

統語構造の反映および付加情報伝達におけるジェスチャーの効果

The effect of gestures that reflect the linguistic structure and transmit additional information

木村 陽菜[†], 安田 哲也[‡], 小林 春美[§]
Hina Kimura, Tetsuya Yasuda, Harumi Kobayashi

[†]東京電機大学大学院, [‡]東京大学, [§]東京電機大学
Graduate School of Tokyo Denki University,
The University of Tokyo, Tokyo Denki University
23rmd14@ms.dendai.ac.jp, t-yasuda@g.ecc.u-tokyo.ac.jp, h-koba@mail.dendai.ac.jp

概要

本研究では、曖昧な句構造において、付加情報がある場合、ジェスチャーが自発的に産出されたのか、また、どのようなジェスチャーが表出されたか検討した。統語境界に基づいて、新たに分析を行なった結果、ジェスチャーは統語構造に応じて表出され、意味情報を表現していることが示唆された。よってジェスチャーは、付加情報の補完に役立っていると考えられ、言語・ジェスチャー両方が提示された場合、必要な情報を過不足なく伝えると考えられる。

キーワード: Co-speech gesture, Sub-merge 構造, Pot-merge 構造, trade-off 仮説, hand-in-hand 仮説

1. 目的

言語は階層構造を持っていることから複数の解釈が可能であり、意味解釈において曖昧性を含む(Fujita & Fujita, 2021)。言語には発話そのものである表層構造と、意味を表す深層構造があるとされる。例えば、“green tea cup”の場合、head が移動する Sub-merge 構造では、“tea-cup for green-tea”の意味を持つ句構造として解釈される。まず green が head となり green と tea のまとまりに対して、tea が head となり cup とまとまりを作る。一方、head の移動がない Pot-merge 構造では、“green-colored tea-cup”の意味を持つ句構造として解釈され、tea と cup のまとまりに対して、head である green がまとまりを作る。このように言語は階層構造を持っているため、発話そのものの表層構造からの判断では解釈に曖昧性が生じる。この曖昧性の解消の役割を果たす可能性がある手がかりとして、音韻情報の付与(Kubozono, 1998; Hirose, 2020 等)やジェスチャー(Kashiwadata et al., 2020)が挙げられる。

Kashiwadata et al.では、ジェスチャーによる非曖昧化の役割について、句構造解釈の曖昧性に着目し、4 語の句、例えば句「黒いしっぽの大きなねこ」を用いて、調べた。その結果、統語構造に応じて言語とジェスチャーが同期することが示唆される結果を得たことから、

これはジェスチャーによる統語的非曖昧化を行った証拠であると議論した。また、この研究をレビューした岡久(2023)は、統語境界を示すためにジェスチャーが有効に使われていた可能性があるとして指摘した。さらに岡久(2024)は、この統語境界において、形容詞については Kashiwadata et al.の知見を追認する結果を得た。こうした考察から、ジェスチャーによる非曖昧化の現象は、Co-speech gesture において起こると考えられる。

Co-speech gesture は発話とともに産出される。その特徴として、発話のある部分を強調(Bull & Connelly, 1985)することや、発話内容を補完すること(Iverson & Goldin-Meadow, 2005; Kita & Özyürek, 2003)、言語の理解を助ける(Kelly, Özyürek, & Maris, 2010)といったものがあるとされる。この Co-speech gesture 現象を説明するの的に確かな仮説として trade-off 仮説 (Bangerter, 2004; de Ruiter, Bangerter, & Dings, 2012), と hand-in-hand 仮説(So, Kita, & Goldin-Meadow, 2009) がある。trade-off 仮説は、話し手は発話情報とそれに伴うジェスチャー情報を過不足なく利用するというものである。つまり、発話で明示されていない対象をジェスチャーにより補完することで発話を適切に解釈させるということである。Bangerter では、ジェスチャーがなされる時に対象指示に関する発話が減ったことから、ジェスチャーが対象の指示すべき情報を補完したことを提示した。一方、So et al.では、対象とすべき事物を発話とジェスチャーの両方で表現していたことを報告した。この仮説は、ジェスチャーが発話の効果を高める可能性を示唆する。

Kimura, Yasuda, and Kobayashi (2023) では、co-speech gesture の観点から、参加者が互いに伝え合うコミュニケーションタスクにより、情報性の違いにおいて、枝分かれ構造の解釈にジェスチャーのタイミングがどのように寄与するのか検討した。情報性の違いとは、実験時に提示されるイラストの背景情報の違いであり、発話には含まれないが聞き手が解釈するために必要と

なる背景情報のことであった。実験の結果、pot-merge 構造(RB)において情報性が高い場合は、最初の名詞の発話の開始時刻に比べて、ジェスチャーのストロークの開始時刻が早くなることが示唆された。また、ジェスチャー産出を視覚的に分析した場合、動詞の意味と句構造の違いを反映しながら、発話と共にジェスチャーも産出されていたことから、発話では表現されない情報をジェスチャーにより補完していた可能性を示す。

Kimura, Yasuda, and Kobayashi (2024) では、事物の動作性に着目し、枝分かれ構造解釈へのジェスチャーの寄与について検討した。動詞「一ている」の表現には、その動作が始まった状態、途中の状態、終わりかけの状態が表現され得ることから、これらの情報がジェスチャーにより補完されるのか調査した。動詞を含む3語の曖昧な句が提示され、情報性の違いとして事物の動作状態が異なるイラストが用いられた。実験の結果、動作情報の違いによってジェスチャーの産出されるタイミングは異なっており、ジェスチャーによりこれらの情報が付加されることが示唆された。

本研究では、曖昧な句構造において、発話されないが正しく解釈するには伝える必要のある付加情報がある場合、ジェスチャーが自発的に産出されるのか、また、どのようなジェスチャーが表出するのか検討した。付加情報として、事物の情報量の違い (Kimura et al., 2023) と、事物の動作性の違い (Kimura, Yasuda, & Kobayashi, 2024) を含み、参加者は情報を過不足なく伝えるのか検証した。以上を検討するために、事物の情報量を取り扱ったデータを Object concept, 事物の動作性を取り扱ったデータを Motion concept とした分析を行った。

2. 方法

2.1 ジェスチャーコーパス

ジェスチャーコーパスは、2人のペアで参加していたビデオデータを Kendon (2004)の定義に基づき、ELAN (ver6.4 と ver. 6.7)を利用してコーディングしたものから構成された。

実験データ 1 のジェスチャーコーパス(以下、実験データ 1)は、31人の日本語を母語とする学生と1人の中国語を母語とする学生が参加し、回答のずれやタスクの遂行が不完全であったグループを除外した24人分のデータで構成されたものであった(M age = 21.6)。実験データ 2(以下、実験データ 2)のジェスチャーコーパス

は、28人の日本語を母語とする学生のデータで構成されたものであった(M age = 21.96)。

2.2 実験条件

実験データ 1 の実験条件は、枝分かれ構造(2: Sub-merge (LB) / Pot-merge (RB) 構造)、情報性として事物情報(2: 高い(More)/低い(Less)), 動詞(2: 落下している(Fall)/飛んでいる(Fly)) であった。

実験データ 2 の実験条件は、枝分かれ構造(2: Sub-merge / Pot-merge 構造)、情報性として動作情報(2: はじめ (Beginning)/おわり (Ending)), 動詞(2: 落ちている (Fall)/飛んでいる(Fly))であった。

2.3 実験刺激

実験で使用された刺激は、動詞 (V: verb) + 名詞 1 (N1: first noun) + 名詞 2 (N2: second noun) の3語から構成される曖昧な句構造であった。例えば、「飛んでいるカメの写真」や「落下している猫の写真」というような句であった。実験データ 1 では、動詞(V)が2種類、名詞 1 (N1)が6種類、名詞 2 (N2)が3種類から構成される12種類の句を用いた。実験データ 2 では、動詞(V)が2種類、名詞 1 (N1)が6種類、名詞 2 (N2)が3種類から構成される12種類の句を使用した。実験データ 1 と実験データ 2 はイラストの種類やその属性は似ているが、イラストに含まれた情報は異なる。

枝分かれ構造条件は、動詞+名詞-1+名詞-2 の句を使用し、Sub-merge 構造と Pot-merge 構造で解釈される3語の曖昧な句構造を使用した。例えば、「飛んでいるタヌキの写真」のような句の場合、Sub-merge 構造では、はじめに「飛んでいる」と「タヌキ」がまとまりとなり、「飛んでいるタヌキが写っている写真」と解釈される。一方、Pot-merge 構造では、「タヌキ」と「写真」のまとまりに対して「飛んでいる」がまとまりとなり、「タヌキが写っている写真が飛んでいる」という解釈がされる。

事物情報条件は、実験刺激で提示するイラスト背景の情報性が高い(More)場合と、低い(Less)場合とした。情報性が高い条件は、背景に描かれた自然物や建物(e.g., 崖やビル) など、発話には表現されないが、聞き手の理解を助ける情報を加えることが可能であると予測されるものであった。一方で、情報性が低い条件は、イラスト背景に、他に加えることが可能な情報はあるが情報性が高い条件に比べると加える情報は少ない(e.g., 平地や木) と予測されるものであった。

動作情報条件においては、実験刺激で提示するイラスト事物の動きが始まった状態を「はじめ(Beginning)」, 事物の動きが終わりに近い状態を「おわり(Ending)」とした。同じ「飛んでいる」という語は、飛び始めた状態にも、飛び終わりの状態にも、両方に使うことが可能である。たとえば、「はじめ」の状態の「飛んでいるカメの写真」という句の Sub-merge 構造を表現するイラストでは、カメが今まさに飛びはじめた様子が写された写真が描かれていた。

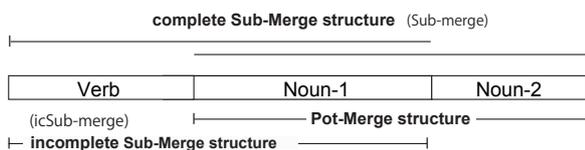
2.4 実験手続き (コーパスデータ)

各データセットに共通する実験手続きは下記の通りである。1) 参加者の役割は、話し手と聞き手の役割にランダムに振り分けられた。2) 話し手は、決められた文言の発話を求められ、イラストの内容を記憶して聞き手に伝えた。3) その後、聞き手は聞いた話に関して、四択のイラストの中から話し手が伝えたと思ったイラストを選んだ。なお、すべての試行が終了した後、役割と場所を交代して実験を行なった。決められていた文言を発すること以外については、表現の制限を行わなかった。加えて、ジェスチャーの利用について参加者から質問があった場合には、自由であることを伝えた。

3. 分析方法

曖昧な句構造におけるジェスチャーの産出の違いを調べるために、岡久(2023)で議論されていた統語境界の概念を利用し、ジェスチャーのタイプをジェスチャーの生起時間を利用して Sub-merge, Pot-merge, No-gesture に分類した (図1)。なお, Sub-merge は、チャンク移動を示さない不完全なものもあったため、不完全な incomplete Sub-merge (icSub-merge) と完全な complete Sub-merge (cSub-merge) の2種類に分けた。

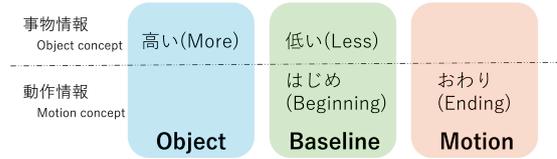
図1 ジェスチャータイプのコーディング



事物情報条件の「低い(Less)」場合と動作情報条件の「はじめ(Beginning)」の場合を Baseline, 事物情報条件における「高い(More)」の場合を Object, 動作情報条件における「おわり(Ending)」の場合を Motion として3つのデータセットに分けた(図2)。これらのデータセット

について、それぞれ χ^2 検定を用い、分析を行なった。

図2 データセットの分類



4. 結果・考察

事物情報(Object concept)条件「低い(Baseline; Less)」と動作情報(Motion concept)条件「はじめ(Baseline; Beginning)」の比較(図3a)において、ジェスチャータイプごとに χ^2 検定を行った結果、有意な偏りが認められた ($\chi^2(df=3, n=366) = 19.4, p < .01$)。調整済み残差を算出したところ、Pot-merge では、低い(Baseline; Less)の方がはじめ(Baseline; Beginning)よりも、多くジェスチャーを産出していた ($z > \pm 1.96$)。また、incomplete Sub-merge では、はじめ(Baseline; Beginning)の方が低い(Baseline; Less)よりも、多くジェスチャーを産出していた ($z > \pm 1.96$)。

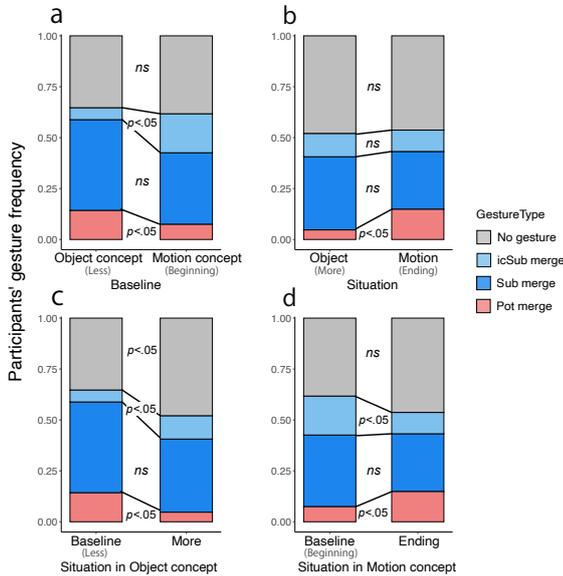
事物情報条件(Object concept)の「高い(Object; More)」と動作情報条件(Motion concept)の「おわり(Motion; Ending)」の比較(図3b)において、ジェスチャータイプごとに χ^2 検定を行った結果、有意な偏りが認められた ($\chi^2(df=3, n=354) = 11.3, p < .05$)。調整済み残差を算出したところ、Pot-merge の場合にのみ、「高い(Object; More)」の方が「おわり(Motion; Ending)」よりも、多くジェスチャーを産出していた ($z > \pm 1.96$)。

事物情報条件における「低い(Baseline; Less)」と「高い(Object; More)」の比較(図3c)において、ジェスチャータイプごとに χ^2 検定を行った結果、有意な偏りが認められた ($\chi^2(df=3, n=396) = 18.6, p < .01$)。調整済み残差を算出したところ、Pot-merge では、低い(Baseline; Less)の方が高い(Object; More)よりも、多くジェスチャーを産出していた ($z > \pm 1.96$)。また、incomplete Sub-merge では、高い(Object; More)の方が低い(Baseline; Less)よりも、多くジェスチャーを産出していた ($z > \pm 1.96$)。

動作情報条件における「はじめ(Baseline; Beginning)」と「おわり(Motion; Ending)」について(図3d)、ジェスチャータイプごとに χ^2 検定を行った結果、有意な偏りが認められた ($\chi^2(df=3, n=324) = 10.5, p < .05$)。調整済み残差を算出したところ、incomplete Sub-merge の場合、はじめ(Baseline; Beginning)の方が、おわり(Motion; Ending)よりも多くジェスチャーを産出していた ($z > \pm 1.96$)。

Pot-merge の場合, おわり(Motion; Ending)の方がはじめ(Baseline; Beginning)よりも多くジェスチャーを産出していた($z > \pm 1.96$).

図3 話し手のジェスチャーの産出



どの状況においても, 統語のチャンクを示すような Sub-merge タイプのジェスチャーの使用は変わらないことがわかった. このタイプのジェスチャーは, まず動詞と名詞 1 のチャンクを描き, 次に名詞 1 と名詞 2 のチャンクを描くというものであった. よって, 岡久(2023)の指摘した, ジェスチャーが統語境界を示す可能性を示唆したと考えることができる.

各 Baseline を比べた図 3a において, 各 concept の incomplete Sub-merge タイプのジェスチャーを比較すると, 動作情報のデータセットの方がジェスチャーの産出が多かった. これは, ジェスチャーする対象が落ち始めたために, 参加者は, 最初の塊である動詞と名詞 1 のチャンクを示す incomplete Sub-merge タイプのジェスチャー, 例えば手首を下方に曲げ「事物の落ち始めの動き」を示すなどを行っていたことが考えられる. 対して, 事物情報と動作情報の違いを比較した図 3b では, Pot-merge タイプのジェスチャーが多かった. Pot-merge タイプのジェスチャーは名詞 1 と名詞 2 のチャンクを示すことができる. よって, 図 3a の落ち始めている状態を表す刺激(Beginning)と比較し, 図 3b の動作の終わりを示している刺激(Ending)の方が, Pot-merge タイプのジェスチャーで表現することは容易であり, このタイプのジェスチャーが産出されたと考えられる. このことは図 3d の結果においても同様のことが示唆さ

れる.

事物情報のデータセット(図 3c)に着目してみると, 事物が多く描かれている「高い(More)」の場合に incomplete Sub-merge タイプのジェスチャーが Baseline (Beginning)と比較し多かった. これは, 背景にある事物を描いていたためにその区間でジェスチャーが行われていたと考えられる. Kimura et al. (2023) でのデータでは発話前にジェスチャーしている場合が観察されていることから, 図 3c において No gesture (発話同期ジェスチャー無し)が多く観察されたことは, 発話をしていない区間で発話とは別にジェスチャーをしていた可能性も示唆する.

以上のことから, 事物に関する情報は発話外でもジェスチャーで描くことが可能であるが, 動作に関する情報は発話と共に描かれやすいことが示唆される. Sub-merge タイプのジェスチャーは, 常に観察されていたことから, 構造を表すようなジェスチャーは基盤的に使用されていた可能性がある.

ジェスチャーは, 統語構造を反映しながら, それに加えて発話で表現されない情報を表現していたことから, 過不足なく情報を伝えるものであると考えられる.

謝辞

本研究は科研費 JP17H06382 (H.K.), JP20H01763 (H.K.), JP20K03375 (T.Y.)の助成を受けた.

主要参考文献

- Fujita, H., & Fujita, K. (2022). Human language evolution: a view from theoretical linguistics on how syntax and the lexicon first came into being. *Primates*, 63, 403-415.
- Kashiwadate, K., Yasuda, T., Fujita, K., Kita, S., & Kobayashi, H. (2020). Syntactic structure influences speech-gesture synchronization. *Letters on Evo Behav Sci*, 11, 10-14.
- Kimura, H., Yasuda, T., & Kobayashi, H. (2023). Spontaneous co-speech gestures with prompt phrases reflect linguistic structures. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 45, 1140-1145.
- Kimura, H., Yasuda, T., & Kobayashi, H. (July, 2024). Co-speech gestures complement motion state information expressed by verbs. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 46, 4582- 4587.
- 岡久太郎, & 白勢彩子. (2024). 統語的曖昧文を区別するイントネーションとジェスチャーに対する聞き手の理解. *認知科学*, 31(1), 138-156.