

令和 6 年 6 月 23 日現在

機関番号：24501

研究種目：基盤研究(C)（一般）

研究期間：2019～2023

課題番号：19K12733

研究課題名（和文）Linking Vision and Language through Computational Modelling

研究課題名（英文）Linking Vision and Language through Computational Modelling

研究代表者

CHANG Franklin (Chang, Franklin)

神戸市外国語大学・英米学科・教授

研究者番号：60827343

交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：本研究では、視覚と言語表現の関係を探るため、過去形（「走った」）と進行形（「走っている」）を選択する際に使用される視覚的な手がかりに焦点を当てた。ビデオ映像に身体動作の「終了状態」があるものとないものを用意して実験を行った結果、日本人の成人と3-5歳の子供の両方が、「終了状態」を説明する際に過去形をより多く使用していることがわかった。さらに、映像内の身体の動きをもとに日本語の動詞を生成するディープラーニングモデルを開発した。このモデルはビデオ実験で被験者が生成した動詞を正しく予測できており、視覚と言語の密接な関連を示している。

研究成果の学術的意義や社会的意義

本研究では、大人と子どもが動画の視覚情報をどのように利用して動詞や動詞の形態を生成するかを調べた。日本人が視覚的な情報からどのように動詞を生成するかを示す、計算AIモデルを開発した。このモデルは、第一言語と第二言語の習得をサポートするための視覚資料作成に役立つ。また、人間が視覚的情報を言語化する方法を解明する一助となる本研究は、日本語を話すAIシステムを作成する際に役立つ。

研究成果の概要（英文）：Language is used to describe events that we see, but the relationship between visual and language representations is still not well understood. In this research, we focused on the visual cues that are used to select past tense (ran) and progressive aspect forms (is running). We created videos of actions by human-like characters where they performed actions like running. Then we added objects into the scene that signaled that endstate had been reached. We found that both Japanese adults and 3-5 year old children used past tense more when the videos has endstate information compared to when it didn't. To understand how they mapped these visual signals into language, we developed a deep learning model that tracked the motion of body parts and objects in the videos and used that to generate Japanese verbs. The model could explain our data and it made predictions that were confirmed in a follow-up experiment. This work demonstrates that vision and language are tightly linked.

研究分野：言語心理学

キーワード：視覚情報 ディープラーニングモデル 動詞 過去形 進行形 終了状態 子ども 大人

## 様式 C - 19、F - 19 - 1 (共通)

### 1. 研究開始当初の背景

私たちが言葉を使うとき、目に見えるものを説明することが多い(例えば、「少年が岩を飛び越えている」、「犬が猫を追いかけている」など)。そのためには、その場面で何が起きているのかを理解する必要がある(例えば、少年は何かをしているが、岩は何もしていない、など)。この理解をするためには、私たちは、場面内のオブジェクトを追跡しなければならない。(例えば、「飛び越える」という動作を見るには、少年の身体を追跡する必要がある。)さらに、少年の身体の動きが「飛び越える」という動作に対応することを学習しなければならない。このように、視覚は私たちが言語の意味を導き出し学習する際の主要な場のひとつであるにもかかわらず、まだ解明されていない側面がたくさんある。

### 2. 研究の目的

この研究では、過去形と進行形を理解するための視覚的手がかりに注目した。例えば、男性が数秒間走っている動画を見た場合、人々は「男性が走っている」というように進行形で表現することが多い。しかし、ゴールロープがあり、映像の最後に男性がロープを断ち切った場合、人々はそれを「男性がレースを走った」と過去形で表現することが多い。映像の最後に訪れる視覚的な変化は、その出来事が終わったという情報を提供するので、過去形が適切なのである。私たちは、このような視覚的な情報を使用して、子どもや大人が日本語の動画描写において過去形と進行形をどのように選択しているのかを調べた。

### 3. 研究の方法

いくつかの動作(「走る」、「歩く」、「跳ぶ」、「投げる」、「刻む」など)を含む動画を作成した。それぞれの動画について、2つのバージョンを作成した。1つは、動作が終了したことを知らせるオブジェクトの終了状態情報(例えば、ゴールのロープ)を含むもので、もう1つは、そのオブジェクトの終了状態情報を含まないものである。日本語母語話者の大人と幼稚園児にこれらの動画を試聴してもらい、そこで起こっている動作を説明してもらうことにより、彼ら・彼女らが過去形と進行形のどちらのアスペクトを生成するかを調べた。

### 4. 研究成果

上記の実験の結果、子供も大人も場面内のオブジェクトに敏感に反応し、終止状態情報のない動画に比べて、終止状態情報のある動画の方が過去形をより多く使用することがわかった。

被験者によるこのようなアスペクト生成を説明するために、私たちは日本語の動詞生成のディープラーニングモデルを開発した。このモデルは、大人と子供の両方のデータを反映することが可能で、動画の中でオブジェクトがどのように動いたか(例えば、男性がゴールロープに近づき、最終的にこれを切る、など)についての情報を使って動詞のアスペクトを選択する。本モデルが生成した動詞のアスペクト生成を、実証実験で確認したところ、モデルは実験内容を正しく予測することができていた。この成果は *Cognitive Science* 誌に掲載された。

- Chang, F., Tatsumi, T., Hiranuma, Y., & Bannard, C. (2023) Visual heuristics for verb production: Testing a deep-learning model with experiments in Japanese. *Cognitive Science*. 47(8), e13324.

さらに、上記のモデルは、私が以前に開発した、事象を理解し言語と結びつける上でオブジェクト追跡がいかに重要であるかというアイデアを利用したものである。たとえば、ボール A が別のボール B を押すことを理解するには、ボール A がボール B に向かって移動する動きを追跡する必要があります。また、ボール A とボール B が接触する様子を見ることも重要です。最終的に、ボール B は押された結果、ボール A から離れるはずですが。私は大学院生の教育・指導の一環として、彼らと共にこのテーマでいくつかの研究を行っていたが、科研費を利用して以下の2つの論文を完成させ、学術誌に発表することができた。

- Jessop, A. & Chang, F. (2022) Thematic role tracking difficulties across multiple visual events influences role use in language production. *Visual*

Cognition. 30(3), 151-173.

- Jessop, A. and Chang, F. (2020). Thematic role information is maintained in the visual object-tracking system. *Quarterly Journal of Experimental Psychology*, 73(1), 146-163

## 5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件/うち国際共著 4件/うちオープンアクセス 4件）

1. 著者名 Chang Franklin, Tatsumi Tomoko, Hiranuma Yuna, Bannard Colin	4. 巻 47
2. 論文標題 Visual Heuristics for Verb Production: Testing a Deep Learning Model With Experiments in Japanese	5. 発行年 2023年
3. 雑誌名 Cognitive Science	6. 最初と最後の頁 1-38
掲載論文のDOI (デジタルオブジェクト識別子) 10.1111/cogs.13324	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 該当する
1. 著者名 Jessop Andrew, Chang Franklin	4. 巻 30
2. 論文標題 Thematic role tracking difficulties across multiple visual events influences role use in language production	5. 発行年 2022年
3. 雑誌名 Visual Cognition	6. 最初と最後の頁 151 ~ 173
掲載論文のDOI (デジタルオブジェクト識別子) 10.1080/13506285.2021.2013374	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 該当する
1. 著者名 Chang Franklin, Tsumura Saki, Minemi Itsuki, Hirose Yuki	4. 巻 43
2. 論文標題 Abstract structures and meaning in Japanese dative structural priming	5. 発行年 2022年
3. 雑誌名 Applied Psycholinguistics	6. 最初と最後の頁 411 ~ 433
掲載論文のDOI (デジタルオブジェクト識別子) 10.1017/S0142716421000576	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 該当する
1. 著者名 Jessop Andrew, Chang Franklin	4. 巻 73
2. 論文標題 Thematic role information is maintained in the visual object-tracking system	5. 発行年 2019年
3. 雑誌名 Quarterly Journal of Experimental Psychology	6. 最初と最後の頁 146 ~ 163
掲載論文のDOI (デジタルオブジェクト識別子) 10.1177/1747021819882842	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 該当する

〔学会発表〕 計2件（うち招待講演 1件 / うち国際学会 0件）

1. 発表者名 Chang, F., Tatsumi, T., Hiranuma, Y., & Bannard, C.
2. 発表標題 The role of endpoints in verb tense/aspect morphology in Japanese children and adults.
3. 学会等名 Plenary talk presented at the TL/Mental Architecture for Processing and Learning of Language conference. Tokyo, Japan. (招待講演)
4. 発表年 2020年

1. 発表者名 Chang, F., Tatsumi, T., Hiranuma, Y., & Bannard, C.
2. 発表標題 A deep learning verb production model using input from visual animations
3. 学会等名 Talk presented at the TL/Mental Architecture for Processing and Learning of Language conference.
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------