

# 物語の大局的理解と局所的理解の相互依存関係の分析

## Analyzing interdependency between comprehension of the whole and parts of story

布山 美慕<sup>†</sup>, 日高 昇平<sup>‡</sup>

Miho Fuyama, Shohei Hidaka

<sup>†</sup> 早稲田大学, <sup>‡</sup> 北陸先端科学技術大学院大学

Waseda University, Japan Advanced Institute of Science and Technology

mih002@sj9.so-net.ne.jp, shhidaka@jaist.ac.jp

### 概要

文章の大局的理解と局所的理解の階層的相互依存的関係は解釈学的循環と呼ばれ、文章理解の基盤の一つとされる。本研究は、この解釈学的循環の認知過程解明を目指し、大局的理解を反映する課題（元童話推定課題）と局所的理解を反映する課題（意味的段落分け課題）を提案し、両課題の回答の構造やその関係性を分析した。その結果、両課題の回答間の一貫した関係性や、意味的段落分けの回答において物語終盤とそれ以外の部分に対し異なる構造が示唆された。これらの構造は解釈学的循環の認知過程の一端を示唆し、その特徴づけにつながる。

**キーワード：**解釈学的循環 (Hermeneutic circulation), 意図推定 (Intension estimation), 物語理解 (Narrative comprehension), 文章分割 (Text segmentation),

### 1. 物語理解を理解するための解釈学的循環過程解明

物語を理解するとはどのようなことだろうか？単語の意味理解、文法理解を含む単文の意味理解を超えて、「物語」の理解には構成する文章の「構造」の理解が含まれると考えられる。こういった文章の構造として、物語文法 [1], global coherence [2] などが考えられてきた。これらの先行研究は物語構造の一部を反映すると考えられるが、説明可能な文章の種類が限定的であるなど [3] の課題が残り、物語構造の特徴づけや定義は必ずしも十分とは言えない。加えて、こういった物語構造が逐次的な読みからどのように構築されるのか、そのダイナミクスは未解明である。

本研究では、文章の局所的理解と大局的理解の相互依存関係を明らかにすることで、読者の物語構造構築過程の解明を目指し、同時に読者が構築する「物語構造」の特徴づけを目指す。ここで、文章の局所的理解と大局的理解の相互依存関係とは、解釈学的循環 [4]

と呼ばれ、文章理解の基盤となる次のような構造である。読者は、文章が全体として伝えたいこと（物語の意図、ある種の物語の要約）を理解するために、その文章を構成する単語や文などの局所的な情報を理解し読み進める必要がある。一方、単語や文の意味は多義性を持ち（比喩的表現の可能性など），“その文章における”単語や文の意図は文脈情報によって定まる。そのため、ある文章における単語や文などの局所的な意味の理解には、文脈情報である上位の意図の理解が必要となる。この理解における階層的相互依存関係である解釈学的循環の認知過程も未だ明らかではない。

我々は、解釈学的循環の構造は階層的・相互依存的な意図のネットワーク構造と見なせ、この構造が文章の構造の一つの表現であり、読者がこの構造を推定することが文章理解であるとの仮説を持つ（図1参照、次段落で説明）。この仮説が正しければ、この意図ネットワーク構造構築過程の解明がすなわち文章構造の構築過程の解明につながる。さらに、物語読書で固有に構築される構造が特徴づけられれば、「物語構造および物語理解とは何か？」への解答に近づくことができると期待する。

図1に示した意図のネットワーク構造と文章理解の仮説について述べる。この仮説では、文章理解を階層的な意図推定・予測過程と見なし、以下の意図ネットワーク構造の推定が文章理解と考える。まず、本仮説において“意図”とは、作者（あるいは作品）がその文表現によって読者に伝達しようとした情報（あるいはそのように読者が思う情報）を指す。読者は各文から推定した局所的な意図から、より上位の意図を推定し、最上位の意図として文章全体の意図を推定する。このときまた、推定された上位の意図から下位の意図が推定され、階層的な相互依存関係をなす（なお、図1では模式的に3層のみ記載しているが、より多層でありうる）。上位の意図ほど情報が圧縮され、階層ごとにそのレベルの意図はその階層レベルのある種の文

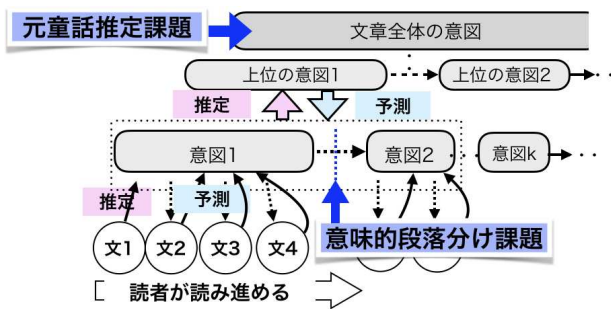


図1 階層的・相互依存的な意図推定・予測過程としての文章理解モデル

章の“要約”（様々なレベル・種類でありうる）と対応すると考える。物語スキーマ [5] や起承転結などの物語構造も、この意図を表現した意図ネットワーク構造の一種とみなせる。読者は現実世界や読書経験から部分的なネットワーク構造の事前知識を持ち、文章理解時の意図推定にはこの知識を利用すると考える。

読書による物語理解は文章理解の一つのクラスであり、説明文に比べ解釈多様性が大きいことから、整合的な意図ネットワークを複数許す（少なくとも部分的には）ような特徴が考えられる。ただし、物語固有の構造の特徴づけは今後の課題とし、本研究では、まず一般に意図ネットワーク中の最上位の意図である文章全体の意図（大局的な意図）の推定過程と下位構造（局所的な意図ネットワーク構造）の推定過程の関係性を行動実験を用いて調べ、解釈学的循環過程の特徴づけを行なった。

## 2. 研究方法

本研究では、最上位の文章全体の意図推定を反映する課題として元童話推定課題を、下位の意図ネットワーク構造推定を反映する課題として意味的段落分け課題を用いる（図1）。順に説明する。

### 2.1 元童話推定課題

文章全体の意図は、ある文章を読み終わった時にその文章全体の最小の「要約」として表現される内容に対応すると考えられる。文章全体の意図の必要条件の一つは、全体の意図のネットワーク構造をある程度制約し、推定可能とすることである。本研究では、よく知られた童話では、この文章全体の意図（最小の「要約」）推定はその童話の“タイトル”推定によって代替可能と考える。この理由は、有名な童話（例えば「浦島太郎」）ではタイトルからその物語内容の要約を想起可能で、全体の意図のネットワーク構造を推定可能

と考えるためである（「浦島太郎」と聞けば、その話の概要が想起でき、「主人公が亀を助ける」→「お礼に亀に竜宮城に連れて行かれる」→「乙姫に歓待される」→「帰る時に玉手箱を渡される」→「数百年経っている」→「玉手箱を開けてしまい老人になる」といった概要に対応する意図ネットワークを元に全体の意図ネットワーク構造を推定できる）。

これを利用し、本研究では、有名な童話の概要的な意図構造を保存したまま新規な物語を実験者（布山）が創作し、その物語がどの童話を元に創作されているか当てる元童話推定課題によって、読者の文章全体の意図推定過程を実験者が推定する。有名な童話そのものでは、被験者が、局所的な情報である固有名詞などを手掛かりにして、全体の意図を推定せずとも、読んでいる物語がどの童話か当ててしまう（たとえば主人公の名前が『浦島太郎』なのでその物語は「浦島太郎」だと推定してしまう）。この固有名詞などの局所的情報によるタイトル推定を避けつつ、読者の文章全体の意図推定過程を調べるため、有名な童話の局所的な情報を改変した上で、その構造を保った物語を創作した（図2参照）。以降、この元童話と構造が同じ創作物語を類推物語と呼ぶ。ここで言う実験者が利用した物語の構造とは、登場人物や登場するアイテムの関係性の有向グラフ構造とその有向の矢印にこめられた意味を指す。部分的な意図ネットワークの一つの表現の仮説として用いた。たとえば、「AさんがBさんを愛す」という物語は

$$A \xrightarrow{\text{愛する}} B$$

というように表せる。この物語に対する類推物語の例は「CがDを愛す」などであり、ラベルAやBは変更されるものの、グラフ構造とその矢印の意味は保たれる。

$$C \xrightarrow{\text{愛する}} D$$

この元童話と対応する類推物語のグラフ構造の例を図2に示した。読者が、表面的なラベル情報ではなく、物語の意図ネットワークの構造の水準で既知の童話の知識を利用できれば、これらの類推物語から元童話のタイトルを当てることが可能と考えられる。

本実験では、10個の童話のリストを渡し、3文読むごとにリスト中のどの童話が元になった童話と思うか被験者に強制回答を求めた。元童話回答の時系列変化が類推物語の文章全体の意図推定過程を反映すると考える（実験手順詳細は後述）。また、実験素材として作成した4話の類推物語のうち2話において、作品の途中で元とする童話を変更し、意図的に文章全体の意図推定過程を変化させることを試みた。

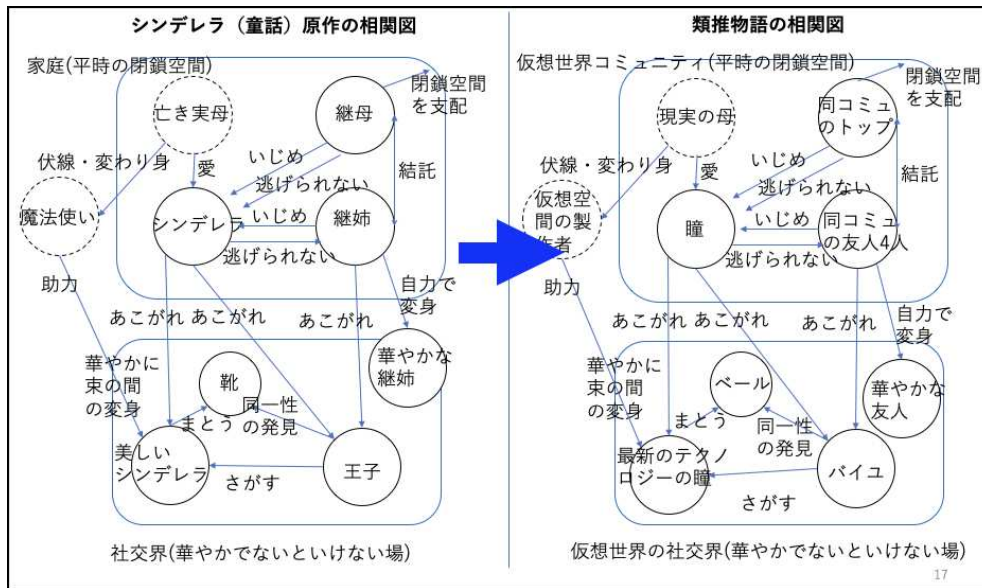


図2 類推物語の構造の対応例：左が元の童話（シンデレラ）、右がその童話を元に創作した類推物語の構造。丸の中が登場人物やアイテムの名称を表し、それらの関係性が矢印を表現する。左右の図で丸の中の固有名詞は変化するが、矢印の関係性は保たれる。

## 2.2 意味的段落分け課題

本研究では下位の意図ネットワーク構造の指標として、文章の意味的段落分けを用いる。読者は文章を読み進める際に、その意味的なまとまりを認識できるとされ、この認知は文章分割の認知過程として研究されてきた [6]。この意味的なまとまりは、一つの物語全体（読了時の話がきちんと「終わった感じ」として認識される）、章、段落、文、単語など様々な単位で認知されうる。この中で、単語や文と言ったまとまりの範囲が明示的なものを除く、最小の意味のまとまりの単位が「段落」である。著者は、これが図1の一番下の階層の意図の変化に対応する単位であり、最小の意図に紐づく文章範囲が一段落に対応すると考える。

以上から、被験者に段落記号を除いた文章を提示し、意味的なまとまりに応じた段落分けをしてもらう課題（意味的段落分け課題）の成績が、読者の下位のネットワーク構造推定の指標と見なせると考えた [7]。ここで、次の文が読めない条件とした理由は、段落をまたぐ2文の間での情報の差分情報を利用し、この差分が大きい部分を段落終端と判断してしまうと、意味のまとまり判断を反映する課題とならないためである。

本実験では、元童話当て課題と同じ類推物語を（元童話当て課題の被験者群とは別の被験者群に）一文ずつ提示し、次の文が読めない状況で、その文が（意味的な）段落の最後の文だと思える程度の回答を五件法で得た。この被験者回答の変化や、回答と作品の元の段

落終端との一致度、被験者間での一致度を下位ネットワーク推定過程の指標とみなして分析した。

## 3. 実験方法

### 3.1 被験者

北陸先端科学技術大学院大学の日本語母語話者の大学院生および教員を被験者とした。10名が元童話推定課題、別の10名が意味的段落分け課題を行った。

### 3.2 実験素材：類推物語

類推物語は、物語1が「鶴の恩返し」、物語2が「シンデレラ」、物語3が「白雪姫」と「赤ずきん」、物語4が「眠れる森の美女」と「浦島太郎」を元に創作した。物語3では「白雪姫」を元にした部分から「赤ずきん」を元にした部分へと、物語4では「眠れる森の美女」から「浦島太郎」を元にした部分へと物語が変化する。各類推物語の文数と文字数は、物語1が72文・2730文字、物語2が109文・3983文字、物語3が92文・3232文字、物語4が143文・4785文字である。

### 3.3 実験手順

元童話推定課題では、被験者は類推物語を読み、3文ごとに元となった童話だと思えるものをリスト中の10話の童話から1つ強制選択にて回答した。回答と同

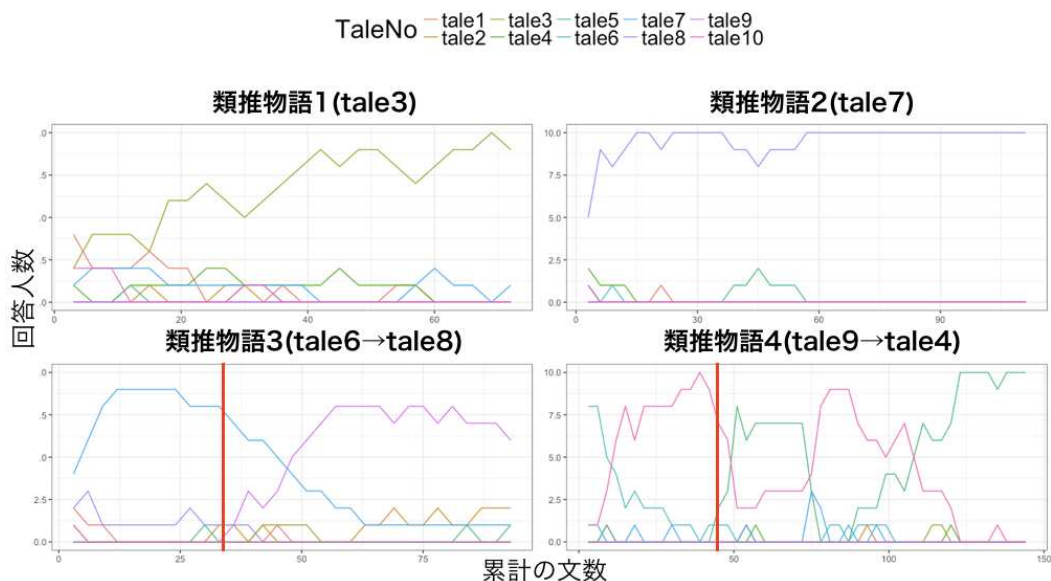


図3 元童話推定課題回答の時系列変化. 被験者10名中何名がその元童話を回答したかを読み進めた文数(累積分数)に対してプロットした. TaleNoは元童話の回答番号を, 各グラフタイトルの括弧内部は正答の番号を示す. 類推物語3と4のグラフ中には元童話の正答の変化点を赤線で示した.

時に, その3文の中で回答の根拠となった箇所があれば赤ペンで下線を引くよう求めた. 回答箇所より以前の文章の読み返しは許可し, 回答箇所より以降の文章を読むことは禁じた. 被験者は以上の教示を受けたのち, まず, 10話の童話タイトルのリストとその回答候補の童話10話の概要を読み, その後4話の類推物語を順に読んだ. 4話の提示順は被験者によって異なるようランダムに定めた. 前述の通り, 4話の類推物語のうち, 2話では途中で元とする童話を変化させた. 4話への回答終了後, 実験への感想や直近一ヶ月間の読書量等を尋ねるアンケートへの回答を求めた.

意味的段落分け課題では, 被験者は元童話推定課題で使用した4話の類推物語を一文ずつ読み, その文で段落が終わると思う程度を5件法(1:全く終わると思わない, 2:あまり終わると思わない, 3:どちらとも言えない, 4:やや終わると思う, 5:強く終わると思う)から強制選択にて回答した. 回答箇所より以前の文章の読み返しは許可し, 回答箇所より以降の文章を読むことは禁じた. 4話の提示順は, 元童話推定課題の提示順と同じにし, 被験者によって異なるようにした. 各類推物語の最後にはフィラーとして無関係の物語の文を5~10文程度追加した. 4話への回答終了後, 実験への感想や, 読んだ物語について気づいたこと, 直近一ヶ月間の読書量等を尋ねるアンケートへの回答を求めた.

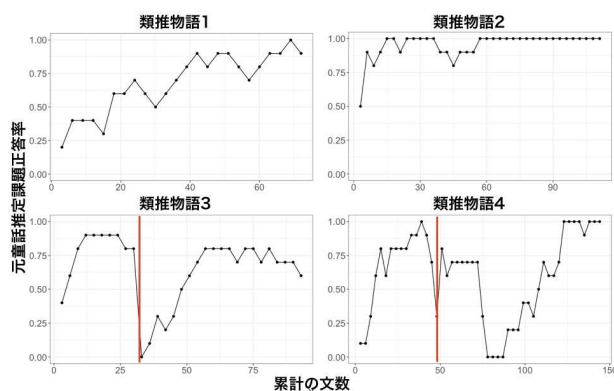


図4 元童話推定課題回答の正答率の時系列変化.

## 4. 実験結果

### 4.1 元童話推定課題回答の時系列変化傾向と課題の妥当性確認

まず, 各課題回答の傾向を独立に確認し, 各課題が物語の大局的意味理解・局所的意味理解を反映した結果とみなせるか検討した.

元童話推定課題の回答の時系列変化を図3に, 対応する元童話正答率を図4に示す. 類推物語1と2では正答の元童話回答が単調増加する傾向が見られ, 類推物語3と4は元童話の変化に伴い回答が変化する傾向が確認された. 図3, 4の類推物語3と4のグラフ中, 赤線で示した時点(30~33文目)で元童話を変化さ

せているが、この変化点に合わせて最大の回答数を得た童話が変化することがわかる。ただし、類推物語4では、実験者（類推物語創作者）の意図は赤線の時点（44～47文目）一回のみ元童話を変化させる（「眠れる森の美女」→「浦島太郎へ」）ことだったが、被験者の回答では3回最大の回答数を得た童話が変化した（「眠れる森の美女」→「浦島太郎」→「眠れる森の美女」→「浦島太郎」）。この理由は、類推物語4では、70文目あたりに再度前半部分の童話（「眠れる森の美女」）が元童話としても解釈可能な記述があり、この解釈多様性をもつ記述によって被験者が再度元童話が変化したと考えたことに起因すると推測される<sup>1</sup>。

全体の傾向として、読み始めを除く多くの時点で、多くとも2つの童話に被験者の回答が集中する。確度が低い時点においても、なんらかの意味で“似ている”童話を被験者が回答する（なんらかの理解がある）ことが示唆される。

以上の結果から、回答がほぼ実験者の意図した正答と一致すること、また一致しない場合であっても被験者間の一貫性は保たれることが示唆された。

加えて、類推物語3と4での元童話回答の変化が緩やかであることから、被験者は局所的情報（一単語など）ではなく、物語全体の構造など、より大局的な情報に依拠して元童話回答を変化させることが示唆された。この確認のため、物語構造を表現した図2のエッジに構造の情報があると考え、各文に対しこの各エッジ内容が含まれる文か否か（その文が構造の情報をもっているか否か）を実験者（類推物語創作者）がアノテーションした。その上で、元童話が途中で変化する類推物語4において、元童話回答と含まれるエッジ情報（どちらの元童話のエッジ情報が含まれるか、またはどのエッジ情報も含まないかの三区分）の関係性をカイ二乗検定で調べた。その結果、含まれる元童話のエッジ情報と同じ元童話が有意（ $p < 0.001$ ）に回答されることが示唆された。この結果も、元童話回答が物語構造の情報に依拠してなされることを示唆する。

以上を総合すると、元童話推定課題の回答は、物語の大局的な意図推定を反映しており、かつ被験者間で一貫性した意図推定がなされたことが示唆される。

## 4.2 元童話推定課題回答の構造

4.1の分析から元童話が1つの場合（類推物語1・2）では正答率が単調増加する傾向や、元童話が2つ

<sup>1</sup>類推物語4におけるこの元童話回答の変化は実験素材の統制の不十分さによるものであるが、実質的に回答の変化点が増えたことで以降の分析では面白い結果が見えた。

の場合はその2つの童話の間で回答が切り替わる傾向が見られた。これらの回答傾向が示唆する大局的理解の構造をより詳しく調べるため、元童話推定課題回答を少数の次元に落としてその構造を調べた。具体的には、元童話の回答をダミー変数に変換し、各データを10名×10種類の元童話の回答＝100次元のデータとし、主成分分析を行った。各データ点の主成分得点を第一・第二主成分の二次元空間上にプロットして示す（図5。各類推物語データの第二主成分までの累積寄与率は40%～65%）。時系列データのため、各点を順に直線でつなぎ、データ点の番号を付した。

全ての類推物語で、第一主成分軸に平行な軌跡と第二主成分軸に並行な軌跡が合わさった直角を成すような軌跡が確認できる。この結果は、回答の時系列変化（図3）と合わせ以下のように解釈できる。

まず、直角を成す二辺のうち、データ点の少ない方の辺（類推物語1, 3, 4では第二主成分軸に、類推物語2では第一主成分軸に並行な辺）は、読み始めの元童話が定まらない探索時期に対応する。一方、もう片方のデータ点が多い辺は、類推物語1・2では元童話の回答が全被験者でほぼ一つに定まった状態、類推物語3・4では元童話の回答が全被験者でほぼ2つに定まった状態に対応する。この解釈は主成分得点と主成分負荷量のbiplotの結果（紙幅の関係で省略）にも整合的で、とくに類推物語3と4では第一主成分軸の両方向がそれぞれ正答の2つの元童話の回答に対応し、この2つの元童話の選択過程が表現されていると解釈できる。

まとめると、元童話推定課題の回答を低次元空間で表現した結果、10話から探索する時期（探索時期）と1話で安定あるいは2話から選択する時期（少数からの選択時期）に分けられることが示唆された。このことは、大局的な意図理解が、広く意図を探索する時期時期と、より少数の意図に限定して意図理解を行う（少数の意図理解を深めるような）時期に分けられることを示唆する。

## 4.3 意味的段落分け回答の一貫性検討と課題の妥当性確認

意味的段落分け課題回答に被験者間で一貫性があるか、また実験者が創作した元の類推物語の段落分けをどの程度当てられているか分析した。被験者10名の相関行列を求めたところ、相関係数は0.01～0.37で全45ペア中34ペア（76%）が $p < 0.05$ 水準で有意な相関を示した。また、被験者が元の文章の段落終端を検

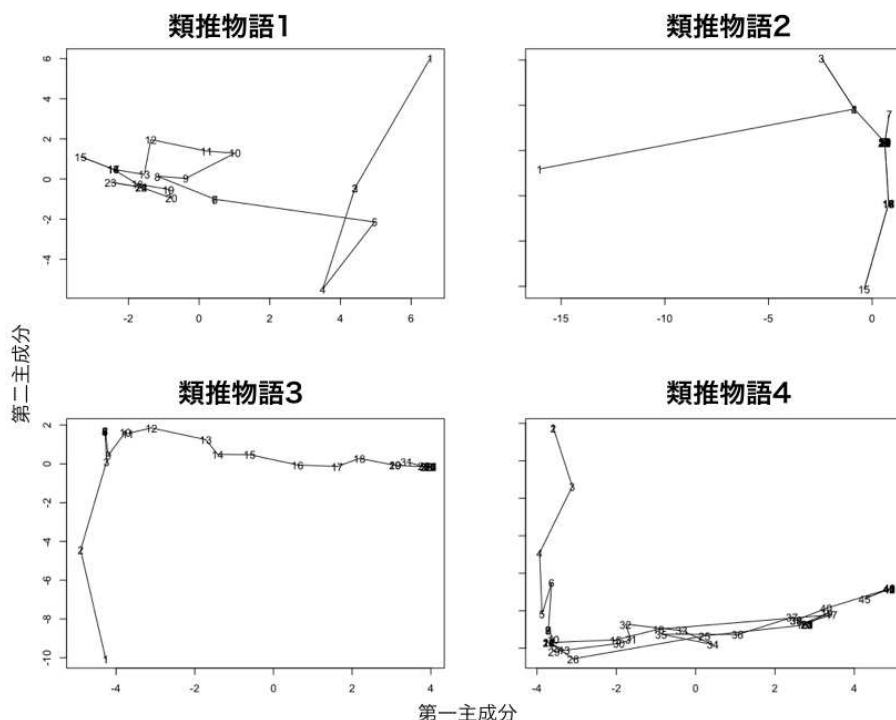


図5 元童話推定課題回答の主成分分析結果。各データ点の番号を記入（文字が重なって見辛いが時系列の順を示す）

出可能か、信号検出理論の  $d'$  prime（信号とノイズをどの程度弁別可能かを示す）を指標として調べた。分析の結果、 $d'$  prime = 0.47 となり、 $d'$  prime は標準偏差相当を単位とみなせるため、ほぼ信号（段落終端）をノイズ（それ以外の文）と標準偏差 0.5 個分程度の精度で弁別可能であることが示唆された。

以上の結果から、被験者は段落分けをなんらかの一貫した手掛かりを元に行っており、その回答は元の段落分けに対応することが示唆された。よって、被験者の文章の段落単位の構造推定を本課題が反映すると考えられる。ただし、段落終端に特徴的な文末表現等のみに依拠した被験者の段落分けの可能性を棄却する分析は行っていない。そのため、段落内部の意味のまとめりではなく、文末表現などより局所的な情報による意味的段落分け判断の可能性が残っており、今後の課題である。

#### 4.4 意味的段落分け回答の構造

次に、元童話推定課題の分析と同様に、意味的段落分け回答を低次元表現し、示唆される局所的構造の構造を調べた。意味的段落分けデータを順序尺度とみなして主成分分析を行った（分析時にはフィラー文への回答は除いた）。結果を図6に示す（第一・第二主成分

の2次元空間上に主成分得点によって各データ点をプロット、第二主成分までの累積寄与率は36%~45%）。図6では、真の段落終端の文を赤丸で、それ以外の文を青丸で示した。文番号は見辛い付していないが、いずれも密度の高い点の集合から外れた端点が類推物語最後の文に対するデータ点である。物語終盤データがそれ以外のデータの密集部分から離れる特徴は、意味的段落分け回答が各類推物語終端部分とそれ以外で異なる特徴を持つことを示唆する。

また、真の段落終端の文（赤丸のデータ点）がそれ以外の文（青丸のデータ点）に比べ、原点（全データ点の重心）からの距離が大きく見える。そこで、距離の平均値に有意差があるかウェルチの検定を行った。その結果、各類推物語での各文集合（赤と青）の原点からの距離の平均値と検定結果の  $p$  値（有効数字2桁）はそれぞれ、類推物語1：段落終端文平均距離 = 2.2, それ以外の文平均距離 = 1.7,  $p = 0.17$ , 類推物語2：段落終端文平均距離 = 2.3, それ以外の文平均距離 = 1.6,  $p = 0.094$ , 類推物語3：段落終端文平均距離 = 1.9, それ以外の文平均距離 = 1.5,  $p = 0.16$ , 類推物語4：段落終端文平均距離 = 2.0, それ以外の文平均距離 = 1.6,  $p = 0.054$ , となった。全類推物語で段落終端文の方がそれ以外の文よりも原点からの距離が大きく、その差は類推物語2と4では有意水準  $p = 0.1$  で

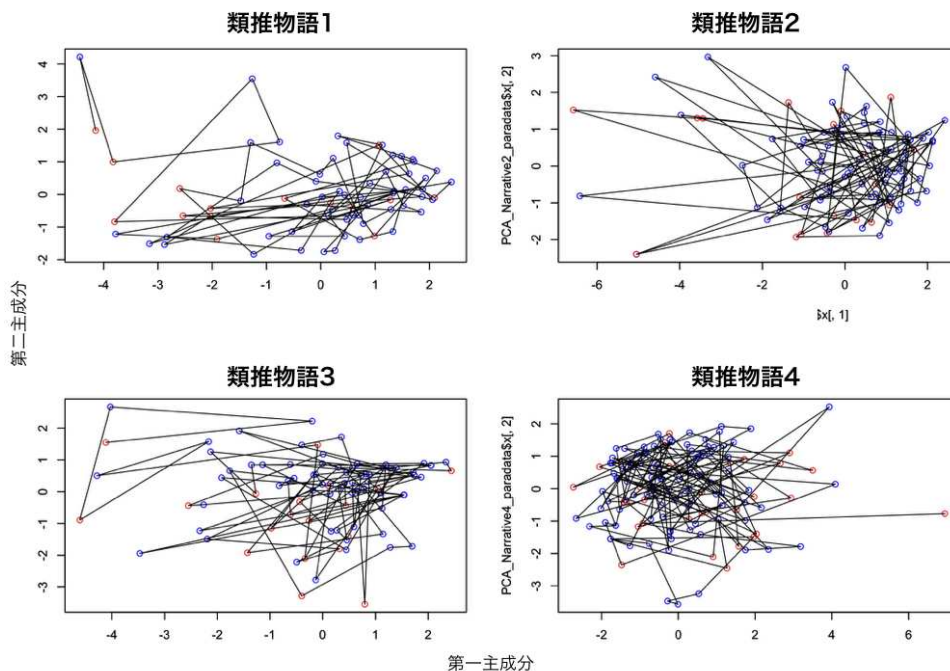


図6 意味段落分け課題の主成分分析結果。類推物語の真の段落終端の文を赤丸で、それ以外の文を青丸で表した。段落終端の文がそれ以外の文に比べて原点から離れた場所にある。

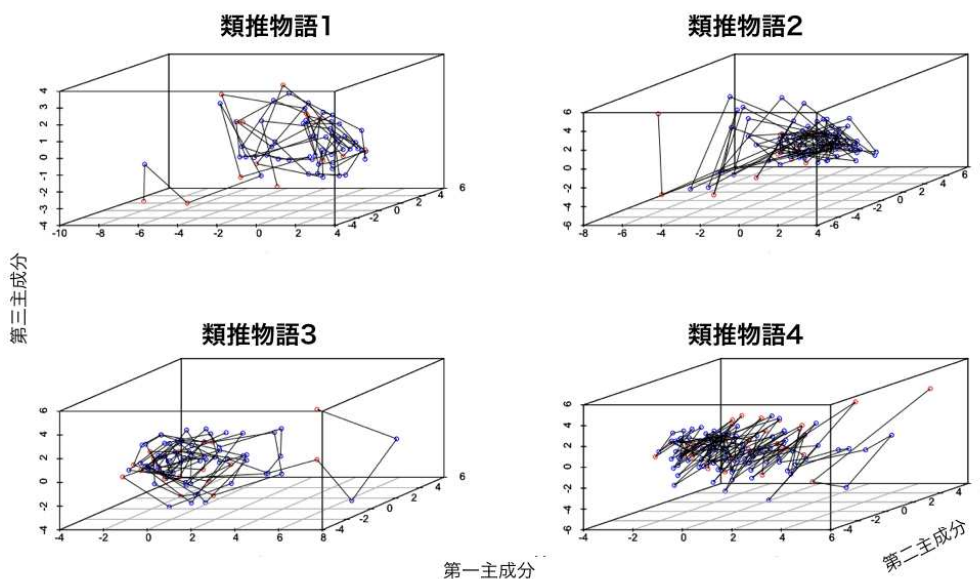


図7 意味段落分け課題の遅延座標系に対する主成分分析結果

有意となった。両主成分の解釈が難しく、原点から距離があることの意味は今のところ明確でないが、被験者が真の段落終端を当てられていた点と合わせ、被験者の意味的段落分け回答の構造と真の段落構造に一貫した関係性があることを示唆する。

加えて、より明確に意味的段落分けのダイナミクスを見るため、遅延座標系への埋め込みを行い、その

データに対して主成分分析を行った。具体的には、各被験者の各文に対する回答データに、2文前までの回答データを加えて1データとし（10名×3時点の回答＝30次元のデータ）、主成分分析を行った。この分析は、2文前までの意味的段落分け評価値を用いた、人の意味的なまとまり判断の構造推定とも解釈できる。

この30次元に対する主成分分析結果を図7に示す。

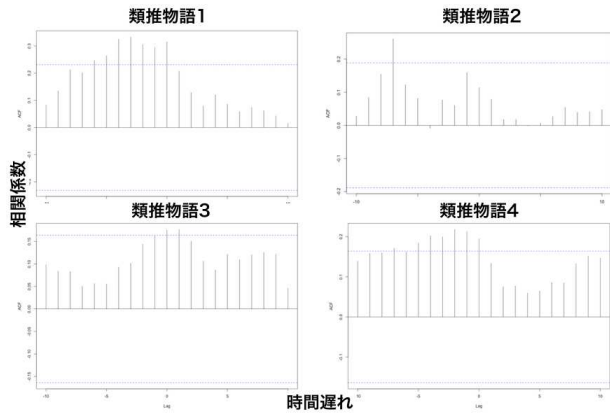


図8 元童話推定課題正答率と意味的段落分け評価被験者間平均値の交差相関分析：横軸は意味的段落分け評価平均値を基準として、元童話推定課題正答率のを示す。青い点線が有意水準。段落分け評価の被験者間平均値に先行して（あるいはほぼ同時に）元童話推定の正答率が同傾向で変化することが示唆された。

第三主成分までの三次元空間上に主成分得点に基づき各データ点をプロットした（第三主成分までの累積寄与率は27%～35%）。各データ点は前述同様に真の段落終端の文か否かで赤/青で色分けした。前述の10次元データに対する主成分分析結果と類似に、物語終端付近のデータがそれ以外のデータ集合から外れて存在することがより明確に確認できる。また、4つの類推物語に共通して、物語終端以外のデータ点は周期的な軌道を描き、物語終端に近づくにつれて周期的軌道の一周の距離が大きくなり（周期軌道が解けていくように見える）、最後に周期的軌道から外れて物語終端を迎える様子が確認できる。この段階的な構造の変化は、意味的段落分け回答が物語理解過程の何らかの段階に対応したフェーズを表現している可能性を示唆する。

#### 4.5 両課題回答の関係性分析

まず、両課題回答の関係性を簡単にみるため、元童話推定課題の正答率6割以上（4話の類推童話合計で325文）と6割未満（同90文）で文の集合を作り、各文集合で意味的段落分けの被験者間相関が有意（ $p < 0.05$ 水準）のペア数を比べた。正答率6割以上の文に対しては36ペアが有意、6割未満に対しては6ペアが有意であり、両指標は独立ではなく、統計的な関係があることが示唆された。両者の時間遅れを含む相関関係を調べる交差相関分析では、段落分け評価の被験者間平均値に先行して（あるいはほぼ同時に）元童話推定の正答率が同傾向で変化することが示唆され

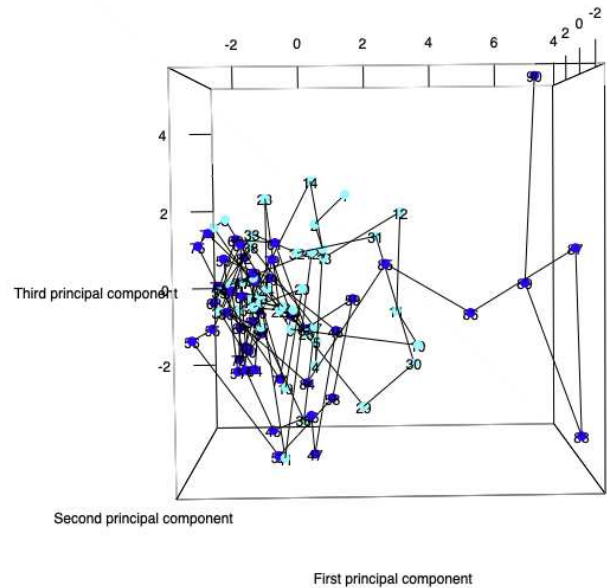


図9 類推物語3の意味的段落分けの時間遅れを含む30次元のデータに対する主成分分析結果を、元童話推定課題の各点での最大の回答数を得た童話で色分けした。

た（図8）。この結果は、上位の意図推定が先行して変化し、下位の意図推定の変化が後続する傾向を示唆する。

次に、4.4の意味的段落データの遅延座標に対する主成分分析で示した通り、意味的段落分け回答は物語理解の何らかのフェーズを反映した構造を持つ可能性が考えられた。この結果を受け、この意味的段落分けの主成分分析結果の各データ点を元童話推定課題の回答に対応させて色分けし、両者の関係性を調べた。具体的には、最大の回答数を得た童話によって色分けした（図9、図10）。ただし、類推物語1と2は元童話推定において開始後即時あるいは数文で正答の回答が過半となっており、本分析の意味がない。このため、類推物語3と4の結果のみを示す。

類推物語3の結果（図9）を見ると、元童話推定回答の2種類（青と水色）が層状に分かれて存在することが確認できる。水色が前半の元童話回答の正答（「白雪姫」）、青色が後半の元童話回答の正答（「赤ずきん」）に対応する。水色が第一主成分軸と第三主成分軸方向に少しほだけたような構造をして層になっているのに対し、青色は（物語終盤を除いて）より密な層として第一主成分の負方向にまとまって見える。もし元童話推定の回答の種類によって層を構成しているならば、大局的な「元童話が何か」という理解と局所的な意味的まとまりの理解の関連性を示唆すると言える



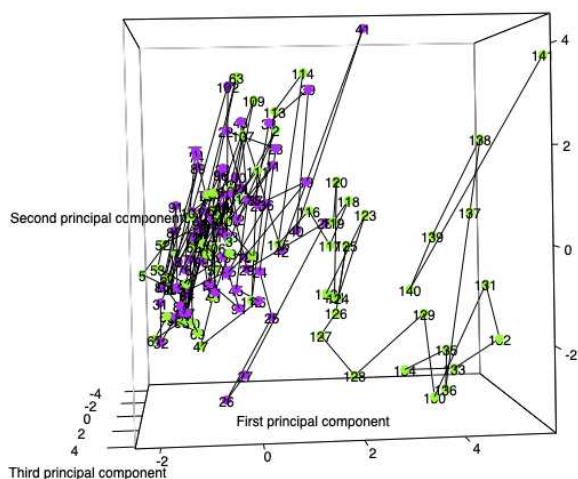


図 10 類推物語 4 の意味的段落分けの時間遅れを含む 30 次元のデータに対する主成分分析結果を、元童話推定課題の各点での最大の回答数を得た童話で色分けした。

だろう。ただし、水色も一部青の集団の中に混ざって見える部分もあり、読み始め部分の特異性や時系列の連続的な変化によって各元童話推定の回答がまとまって見える可能性も現時点では否定できない。

次に類推物語 4 の結果 (図 10) を、元童話推定回答傾向のグラフ (図 3) を参照しつつ検討する。まず、大まかに見て、このプロットには 3 層の構造が見て取れる (図 11 参照)。1 層目は第一主成分の値が  $-2 \sim 1$  程度の間位置する最も密なデータ集合の層 (図 11 赤の楕円で囲った部分)、2 層目は 1 層目から少し離れ 3 層目との中間に位置する層 (同青の楕円)、3 層目は物語の最終盤部分に対応する層 (同水色の楕円) である。

1 層ずつ詳しく見る。1 層目は密度が高く元童話回答の種類も混在しているため、3 層目から考察する。3 層目は、データ番号 129 以降 (第一主成分の得点が 4 以上)、元童話推定回答が被験者 9 名以上で一致する時点以降のデータを含み、他の類推物語の結果と同じように物語最終盤に特徴的な、周期的構造から抜け出す部分に対応すると考えられる。

次に、2 層目には紫 (元童話「眠れる森の美女」に対応) のデータ集合と緑 (元童話「浦島太郎」に対応) のデータ集合が含まれるため、順に考察する。まず紫 (元童話「眠れる森の美女」に対応) の 24-42 番は、元童話推定で読み始めの探索時期を終えて、8 名以上の被験者の回答が正答の「眠れる森の美女」に集中し、次の童話「浦島太郎」に変化する直前までの部分であ

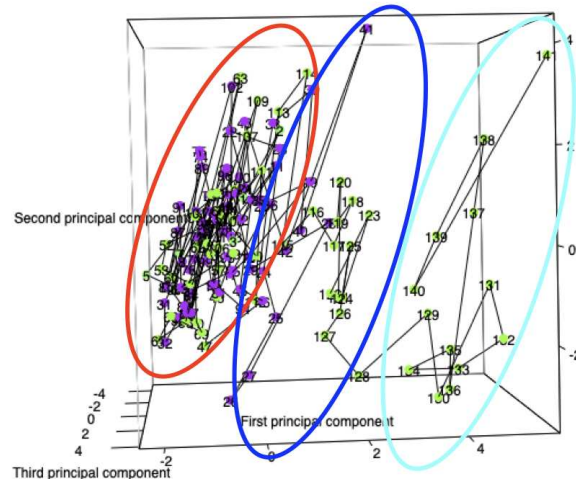


図 11 図 10 に対し、分析者の層の解釈を加えた。1 層目=赤、2 層目=青、3 層目=水色で囲った集合と見なす。

る。元童話回答の 1 回目の変化点 (「眠れる森の美女」から「浦島太郎」へ変化。なお童話の変化点のカウントは読み始め部分の回答変化を除きカウントする) は 44 文目 (43 文目まで「眠れる森の美女」) であり、この遅延座標系の主成分分析では 2 時点ずれるのでちょうど 42 番目のデータに対応する。物語の内容は、この変化点において、物語中 100 年の時間が経過し、主な登場人物にも変化がある (話自体は 1 つの話としてまとまるよう変化点以前のストーリーと繋がっている)。また、3 層目の最終盤とこの部分の軌跡の形状は、大きく第二主成分軸上で値が変化する点で類似する (図 11 では見えないが第三主成分の絶対値も共通して大きな値をとる)。これらを総合すると、この 2 層目 24-42 番目の部分は、元童話の推定が定まった後の物語終端特有の読解に対応する構造と解釈できる。つまり、被験者はこの部分を、物語終盤と捉えて処理していた可能性がある。一方、緑 (元童話「浦島太郎」に対応) の 114 番目から 128 番目は、元童話推定課題の最大の回答数を得た童話が「眠れる森の美女」から「浦島太郎」に変化し、「浦島太郎」回答が全 10 名まで増加する部分である。紫と同様に物語の変化点でもあるが、むしろ他の類推物語 1~3 の物語終盤に向かう周期的軌道がほどけていく部分に対応すると解釈できるかもしれない。つまり、物語終盤に向かうことを予想し、第 1 層目から徐々に離れた軌道を取る部分との解釈ができる。<sup>2)</sup>

<sup>2)</sup> 2 層目、特に第二主成分の値が大きいデータ点に関連し、元童話の文章には共通点がある。43 文目 (データ番号 41) 「そうして、百年後に目覚めたときまで安全な場所で守られることになった。」

1層目は、読み始めおよび75文目あたりの2回目の元童話回答の変化点のデータを含むが、それ以外は物語終端に関連しないデータ点の集合と見なせる。類推物語3と異なり、元童話ごとに層が分かれる構造は確認できない。

以上を総合し、意味的段落分けの遅延座標への埋め込みを行ったデータの構造は、元童話の回答自体の変化よりも、物語の終端部分での理解とそれ以外の部分での理解の処理の違いをより強く反映すると考えられる。類推物語4では、元童話の変化点直前が物語の終盤部分と被験者に判断されたため、一つの類推物語の読書の中で二回物語終端部分における理解の処理が観測されたと考える。類推物語3での元童話変化点ではこの構造が確認できない理由としては、類推物語3の元童話の変化点では時間や空間の変化が大きい一方、類推物語4では変化点で100年間の時間変化や登場人物も変化するため、より（ある種非連続的な）構造となっていたことが影響した可能性が考えられる。各類推物語の終盤においては、残文が少ないというメタ情報から意味的段落分け回答が特徴的になる可能性も考えられた。しかし、類推物語4の1回目の回答変化点において物語終盤と対応する意味的段落分け回答の構造が確認できたことによって、読者が残分数情報によらず物語終端の判断と関連づけて意味的段落分け回答の構造を変化させることが示唆された。意味的段落分けという最小単位の意味のまとまり判断のデータが、物語全体の終端（物語全体のまとまり）を示唆する点は興味深い。

## 5. 総合議論と今後の課題

本研究では、文章の大局的理解と局所的理解の相互依存関係の解明に向けて、元童話推定課題と意味的段落分け課題の回答の関係性分析を行った。現在までの分析から、新規に提案した元童話推定課題が大局的意図理解を、意味的段落分け課題がより局所的な意図理解を反映する課題と見なせることが示唆された。また、元童話推定課題の正答率と意味的段落分け課題の被験者間相関の関係性から、両者は同時点（あるいは多少の時間遅れを伴って）に一貫した相関関係を示すことが示唆された。加えて、意味的段落分け回答データの遅延座標系での主成分分析結果に元童話推定課題

116 文目（データ番号 114）「それで、彼に百年程度仮死状態となる薬を飲ませ、自分も同じ薬を再度飲んだのだった。」など、長い時間の経過を表す文に対応するデータ点において、2層目の軌道中第二主成分の値が突出して大きい。このことは、物語の大きな時空間的な変節点（ある種の終端と始点を示唆し、場合によっては物語自体の終端を示唆する）の予測が第二主成分の大きな絶対値に対応するとも解釈できる。

の回答情報を加えて分析することで、「物語終盤」という大局的な意図理解のフェーズが、意味的段落分けというより局所的な読解のデータから推測可能なことが示唆された。この結果は、意味的段落分けの認知処理の特徴が、より大局的な物語終端に関する推定・予測と関連して変化することを示唆し、解釈学的循環の一つの構造を示すものと著者らは考える。

著者らは、意味的段落分けの主成分分析結果（元の座標系および遅延座標系における）において、少数の次元のみでは寄与率があまり高くない点や、構造に周期性が見出せる点から、元データが非線形の構造を持つと考え、多様体学習による構造分析も試みた。しかし、データ点の少なさから、現在までの分析では構造を見出せない。また、ベクトル自己回帰モデル等やマルコフ転換モデルを用いた時系列分析も試みたが、現在まで報告できる結果は少ない。今後、両課題の回答データに加えて、図2の物語ダイアグラムの分析や、元童話推定課題で回答の根拠部分に引かれた赤線データなどを活かして分析を進め、議論を深めたい。加えて、より長文・より多人数による実験を計画し、十分なデータ数に対する非線形の分析も試みたい。

## 謝辞

本研究は科学研究費補助金若手研究A（16H05860）の補助、および早稲田大学特定課題研究助成費（課題番号：2019C-602）を受けた。

## 文献

- [1] Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive psychology*, 9(1), 77-110.
- [2] McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. *Psychology of learning and motivation*, 51, 297-384.
- [3] De Beaugrande, R. (1982). The story of grammars and the grammar of stories. *Journal of Pragmatics*, 6, 383-422.
- [4] Gadamer, H. G. (1975) *Wahrheit und Methode. Grundzüge einer philosophischen Hermeneutik.* J.C.B. Mohr (Paul Siebeck) Tübingen (饗田収訳 (1986) 『真理と方法』法政大学出版局.)
- [5] Alba, J. W., & Hasher, L. (1983). Is memory schematic?. *Psychological Bulletin*, 93(2), 203-231.
- [6] Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*, 138(2), 307-327.
- [7] Fuyama, M. & Hidaka, S. (2018). Can readers recognize unit of summarization for reading?: an analysis of text segmentation task. *The Proceedings of the 28th Annual Conference of the Japanese Neural Network Society*, 108-109.