlin, J., Chang, M-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. <https://arxiv.org/abs/1810.04805>
- Firth, J. R. (1957). A synopsis of linguistic theory 1930–1955. In *Studies in Linguistic Analysis*. (pp.1–32). University of Birmingham Press.
- Fillmore, C. J. (1982). Frame semantics. In The Linguistic Society of Korea (Ed.), *Linguistics in the morning calm: Selected papers from SICOL-1981*. Hanshin Publishing Company.
- Fonteyn, L. (2021). Varying abstractions: a conceptual vs. distributional view on prepositional polysemy. *Glossa: a journal of general linguistics*, 6(1), 1–28.
- Geeraerts, D. (2010). *Theories of Lexical Semantics*. Oxford University Press.

- Green, G. (1974). *Semantics and Syntactic Regularity*. Indiana University Press.
- Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In K. Denning & S. Kemmer (Eds.), *Universals of language*. (pp.73–113). MIT Press.
- Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair*. (pp.157–176). John Benjamins.
- Levin, B. (1993). *English Verb Classes and Alternations: A preliminary investigation*. The University of Chicago Press.
- Madabushi, H. T., Romain, L., Divjak, D. & Milin P. (2020). CxGBERT: BERT meets Construction Grammar. <https://arxiv.org/abs/2011.04134>
- Stefanowitsch, A. & Rohde, A. (2004). The goal bias in the encoding of motion events. In G. Radden & K-U. Panther (Eds.), *Studies in linguistic motivation* (pp.249–267). Mouton de Gruyter.
- Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Languages*, 2, 1–33.
- Szmrecsanyi, B. (2006). *Morphosyntactic Persistence in Spoken English: A corpus study at the intersection of variationist sociolinguistics, psycholinguistics, and discourse analysis*. Mouton de Gruyter.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description, vol.2* (pp.36–149). Cambridge University Press.
- Talmy, L. (1991). Path to realization. *Proceedings of the seventeenth annual meeting of the Berkeley Linguistics Society*, 480–519.
- Talmy, L. (2000). *Toward a cognitive semantics, vol. 2: Typology and process in concept structuring*. MIT Press.

