

# 枝分かれ構造は各チャンクの発話継続時間に反映し 聞き手の理解を向上させる

## Branching structure is reflected in the utterance duration of each chunk to improve listeners comprehension

木村 陽菜<sup>†</sup>, 安田 哲也<sup>‡</sup>, 小林 春美<sup>§</sup>  
Hina Kimura, Tetsuya Yasuda, Harumi Kobayashi

<sup>†</sup>東京電機大学大学院, <sup>‡</sup>東京大学, <sup>§</sup>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、コミュニケーションタスクにおいて、伝え手により自発的に産出された発話の継続時間が、曖昧な句構造の解釈にどのように寄与するのか検討した。発話を聞いた回答者が句構造について判断した回答の正誤を基準とした分析を行った。その結果、左枝分かれの句よりも右枝分かれの句の方が、発話継続時間が長いこと、またこの長さの違いが回答の正誤にも関係することがわかった。動詞と名詞のまとまりからなる最初のチャンクにおいて、発話継続時間が句構造の違いを反映していたこと、またそうした情報が発話を聞いた回答者の解釈に影響を与えていたことが示唆された。

キーワード: 枝分かれ構造, 発話継続時間, チャンク化

### 1. 目的

発話とともに産出される co-speech gesture は言語情報の解釈に影響を与えるとされる(Kendon, 2004; Krauss & Hadar, 1999)。本研究では、自発的に産出された参加者の発話情報がどのように曖昧な句構造の解釈に寄与するのか、チャンク部分の発話継続時間に着目し検討した。

言語は階層構造を持っており、言語そのものの表面上の構造(表層構造)とは別に深層構造があるとされる(Fujita & Fujita, 2021)。例えば、Fujita and Fujita では“green tea cup”の例を取り上げた。この例は、統語構造において、左枝分かれ(LB: Left Branching<sup>1</sup>)することで“tea-cup for green-tea”の意味を持つ句構造ができる。一方、右枝分かれ(RB: Right Branching)することで“green-colored tea-cup”の意味を持つ句構造ができる。言語はこのように階層構造を持っているため、言語の表面からだけでは想定される構文に多様性があり解釈に曖昧性が生じる。そこでこうした句を解釈するためには音韻情報は重要であり、音韻情報により非曖昧化されるとしている(Ito, Arai, & Hirose, 2015)。

この曖昧な句構造に関する非曖昧化は、ジェスチャーによりなされるという主張もある (Kashiwade et al., 2020)。Kashiwade et al.では、この句構造解釈の曖昧性に注目して、ジェスチャーによる非曖昧化の役割を、例えば句「黒いしっぽの大きなねこ」を用いて調べた。この句は、大きなしっぽを持つ黒い猫(big-tail 構造)と、黒いしっぽを持つ大きな猫(big-cat 構造)という、異なる統語構造に対応する2つの意味に解釈することができる。この結果、big-tail 構造では「しっぽの」と「大きな」のチャンクとジェスチャーが同期しやすく、big-cat 構造では「黒い」と「しっぽの」のチャンクと「大きな」「猫」のチャンクの2つがジェスチャーと同期しやすかった。このことから、統語構造に応じてジェスチャーが同期することが示唆された。

Handa et al. (2021)では、動詞が含まれた3語から構成される曖昧な句構造「落下している男の携帯」について、ジェスチャーの頻度とそのタイミングがどのように意味解釈に寄与するのか検討した。Handa et al.は左枝分かれ(LB)の句構造と右枝分かれ(RB)の句構造において、ジェスチャーの発生するタイミングに差があり、この差は句構造におけるまとまりに対応していたことを報告した。よって、構文的に曖昧な構造の曖昧性を解消することにジェスチャーが利用されることを示唆した。音韻情報とジェスチャーの観点から調べた Okahisa and Shirose (2018)では、「黒いテレビの画面」という3語の合成語を利用し、ジェスチャーと音韻情報は同時に使用できないことを議論している。ジェスチャー使用を強制した場合、音韻情報の利用があまりなされないことから、ジェスチャーが話し手に認知的な負荷を与え、認知資源の割り当ての範囲内でジェスチャーが韻律情報よりも優勢になったと議論した。この Okahisa and

<sup>1</sup> Fujita & Fujita(2021)では枝分かれ構造を sub-merge, pot-merge から説明している。

Shirose の音韻情報とジェスチャーが同時使用されないという知見は、ジェスチャーが発話情報を補完するといった co-speech gesture の観点に新たな示唆を加えるもので興味深い。

Kimura, Yasuda, and Kobayashi (in-press) では、co-speech gesture の観点から、コミュニケーションタスク(互いの参加者が伝え合う)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、発話されないが聞き手が伝えられた情報を他の情報と区別し解釈するために必要となる背景情報に関する違いであり、この背景情報を含むものを情報性が高い刺激、含まないものを情報性が低い刺激として設定した。刺激は Handa et al. で利用した刺激文を修正し、モーラ数等を統制したものをを用いた。実験の結果、RB 構造において情報性が高い状況では、最初の名詞の発話の開始時刻と、ジェスチャーのストロークの開始時刻に差があることが示唆された。なお、この知見は Handa et al. が報告した知見と似ていた。RB において情報性が高い場合にジェスチャー開始時刻に差があったことは、Okahisa and Shirose が指摘した観点から考えると、認知的負荷が掛かった可能性を示唆する。しかしながら、Kimura et al. では、ポーズ等の音韻情報の比較はなされていないため、これらの検討は不十分である。

本研究では、Kimura et al. のデータセットを用い、参加者の回答を基に、回答者が正答をした場合に聞いていた発話と、回答者が誤答をした場合に聞いていた発話の 2 つのデータセットを作成し、自発的に産出された発話情報、特にチャンクに関する発話の継続時間が、どのように曖昧な句構造の解釈に寄与するのか、検討した。なお、意味が曖昧な句として「飛んでいるタヌキの写真」を利用した。

もし階層構造の基準となるチャンキング(図1)により非曖昧化が促されるとしたら、1) LB 構造においては、意味の塊を表出するために「飛んでいる」、「タヌキの」に対して発話開始早々にチャンキングを行うことになる。よって、このチャンクの継続時間(chunk 1)は LB 構造の方が RB 構造よりも短くなるだろう。

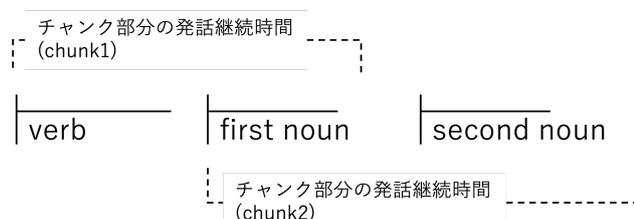


図1 予測されるチャンク部分とその発話継続時間

2) RB 構造においては、「タヌキの」、「写真」に対してチャンキングを行うことになる。「タヌキ」と「写真」を結びつけようとするので、このチャンクの継続時間(chunk 2)は RB 構造の方が LB 構造よりも短くなるだろう。

## 2. 方法

### 2.1 実験参加者

データセットに含まれていた人数は、28 人の日本語を母国語とする学生(平均: 21.7 歳、範囲: 20-23 歳)であった。なお、2 人のペアで実験に参加していた。

### 2.2 実験条件

実験条件は、枝分かれ構造(2: 左枝分かれ(LB)/右枝分かれ(RB) 構造)、情報性(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。枝分かれ構造条件は、句構造が左枝分かれで解釈されるものと、右枝分かれで解釈されるものであった。例えば、3 語の曖昧な句構造として、「飛んでいるタヌキの写真」、動詞(verb)+名詞-1(first noun)+名詞-2(second noun)を使用した(図 2)。

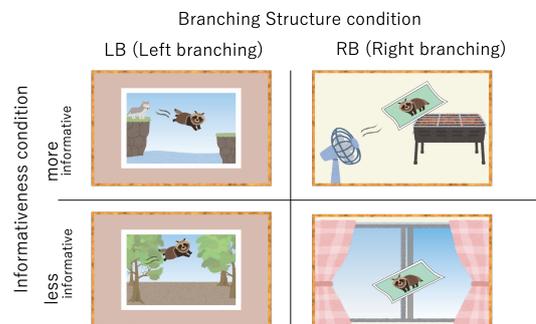


図2 実験条件

この例の場合、左枝分かれ構造では、「飛んでいる」と「タヌキの」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いう解釈ができる。一方で、右枝分かれ構造では、「タヌキの」と「写真」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できる。なお、情報性条件は、情報性が高い場合と、低い場合とした。情報性が高い条件(more informative)は、背景に描かれた自然物や建物(e.g., 崖やビル)など他に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(less informative)は、背景に加える情報はあがるが情報性が高い条件に比べると加える情報は少ない(たとえば平地や木など)と予測されるものであった。動詞条件は、落下している/飛んでい

るという2つの動詞から構成された。これらの動詞は重力のような強い方向性があるものとならないものであり、動き方の表現に差が出ると予測されるものであった。要因計画に関しては、枝分かれ構造条件と動詞条件は参加者内要因、情報性要因は参加者間要因であった。

## 2.3 実験刺激

Handa et al. (2021)の基準を用い、実験刺激は動詞(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を用いた。例えば、「飛んでいるタヌキの写真」や「落下している猫の写真」というような句であった。動詞(V)が2種類、名詞(N1)が6種類、名詞(N2)が3種類から構成される12種類の句を作成した。これらの句について、左枝分かれ構造、および右枝分かれ構造を表すイラストに、情報性の異なる背景、例えば崖から落ちている/木から落ちているというような情報を付加したイラストを作成した。なお、句構造に用いられた語のモーラ数は、条件間で同程度になるように統制した。

## 2.4 実験手続き

本研究では、2人のペアで実験を行った。参加者を、イラストの内容を伝える話し手の役割(参加者A)と、イラストの内容を聞く聞き手の役割(参加者B)にランダムに分けた。話し手(A)と聞き手(B)は、刺激が出力されるモニターが互いに見えないよう、テーブルを挟んで座った。話し手(A)には動画を視聴してもらい、そのイラストの内容を記憶して、聞き手(B)に話してもらった。その後、聞き手(B)には、聞いた話に関して、四択のイラストの中から選んでもらった。参加者は練習試行を行ったあと、本試行に取り組んだ。すべての試行が終了した後、参加者AとBは役割と場所を交代して、再び実験を行った。参加者には、実験後に謝礼を渡すことを伝え、正解率が良いほど謝礼がよくなることを伝えた。

## 3. 分析方法

分析はELAN(ver.6.4)を利用し、枝分かれ構造ごとに発話とジェスチャーをコーディングした。ジェスチャーに関しては、Kendon (2004)のジェスチャーフェーズに対応したコーディングを行った(sec, 細馬, 2008)。句

構造の違いにおけるチャンク部分の発話継続時間を調べるために、聞き手の回答が正答のセット(全て co-speech gesture で表現されていた)と誤答のセット(70%弱が co-speech gesture で表現されていた)の2つのデータセットに分けた。これらのデータセットについて、それぞれ線形混合モデル(LMMs: Linear Mixed Models)を用いた分析を行った<sup>2</sup>。各データセットは、動詞発話の開始時点と名詞-1発話の終了時点とをチャンク1(chunk1)、名詞-1発話の開始時点と名詞-2発話の終了時点とをチャンク2(chunk2)とした。各条件のコード化は、それぞれのデータセットにおいて、ダミーコーディングした後、中心化するようなエフェクトコーディングを行った。各LMMを選定するために、固定効果に実験条件とそれらの交互作用が含まれた最大モデルを構築した後に、前向きステップワイズ法を用いた。解析に利用した各LMMは赤池情報量基準の一番低いものを利用した。

## 4. 結果・考察

句構造の違いによるチャンク部分の発話継続時間において、正答の場合のchunk1のLMMの結果、切片( $\beta = 1.8721, t = 17.174, p < .01$ )、枝分かれ構造(branch;  $\beta = -0.1772, t = -2.266, p = .025$ )、動詞(verb;  $\beta = 0.3186, t = 3.372, p = .007$ )が有意であった。左枝分かれ構造の場合、動詞と名詞-1のまとまりであるchunk1の継続時間が右枝分かれ構造に比べ、有意に短かった。動詞(verb)に関しては、fallのほうがflyよりも有意に長かった。

chunk2の発話継続時間のLMMの結果、切片( $\beta = 1.239, t = 16.42, p < .001$ )が有意であった。chunk2において、RBであれば、発話継続時間が短くなると予測していたが、枝分かれ構造に関わらず、発話継続時間は変わらなかった。これは、より解釈が難しいとされる右枝分かれ構造であると聞き手が解釈するには、名詞-1と名詞-2がまとまりであることが伝わるよう発話する必要があるため、動詞と名詞-1のまとまりであるchunk1が長くなったと考えられる。

誤答に関してchunk1のLMMの結果、切片( $\beta = 1.725, t = 18.430, p < .01$ )、動詞(verb;  $\beta = 0.4341, t = 4.048, p = .0030$ )が有意であった。動詞(verb)に関しては、fallのほうがflyよりも有意に長かった。なお、branchは固定効果として採用されなかった。

<sup>2</sup> 各LMM構築にはlme4パッケージ(Bates, Mächler, Bolker, & Walker, 2015)を用いた。

正答のデータセットにおけるchunk1では、枝分かれ構造(branching)と動詞(verb)が固定効果として選定された( $df=6$ )。chunk2では、nullモデルが選定された( $df=4$ )。誤答のデータセットにおけるchunk1では、動詞(verb)が固定効果として選定された( $df=5$ )。chunk2では、枝分かれ構造(branching)が固定効果として選定された( $df=5$ )。

chunk2 の LMM の結果、切片( $\beta = 1.225, t = 17.338, p < .001$ )、枝分かれ構造(branch;  $\beta = 0.157, t = 2.645, p = .009$ )が有意であった。右枝分かれ構造の場合、名詞-1 と名詞-2 のまとまりである chunk2 の継続時間が左枝分かれ構造に比べ、有意に短かった。

なお、図3に chunk1 の発話継続時間、図4に chunk2 の発話継続時間を示す。なお、これらの図は、枝分かれ構造に焦点を当てプロットした。

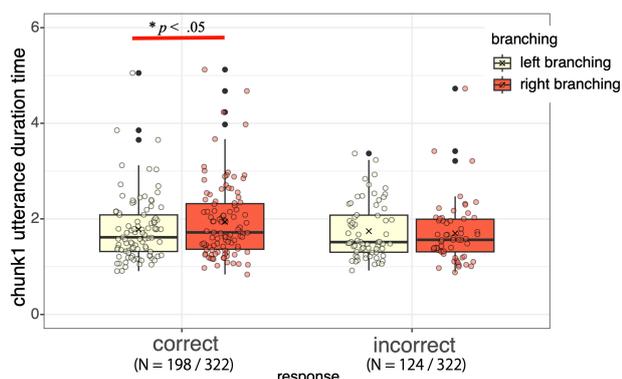


図3 聞き手の回答における句構造の違いによる chunk1 の発話継続時間

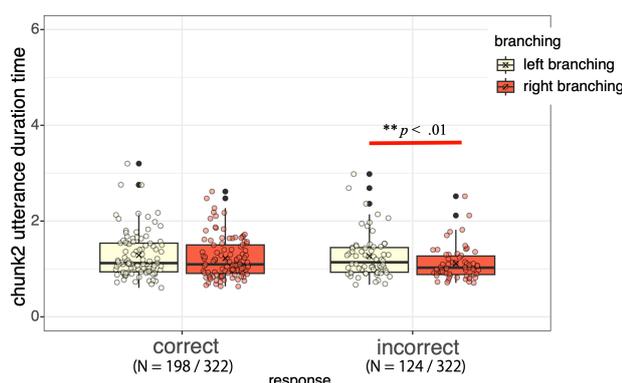


図4 聞き手の回答における句構造の違いによる chunk2 の発話継続時間

以上のことから、chunk1 に関しては、正答でも誤答でも同じように fall のほうが fly よりも長く表現していることがわかった。よって、正答/誤答ともに動詞の意味的情報を付与できていたことが予想される。しかしながら、正答において枝別れ構造に応じた発話継続時間の差異があった一方で、誤答においてはそうではなかった。興味深いことに、chunk2 に関しては、正答においては枝別れ構造に応じた発話継続時間の差異はなく、むしろ誤答において枝分かれ構造の差異があった。これは「適切ではなかった」チャンク構造で表現された発話を聞いたことにより「誤った」答えをするよう導かれた可能性が考えられる。よって、chunk1 の表現の仕方が枝分かれ構造の解釈を正しく行うために、重要であることが示唆される。

この知見は、枝分かれ構造をチャンク化してジェスチャーするという Kashiwade et al. の知見とも異なる。Kashiwade et al. のような想起タスクでは、情報を振り返りながらジェスチャー産出ができるのに対し、本研究でのコミュニケーションタスクでは、伝えるべき相手がいるため、時間という概念も手がかりとなり得る。そこで、枝分かれ構造をまず決める chunk1 のような箇所で、句構造を特定するために必要な情報を付加しようとしていた可能性が考えられる。また、Kimura et al. でのジェスチャー分析では、RB 時に情報性の程度により違いがあったが、本研究での発話継続時間に関して、情報性の違いは確認されなかった。つまりジェスチャーが情報性を表現する点において有効に働いていたことが考えられる。

## 5. 結論

本研究では、伝え手により自発的に産出された発話継続時間が、どのように曖昧な句構造の解釈に寄与するかを検討するために、解釈を行った回答者の正誤を基準とした分析を行った。動詞と名詞のまとまりからなる最初のチャンクにおいて、発話の継続時間は、句構造の違いを反映していた。コミュニケーションのタスクにおいて、発話の早い段階で句構造を特定するために必要な情報が提供されていると言える。今後、無音区間の継続時間を調べることで、より精緻にチャンキング方略等が検討できる可能性がある。

## 謝辞

この研究に協力してくださった全ての参加者に感謝申し上げます。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

- [1] ELAN (Version 6.4) [Computer software]. (2022). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from <https://archive.mpi.nl/ta/elan>
- [2] 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.
- [3] Handa, Y., Yasuda, T., & Kobayashi, H. (2021). The Use of Co-Speech Gestures in Conveying Japanese Phrases with Verbs. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 43, 1555-1559.
- [4] Kashiwade, K., Yasuda, T., Fujita, K., Kita, S., & Kobayashi, H. (2020). Syntactic structure influences speech-gesture synchronization. *Letters on Evo Behav Sci*, 11, 10-14.
- [5] Kimura, H., Yasuda, T., and Kobayashi, H. (July, 2023; in press). Spontaneous co-speech gestures with prompt phrases reflect linguistic structures. *Proceedings of CogSci*, 45
- [6] Okahisa, T., & Shirose, A. (2018). Influence of hand gestures on prosodic disambiguation of syntactically ambiguous phrases. *Acoustical Science and Technology*, 39, 171-17.