

単一述語項関係アノテーション課題における 視線情報の収集と分析

光田 航^{1,a)} 飯田 龍^{1,b)} 徳永 健伸^{1,c)}

概要：テキストアノテーション時の作業者の振舞いは，作業者が文章を理解する認知的な過程を反映していると考えられる．したがって，この振舞いを分析し，アノテーションを行う際に作業者がどのような点に着目しているかを明らかにすることで，深い言語処理やコーパス構築に役立つ知見を獲得できると考えられる．我々はこれまでに日本語述語項関係アノテーションにおける作業者のアノテーションの履歴と視線情報を収集したが，この際に採用したデータ収集の方法では，一つの文章全体を単位としてデータ収集を行ってしまうため，個別の述語項関係のアノテーション過程に関する分析が困難になるという問題が発生した．そこで，本研究では，文章中に出現する一つの述語に対してその項をアノテーションする作業を設定し，作業者の振舞いのデータを収集する．この収集方法にしたがってデータを収集することで，個別の述語項関係の分析が容易になると考えられる．本稿では，このデータ収集の方法とその収集結果，また収集したデータに基づく予備的な分析結果について報告する．

1. はじめに

個々の事例に対して正解のタグをアノテーションしたコーパスを構築し，そのコーパスを用いて学習を行うアプローチは，自然言語処理の多くの問題解決に貢献してきた．特に，形態素解析や統語解析のように単語や品詞などの表層的な情報が有効な問題では，このアプローチを採用することで高い解析精度を得ることができている．一方，述語項構造解析や談話関係解析などの問題では，単語などの表層的な情報に加えて，WordNet[7]などの整備された知識やセンタリング理論[2]などの談話の理論から得た知見を導入して問題の解決を試みているが，それだけでは形態素・統語解析ほどの解析精度を得ることはできない．このため，これまでに行われてきたエラー分析のように人間の内省に基づいて言語現象を調査することが必要とされるが，必ずしも内省だけから解析に有益な手がかりが明らかになるとは限らない．そこで，本研究では，別の分析の方向性として認知科学的なアプローチを採用する．具体的には，作業者がアノテーションを行う際に文章のどこに着目し，どのように作業を行ったかを分析することで，内省では得ることができない知見を得ることを目的

とする．

我々の先行研究[12]では，日本語述語項関係アノテーションを対象に，アノテーション時の作業者の視線情報とツールの操作履歴を収集した．さらに，収集したデータを利用して，作業者が各述語に対して項をアノテーションする過程の分析を行った[3]．しかし，この研究で採用したデータ収集では，作業者が文章中に出現する全ての述語に対し一度にアノテーションを行ったため，作業者が個々の述語に対してどのように項を探索してアノテーションを行ったのかを分析することが困難になる問題が発生した．そこで，本研究では，この問題を解決するために，文章中の一つの述語に対して作業者が項をアノテーションする際の振舞いを個別に収集することで，個々の述語項関係についての作業者の振舞いの分析を可能にする．以降，このアノテーション課題を単一述語項関係アノテーション課題と呼ぶ．

本稿では，単一述語項関係アノテーションを対象に，アノテーション時の作業者の振舞いを収集する方法とその収集結果，また収集したデータに基づく予備的な分析結果について報告する．まず2節で先行研究で採用したデータ収集の方法と問題点について説明し，3節でその問題点を解決するために設計した単一述語項関係アノテーション課題について述べる．4節でこのアノテーション課題を対象にしたデータ収集実験について紹介し，5節で収集したデータの概要を述べる．6節で収集したデータに関する予備的

¹ 東京工業大学
Tokyo Institute of Technology
^{a)} mitsudak@cl.cs.titech.ac.jp
^{b)} ryu-i@cl.cs.titech.ac.jp
^{c)} take@cl.cs.titech.ac.jp

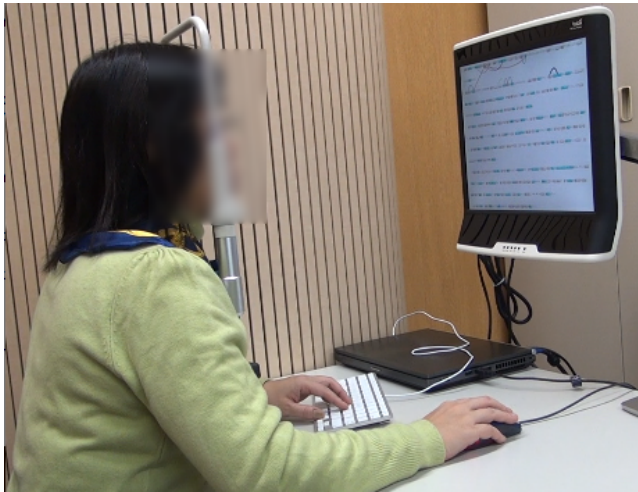


図1 Tobii T60 を使ったアノテーションのスナップショット

な分析の結果について述べ、7 節でまとめと今後の課題を述べる。

2. 先行研究におけるデータ収集実験

先行研究 [12] では、日本語述語項関係アノテーションを対象に、現代日本語書き言葉均衡コーパス (BCCWJ) [6] から選択した 43 記事に対して、5 名の作業者がアノテーションを行う際の視線情報とアノテーションツールの操作履歴を収集した。視線情報の記録には図 1 に示すように視線計測装置 Tobii T60 を使用し、アノテーション履歴の記録には図 2 に示すようにアノテーションツール Slate [5]^{*1} を修正したものを使用した。述語項関係アノテーションでは、作業者は文章中出现する全ての述語に対して、ガ格、ヲ格、二格のアノテーションを行った。述語やその項となりうる名詞句はあらかじめセグメントとしてアノテーションされており、作業者はマウスを利用して述語とその項の関係をアノテーションした。図 2 は一つの文章全体のアノテーション作業が終了した時点のスナップショットである。図中のセグメント間をつなぐ線は作業者によってアノテーションされた述語項関係を表し、線の色は格関係を表している。

収集したデータを分析する際、我々は作業者が述語に対して項をアノテーションする過程を以下の 3 つの段階に分割して調査を行った [3]。

適応段階 アノテーション作業者が与えられた文章を読み、文脈を理解する。

評価段階 アノテーション対象となる述語を注視した後、述語との関係 (例えば、ガ格) に関して、項候補集合の中から正解となる項を探索する。

確認段階 アノテーション対象とすべき項を見つけた後、述語項関係を確認するために周辺文脈を探索し、述語と項の関係を確定する。

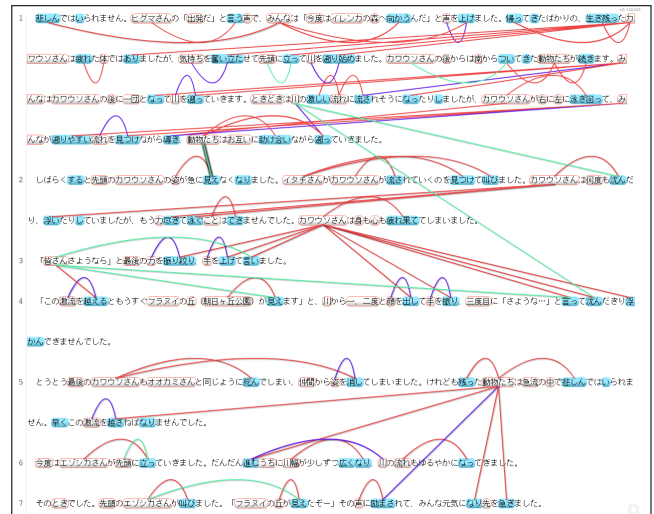


図2 アノテーションツールのスナップショット (作業終了時)

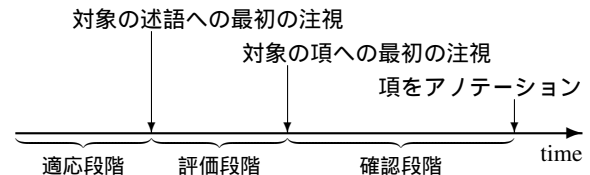


図3 アノテーション過程の分割

認知科学の分野では、認知過程をいくつかの機能的段階に分割するという考え方が一般的に用いられており、ここで示したアノテーション時の認知過程の分割には Gidlöf ら [1] や Russo ら [10] が議論している考え方が反映されている。図 3 はこの分割基準と各段階の関係をまとめた図である。

ここで示した 3 つの段階の中で、述語項関係解析に役立つ素性を検討する上で最も有益だと考えられるのは、作業者がアノテーション対象の述語に対して適切な項を同定する段階、すなわち対象の述語を注視してから述語と特定の格関係を持つ項を発見するまでの評価段階である。しかし、先行研究で採用したデータ収集の方法では作業者が文章全体に対して網羅的にアノテーションを行ったため、個別の述語に関して厳密に評価段階の範囲を決定することが困難となる問題が発生した。この問題を解決するために本研究では、先行研究のように一つの文章全体を対象としてデータ収集を行うのではなく、文章中のあらかじめ指定した一つの述語に対して対応する一つの項をアノテーションする課題を設定し、このときの作業者の振舞いを収集する。

3. 単一述語項関係アノテーション課題

3.1 アノテーション仕様

本研究で行うアノテーションの仕様には、NAIST テキストコーパス [13] の仕様を簡略化したものを採用した。ただし、述語項関係のうち、自動解析として興味深い問題は項が省略されるゼロ照応の関係であるため、項が述語の近く

^{*1} <https://bitbucket.org/dainkaplan/slate/>

から遠くまでさまざまな位置関係として出現するガ格のみを対象にアノテーションを行った。

アノテーションは述語の原形に対して行うため、述語が格交替を伴って出現している場合は、表層的なガ格に対して項をアノテーションしないよう注意する必要がある。省略された格要素（ゼロ代名詞）に対して述語項関係を付与する場合は、述語の前方文脈に出現する先行詞に対してアノテーションを行う。述語の項として適切な表現が複数個出現している場合は、いずれか一つにアノテーションを行う。また、例えば並列構造のように、名詞句の範囲が文節を越える場合は最右の文節にのみアノテーションを行う。

3.2 アノテーション対象データの作成

アノテーション対象の文章には、BCCWJ[6] のコアデータの書籍データ (PB) を採用した。NAIST テキストコーパスではアノテーション対象として毎日新聞の記事を利用しているが、新聞記事の場合、例えばリード文に出現する主題がその後省略されやすいというように、分析内容が新聞特有の現象に偏ることが考えられる。本研究では、作業者が項を探索する過程を分析の対象としており、より多様な現象が含まれる文章を対象にアノテーションを行うことが望ましいため、書籍がこの要件を満たすという仮定に基づき書籍データを選択した。

本研究で行うデータ収集では、収集した視線情報と画面上の文字の対応付けを簡単にするために、アノテーション時にテキストのスクロールが起らないようにした。したがって、画面上に表示されている項候補のセグメントがアノテーションの対象となる。画面上に表示する内容としては、本実験で利用する視線計測装置 Tobii T60 の計測誤差やアノテーション時の文章の可読性を考慮し、この長さを最大 962 文字 (74 文字 × 13 行) とした。したがって、まずこの長さに収まるようにコーパス中の文章を分割し、アノテーション候補文章とする。

次に、各候補文章中からアノテーションの対象とする述語項関係の一つを選択する。述語項関係のアノテーションでは述語よりも前方文脈に出現している項に対してアノテーションを行うため、項が述語の前方に出現しているものを選択する。また、項候補の数が少なくなりすぎないように、述語が文章の後半に出現している述語のみから選択を行う。候補文章から適切な述語項関係の一つを選択できた場合は、その候補中からそれ以上アノテーションの対象を抽出することはしない。このようなアノテーション対象データ作成の方法を採用することで、作業者が常に新規の文章に対して作業することになり、一度作業した文章の内容を記憶していて、アノテーションを行うということがなくなる。

最後に、抽出された述語項関係の中からランダムに事例を選択する。このとき、述語とそのガ格の項が同一文内にある場合 (文内) と同一文内にない場合 (文間) の事例の割

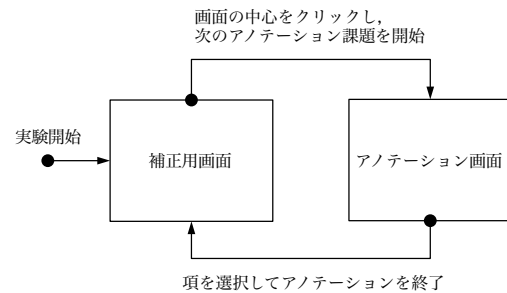


図5 アノテーションツールの画面遷移

合を1対1程度に調整することで、それぞれの場合におけるアノテーション作業者の振舞いを分析可能にする。このようにして抽出した221事例をアノテーション対象のデータとして利用する。アノテーションの際には、文章中の対象となる一つの述語とその述語の前方文脈に出現する項候補を明示しておき、作業者はアノテーション対象の述語のガ格に対応する適切な項を項候補の中から一つ選択することになる。

4. データ収集実験

4.1 アノテーションツールのインタフェース

単一述語項関係のアノテーションを行うツールには、データ収集の目的に合わせて作成した専用のアノテーションツールを利用した。このアノテーションツールのスナップショットを図4に示す。図4では、アノテーション対象の述語のセグメントは薄青色の長方形で、述語に対応する項候補のセグメントは赤線で囲まれた灰色の長方形で表示されている。アノテーション時の視線情報の分析を容易にするために、画面内には文章とセグメントのみが表示される。アノテーションの際は、作業者は項候補のセグメントの中から述語のガ格の項として適切なセグメントをマウスでクリックすることで選択する。これにより作業者がキーボードなどの手元を見て作業を行い、その結果視線情報が連続的に収集できなくなるという問題を回避した。作業者がセグメントを選択すると自動的に画面が切り替わり、視線位置を補正するための画面が表示される。この画面の中心のマーカーをクリックすることによって次のアノテーション課題へ進む。図5はこの画面遷移をまとめたものであり、作業者は補正用画面とアノテーション画面を交互に遷移しながらアノテーション作業を進める。

4.2 アノテーション作業へのインストラクション

実際のデータ収集を行う前に、実験で採用したアノテーションの仕様を作業者に把握させるために、以下に示す通りの順序でアノテーションのインストラクションを行った。

(1) 実験概要の説明

作業者が取り組む単一述語項関係アノテーション課題の内容について簡単な説明を行う。また、アノテ

の姿であったように思えてくる。

兄の正樹が、母と佐倉は昔からの知り合いだったに相違ないと、いつか興奮気味に自分に告げてきたことまでが思い出されていたからかも知れない。正樹に言わせれば、いくらなんでも母が見も知らぬさいはての地の、それも小さな病院の外科医を頼って行くはずはない、きっとどこかで佐倉のことを知っていたはずだ、と。どこかと言っても、お母さんは大学病院を出た後は名取の町立病院にしか勤めていないのだからどちらかでしょ、と言い返すと、ウン、どちらかだよ、でも、多分町立病院だ、と兄は主張した。その根拠はと尋ねると、大学病院はあまたの医者がいただろうし、母も看護婦になりたてで右も左もわからない状況だったろうから、特定のひとりの医者だけが印象に残ったとは考え難い、というものだった。では母はなぜ佐倉とは一面識もなかったと嘘をついたのか、との疑問には、それはたぶん父を氣遣ってのことだろう、と答えた。

昔職場で一緒だった医者がその道の大家となって、たとえば仙台の大学病院や東京の大病院の偉い外科医になっている、ということなら、その医者が旧知の人であったとしてもなんら父も不思議に思わないだろうが、秋田のさびれた町の小さな病院に勤めている外科医をいくら旧知でもわざわざ訪ねていくはずはない。何か男女の秘密めいた関係が二人の間にあったのではないかと父は疑うだろう、だから知らないと言い張ったんだよ、と。兄はもう、母と佐倉は名取の町立病院で出会っている、と固く信じ込んでいるようだ。だが三宝は、半信半疑のまま兄の言い分を聞いていた。

朝食の最中、正樹から電話が入った。改めて志津の今日の予定を問いただしてきた。

☐ 正解の項

「お昼ごはん、どこかで一緒に食べようか？」

図4 単一述語関係アノテーションツールのスナップショット

ション作業中に視線情報を記録すること、記録した情報を利用して作業者がアノテーションを行う過程を分析することについて説明を行う。

(2) アノテーション作業に関する注意事項の説明

作業方針を指示しないと、作業者が文章頭からアノテーション対象の述語まで読み進め、その後で正解の項を探索する可能性がある。しかし、任意の文章に対して作業者がこの方針を取ると、局所的な探索のみで問題が解決する場合も広く文脈を見渡さなければならぬ場合も区別なく同じ作業が行われてしまい、収集した視線にその違いが反映されないという問題が起こる。これを防ぐために、アノテーションの際は画面内に表示された全てのテキストを読む必要は無く、正解となる項を発見し、十分な確信度が持てる場合にはそのまま残りの文脈を見ることなくアノテーションするよう指示した。これは、ガ格の項を発見し妥当性が確認できた場合は即座にアノテーションを完了し次の問題へと移るようにすることで、正解をアノテーション

する際に起こる冗長な確認作業や直接対象述語の項の選択に無関係な読みに関する記録を排除するためである。

(3) アノテーションの練習セッション

必要に応じてアノテーション仕様の説明をしながら、作業者に5問の練習課題をアノテーションさせ、ツールの使い方を習得させる。具体的な注意事項として以下を伝えた。

- 画面内の一つの述語の原形に対して、ガ格となる項を事前にアノテーションされた項候補の中から一つ選ぶこと、
- その項は文内に出現している場合もあれば文を越えて出現する場合もあるということ
- 共参照の関係にある複数の項候補が正解となる場合には、その中のどれを選んでもよいこと
- 並列構造を持った名詞句全体にアノテーションを行う場合は最右のセグメントを選択すること
- 述語が受動化や使役化によって格交替を起している

場合は述語の原形を基礎としてガ格の項を選択すること

(4) 確認問題を用いた仕様の理解度の確認

練習課題で説明したアノテーション仕様が理解できているかを確認するために、別の5課題をアノテーションさせ、その結果によって仕様の理解度を確認する。この際間違った作業を行った場合はその作業を行った時点もしくは全5問の確認課題の作業完了時に間違いの内容を説明し、仕様の理解度を向上させた後でデータ収集を開始する。

4.3 実験設定

アノテーション時の作業者の視線の記録には視線計測装置 Tobii T60 を利用した。Tobii T60 は解像度 1,280×1,024 のディスプレイを持ち、作業者が画面上のどの位置を見ていたかを 1/60 秒で記録することができる。アノテーション作業時にはディスプレイと作業者の間の距離が約 50cm になるよう調整した。また、視線検出のエラーを抑制するために図 1 のように顎台を利用することで、作業者の頭の動きを固定した状態でデータ収集を行った。

アノテーション作業には計 20 名の作業者 (以降、各作業者を a t と表記する) を雇用し、各作業者は 3 節で作成した 221 文章に対してアノテーション作業を行った。このうち、5 名 (c, k, n, o, t) の作業者は過去に述語項関係アノテーションの経験がある。作業の際は、アノテーション対象の文章を 5 セットに分け、1 セットをアノテーションするごとに休憩を入れるようにし、各セットの作業開始時には視線計測装置のキャリブレーションを行った。作業者の 221 事例の作業時間の平均は約 2 時間であった。

5. 収集したデータの内訳

5.1 アノテーション結果

20 名のアノテーション作業 (a t) が全 221 文章に対してアノテーションした結果の作業者ごとの正解率を表 1 に示す。表 1 より、各作業者の正解率は約 6 割から 9 割と作業者ごとに大きく異なる一方、述語項関係アノテーションの経験者は全員 8 割以上の高い正解率になっていることがわかる。これは、これらの作業者の過去の経験が直接アノテーション結果に反映されたものと考えられる。

このアノテーション結果を調査するために、各問題を言語的な特徴から分類し、分類したカテゴリごとの作業結果の正解率を計算した結果を表 2 に示す。この表から、表 1 の平均正解率と比較して、特徴的なカテゴリに分類された問題の正解率が低く、またその他問題では正解率が高いことがわかる。これはインストラクション時にアノテーションに影響を及ぼす典型的な言語現象とその対応を示したにも関わらず正解率が低いことを表しており、個別の言語現

表 1 各作業者のアノテーション作業結果の正解率

作業者	正解率	作業者	正解率
a	0.63 (140/221)	k*	0.82 (182/221)
b	0.81 (178/221)	l	0.61 (134/221)
c*	0.85 (188/221)	m	0.59 (130/221)
d	0.64 (142/221)	n*	0.91 (201/221)
e	0.78 (173/221)	o*	0.82 (182/221)
f	0.85 (188/221)	p	0.76 (168/221)
g	0.69 (152/221)	q	0.63 (139/221)
h	0.68 (151/221)	r	0.67 (147/221)
i	0.74 (163/221)	s	0.78 (173/221)
j	0.51 (112/221)	t*	0.83 (183/221)
		平均	0.73 (161/221)

*は述語項関係アノテーション経験者。

表 2 カテゴリごとの問題数と正解率

問題カテゴリ	問題数	正解率	問題の特徴
格交替 (受動)	12	0.35	述語: れる・られる
格交替 (使役)	3	0.06	述語: 使役動詞, せる・させる
可能	8	0.24	述語: 可能動詞, できる
慣用表現	7	0.24	述語が慣用表現の一部
換喩	9	0.33	項に換喩の関係が成立
全体-部分関係	7	0.28	項に全体-部分関係が成立
ガ-ガ構文	5	0.16	述語がガ格を 2 つ持つ
形式名詞	5	0.11	項が「こと」や「ところ」など
複数の登場人物	17	0.47	項候補として複数の人物が出現
知識+推論	2	0.01	項の選択に知識と推論が必要
曖昧性	2	0.03	項に曖昧性が存在
その他	151	0.79	上記以外

複数のカテゴリに含まれる事例も存在する。

象を理解させるためにより丁寧なインストラクションが必要であることがわかる。

5.2 視線データ

図 6 に Tobii T60 が出力する視線計測結果の例を示す。それぞれの行は 1/60 秒ごとに計測された視線の計測位置を表しており、タイムスタンプ、視線の画面上 X 座標、視線の画面上 Y 座標の 3 つ組で構成されている。X 座標と Y 座標は左右の視線の平均座標を表しており、画面領域外の座標を示していた場合は、その時刻で視線計測に失敗したということを示している。例えば図 6 では、時刻が 14026 から 14076 までの約 50 ミリ秒の間、視線計測に失敗したことを表している。計測失敗が起きる原因としては、作業者が目を閉じていること、画面外を見ていること、視線計測装置と作業者の相性が悪いことなどが挙げられる。

表 3 に、文章ごとに計算した各作業者の視線計測の平均失敗率を示す。表 3 中の平均と最大の列は文章ごとに計算した視線計測に失敗したデータ数の割合の平均値とその最大値を表している。この結果に基づき、経験的に計測失敗率の平均値が 0.15 以上かつ最大値が 0.70 以上という条件

\$TIMESTAMP	\$GAZE_X	\$GAZE_Y
13942	165	643
13959	168	646
13975	175	643
13992	159	639
14009	162	632
14026	-1	-1
14042	-1	-1
14059	-1	-1
14076	-1	-1
14092	166	647
14109	170	645
⋮	⋮	⋮

図 6 Tobii T60 が出力する視線計測データ

表 3 各作業者の視線計測失敗率 (文章ごと)

作業者	平均	最大	作業者	平均	最大
a	0.10	0.34	k	0.09	0.23
b	0.04	0.18	l	0.07	0.18
c	0.04	0.11	m	0.06	0.70
d	0.08	0.21	n	0.03	0.10
e	0.12	0.36	o	0.19	1.00
f	0.16	0.72	p	0.08	0.20
g	0.12	0.26	q	0.05	0.15
h	0.09	0.40	r	0.23	0.59
i	0.15	0.47	s	0.08	0.28
j	0.28	0.79	t	0.13	0.42
平均			0.11	0.38	

を設定し、この条件に該当する 3 名の作業者 (f, o, j) を信頼性の観点から以降の分析の対象外とした。

また、先行研究 [12] で収集した視線データを分析した結果から、Tobii T60 では画面鉛直方向に大きく視線計測誤差が生じることがわかっている。そのため、視線データが文章中のどの行に対応しているかという鉛直方向の誤差修正を手で行った。このとき、視線の計測誤差が大きい作業者を同様に信頼性の観点から分析の対象外とし、この条件を誤差が行の高さ以上と設定した結果、2 名 (r, s) が除外された。したがって、以降の予備分析では 15 名の作業者 (a, b, c, d, e, g, h, i, k, l, m, n, p, q, t) のデータを対象に分析を行った。

6. 予備分析

文間の事例は文内とは異なり、アノテーションを開始した後、適切と思われる項を発見しても、その後周辺文脈の内容を把握して項の妥当性を確認する必要があるため、文内の場合とは異なる振舞いが観測できると考えられる。そこで予備分析では、これら文間と文内の事例について作業者の振舞いの違いを収集データから捉えることを試みた。

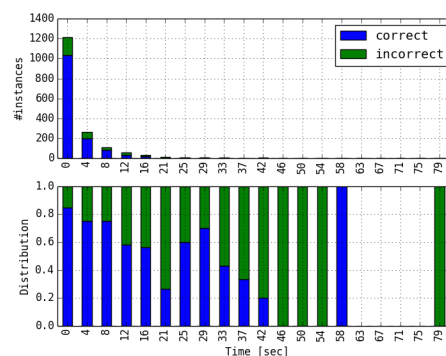


図 7 作業時間と正解率の関係 (文内事例, 評価段階)

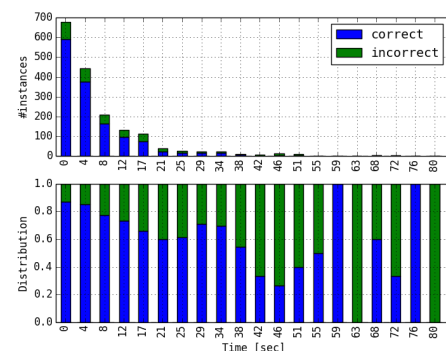


図 8 作業時間と正解率の関係 (文内事例, 確認段階)

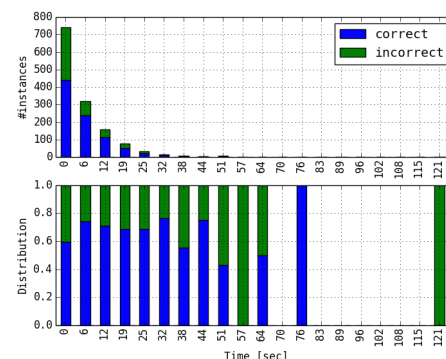


図 9 作業時間と正解率の関係 (文間事例, 評価段階)

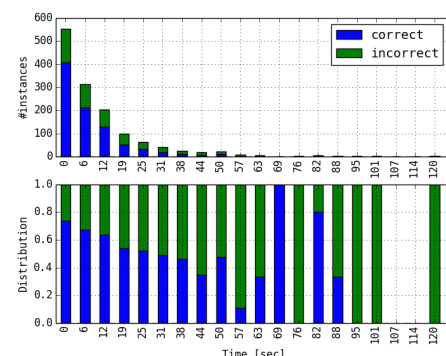


図 10 作業時間と正解率の関係 (文間事例, 確認段階)

6.1 評価段階と確認段階の分析

図 3 に示したアノテーション過程の分割では、作業者がアノテーション対象の項を最初に注視した時刻を評価段階と確認段階の境界とした。以下の分析でもこの規準を採用

表 4 視線ボタンを表す記号の定義

カテゴリ	記号
注視セグメント	対象述語 (p), 選択した項 (sel) 選択しなかった項の候補 (a),
注視の遷移	順方向 (+), 順方向サッケード (+j), 逆方向 (-), 逆方向サッケード (-j)

する．図 7-図 10 に作業者が評価段階と確認段階に費やした作業時間と正解率の関係を示す．図 7, 図 8 が文内の事例, 図 9, 図 10 が文間の事例に対する結果である．各図の上部のグラフは, 作業時間 (x 軸) と正解 (青)/不正解 (緑) 事例の頻度の関係 (y 軸) を表している．また, 下部のグラフは, 作業時間 (x 軸) と, その区間での正解/不正解事例の頻度の比率 (y 軸) を示している．図 7-図 10 から, 文間の事例では, 確認段階の作業時間が長くなるにつれて, 正解率が減少することがわかる．

6.2 頻出する視線ボタンの分析

我々は先行研究 [8] において, 文献 [12] で収集したデータを利用して, アノータ作業者間の作業結果の不一致を視線データに基づいて検出する問題に取り組んだ．具体的には, アノテーション中の作業者の注視の系列を, 注視が起きたセグメントの種類や述語からの相対位置によって記号列に変換し, 頻出する記号ボタンを素性として利用し, 不一致検出モデルを構築した．このモデルを評価した結果, 頻出する一部の注視ボタンが検出に有効であることがわかったが, この記号化は視線の移動に関する粒度が細か過ぎたために, マクロな視線の移動をうまく捉えきれない可能性が残されていた．本研究では, より抽象度の高い作業者の振舞いを捉えるために, 注視の系列を記号化する新しい手法を提案し, これを用いて分析をおこなう．

アノテーション時の特徴的な視線の動きを捉えるために, 先行研究同様にアノテーション中の注視^{*2}の遷移を記号化する．先行研究の方法では, 注視が起きたセグメントのみを記号化したが, 本研究ではそれに加えて遷移の種類も記号化する．表 4 に注視と注視間の遷移の記号化の一覧を示す．表 4 にしたがって, 注視は注視したセグメントの種類を表わす記号に, 注視間の遷移はどの方向にどの程度移動したかを表わす記号に変換する．遷移は文章を読み進める方向に起きた場合は+, 読み返す方向に起きた場合は-となる．遷移の移動距離が一定距離^{*3}を越えた場合はサッケード^{*4}として扱い, 通常の文章を読み進む視線の動きと区別して扱う．

注視系列を記号の系列に変換したのち, 文内と文間の事

^{*2} 視線データから注視を抽出する際には, Dispersion-Threshold Identification (I-DT) アルゴリズム [11] を利用した．

^{*3} 日本語文章における周辺視野 [4], [9] を考慮して, この距離閾値を 7 文字とした．

^{*4} 一点から別の点に素早く移動する視線の動き

例にそれぞれ頻出するボタンを調べる．ここでは, 15 名の作業者が文内と文間の事例に対してアノテーションを行った結果を正解と不正解の 2 つに分け, 正解と不正解の事例に偏って頻出するボタンが存在するかどうかをそれぞれの事例について調べた．出現の偏りはボタンの頻度と出現確率を考慮し, 式 (1) に基づいて計算したスコアを昇順に並べることで調べた．

$$S(p_i) = \log(C(p_i)) \frac{C(p_i)}{N^+} - \log(C(p_i)) \frac{C(p_i)}{N^-} \quad (1)$$

ここで, $C(p_i)$ は視線ボタン p_i の頻度, N^+ と N^- はそれぞれ正解と不正解の事例に出現した視線ボタンの総数である．視線ボタンは, 遷移元セグメント, 遷移記号のリスト, 遷移先セグメントの 3 つ組を利用した．視線ボタンの頻度とスコアをスコアの降順にソートし, その上位, 下位 20 位ずつを表 5 に示す．表 5 の中で上位に出現しているボタンはアノテーションの結果が正解の事例に偏って出現していることを表しており, 下位に出現しているボタンは不正解の事例に偏って出現していることを表す．表 5 では, 最終的に選択された項以外の項候補間の遷移を表わす視線ボタン (a/.../a) を赤色で示した．文間の事例では正解の側にこのボタンが頻出している一方, 文内の事例では不正解の側に頻出していることがわかる．これは, 文間の事例では項の候補を見渡して慎重にアノテーションする方が正解になる割合が高い一方, 文内の場合は余計な項の候補を見ずに決めている, つまり迷っていない場合に正解率が高くなるということを示唆している．

7. おわりに

本研究では, 日本語述語項関係アノテーション時の作業者の振舞いを収集し, それを分析することで, これまでに得られていない述語項構造解析のための知見を得るという問題に取り組んだ．先行研究では文章中すべての述語項関係をアノテーションしたときの振舞いを収集したのに対し, 本研究では, 文章中の一つの述語に対してそのガ格の項をアノテーションする課題を設定し, 20 名の作業者に 221 文章のアノテーション作業を行わせた．アノテーション中の作業者の視線とアノテーションツールの操作履歴を収集し, そのデータを分析することによって, 項が述語と同一文内にある場合 (文内事例) とない場合 (文間事例) で頻出する視線のボタンが異なることを明らかにした．今回は, 抽出した視線ボタンの頻度のみを利用して分布の違いを分析したが, テキストマイニング技術やクラスタリング技術を利用することでより特徴的なボタンを抽出することが必要である．また, 今回利用した記号化の方法では単純な情報をヒューリスティクスに基づいて導入したが, 今後はセグメントの意味カテゴリの情報や遷移の読み返しなどの情報を利用することで, より抽象度の高い観点からの分析を行うことを考えている．また, 今回はアノテーション対象の文

章を文間事例と文内事例で分けて分析を行ったが、個々の問題カテゴリに応じた細かい分析をすることで、全問題を対象にした分析では捉えられない振舞いを明らかにすることも必要だと考えている。今後は、こうした分析から特徴的な振舞いを明らかにすることで、述語項構造解析、特にゼロ照応解析に役立つ知見を明らかにしたい。

参考文献

- [1] Gidlöf, K., Wallin, A., Dewhurst, R. and Holmqvist, K.: Using eye tracking to trace a cognitive process: Gaze behaviour during decision making in a natural environment, *Journal of Eye Movement Research*, Vol. 6, No. 1, pp. 1–14 (2013).
- [2] Grosz, B. J., Joshi, A. K. and Weinstein, S.: Centering: A Framework for Modeling the Local Coherence of Discourse, *Computational Linguistics*, Vol. 21, No. 2, pp. 203–225 (1995).
- [3] Iida, R., Mitsuda, K. and Tokunaga, T.: Investigation of annotator's behaviour using eye-tracking data, *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*, pp. 214–222 (online), available from <http://www.aclweb.org/anthology/W13-2326> (2013).
- [4] Ikeda, M. and Saida, S.: Span of recognition in reading, *Vision Research*, Vol. 18, No. 1, pp. 83–88 (online), DOI: 10.1016/0042-6989(78)90080-9 (1978).
- [5] Kaplan, D., Iida, R., Nishina, K. and Tokunaga, T.: Slate – A tool for creating and maintaining annotated corpora, *Journal for Language Technology and Computational Linguistics*, Vol. 26, No. 2, pp. 89–101 (2012).
- [6] Maekawa, K., Yamazaki, M., Maruyama, T., Yamaguchi, M., Ogura, H., Kashino, W., Ogiso, T., Koiso, H. and Den, Y.: Design, compilation, and preliminary analyses of balanced corpus of contemporary written Japanese, *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2010)*, pp. 1483–1486 (2010).
- [7] Miller, G. A.: WordNet: A Lexical Database for English, *Communications of the ACM*, Vol. 38, pp. 39–41 (1995).
- [8] Mitsuda, K., Iida, R. and Tokunaga, T.: Detecting missing annotation disagreement using eye gaze information, *Proceedings of the 11th Workshop on Asian Language Resources*, pp. 19–26 (2013).
- [9] Osaka, N.: Size of saccade and fixation duration of eye movements during reading: Psychophysics of Japanese text processing, *Journal of Optical Society of America*, Vol. 9, No. 1, pp. 5–13 (1992).
- [10] Russo, J. E. and Leclerc, F.: An eye-fixation analysis of choice processes for consumer nondurables, *Journal of Consumer Research*, Vol. 21, No. 2, pp. 274–290 (1994).
- [11] Salvucci, D. D. and Goldberg, J. H.: Identifying fixations and saccades in eye-tracking protocols, *Proceedings of the 2000 symposium on Eye tracking research & applications (ETRA '00)*, pp. 71–78 (online), DOI: 10.1145/355017.355028 (2000).
- [12] Tokunaga, T., Iida, R. and Mitsuda, K.: Annotation for annotation – Toward eliciting implicit linguistic knowledge through annotation –, *Proceedings of the 9th Joint ISO - ACL SIGSEM Workshop on Interoperable Semantic Annotation (ISA-9)*, pp. 79–83 (2013).
- [13] 飯田 龍, 小町 守, 井之上直也, 乾健太郎, 松本裕治: 述語項構造と照応関係のアノテーション: NAIST テキストコーパス構築の経験から, 自然言語処理, Vol. 17, No. 2, pp. 25–50 (2010).

表 5 視線バタンの出現分布

文内事例						文間事例					
視線ボタン p_i	正解		不正解		$S(p_i)$	視線ボタン p_i	正解		不正解		$S(p_i)$
	頻度	相対頻度	頻度	相対頻度			頻度	相対頻度	頻度	相対頻度	
sel/+/a	1095	0.0407	219	0.0173	0.0834	a/+/a	5546	0.1790	3662	0.1700	0.0645
a/-/sel	693	0.0258	115	0.0091	0.0545	a/-j/sel	595	0.0192	223	0.0104	0.0290
p/-j/sel	628	0.0234	117	0.0092	0.0463	sel/+/a	826	0.0267	411	0.0191	0.0279
a+/p	1349	0.0502	533	0.0420	0.0425	a/-/a	2058	0.0664	1350	0.0627	0.0240
a/-j/sel	531	0.0198	177	0.0140	0.0224	p/-/a	891	0.0288	507	0.0235	0.0212
sel+/p	538	0.0200	185	0.0146	0.0216	a/-j/a	3144	0.1015	2164	0.1005	0.0199
sel/+j/p	379	0.0141	99	0.0078	0.0208	a/-/sel	422	0.0136	183	0.0085	0.0165
p/-/a	854	0.0318	368	0.0290	0.0187	a+/p	1176	0.0380	751	0.0349	0.0163
a/-/a	1535	0.0571	728	0.0574	0.0176	a/+j/a	3771	0.1217	2651	0.1231	0.0141
p/-/sel	408	0.0152	154	0.0121	0.0130	sel/+j/a	374	0.0121	182	0.0084	0.0120
a/+j/sel	495	0.0184	214	0.0169	0.0103	a/+,+j/a	450	0.0145	281	0.0130	0.0066
sel/-j/a	413	0.0154	179	0.0141	0.0084	sel/+j/p	274	0.0088	155	0.0072	0.0058
sel/+j/a	394	0.0147	169	0.0133	0.0083	a/+, -j/a	355	0.0115	220	0.0102	0.0053
sel/-,+/sel	134	0.0050	27	0.0021	0.0075	a/+j, -j/a	434	0.0140	295	0.0137	0.0031
p/-j/a	876	0.0326	427	0.0337	0.0073	a/-j, -/a	197	0.0064	125	0.0058	0.0024
p/+, -/p	209	0.0078	73	0.0058	0.0073	a/-, -j/a	168	0.0054	106	0.0049	0.0021
a/+j/p	752	0.0280	364	0.0287	0.0069	a/+j, +, -j/a	87	0.0028	46	0.0021	0.0019
p/+, -j/sel	90	0.0033	9	0.0007	0.0059	a/+j, +, +j/a	56	0.0018	25	0.0012	0.0015
a/+j, +/p	116	0.0043	37	0.0029	0.0043	sel/+j, +/a	51	0.0016	22	0.0010	0.0014
a/+, -j/sel	57	0.0021	12	0.0009	0.0027	a/+j, +/p	142	0.0046	93	0.0043	0.0014
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
p/-j, -, +/a	3	0.0001	10	0.0008	-0.0008	p/-j, -/a	57	0.0018	51	0.0024	-0.0008
a/+, +j, -/a	8	0.0003	12	0.0009	-0.0008	p/-j, +j/a	40	0.0013	39	0.0018	-0.0008
sel/+j, +/a	17	0.0006	18	0.0014	-0.0010	sel/+, -/p	0	0.0000	17	0.0008	-0.0010
sel/-j, -/a	10	0.0004	15	0.0012	-0.0010	p/-j, +/sel	16	0.0005	27	0.0013	-0.0012
p/+j, -j/a	23	0.0009	21	0.0017	-0.0010	p/+, -/sel	0	0.0000	20	0.0009	-0.0012
a/+j, -j, +/a	18	0.0007	19	0.0015	-0.0011	a/+, -, +/a	33	0.0011	39	0.0018	-0.0013
a/-j, -/a	90	0.0033	58	0.0046	-0.0015	sel/-, +/sel	52	0.0017	54	0.0025	-0.0015
a/+j, -/a	81	0.0030	54	0.0043	-0.0016	a/-, +/a	377	0.0122	290	0.0135	-0.0018
a/+, +j/p	51	0.0019	39	0.0031	-0.0017	sel/+, -/sel	59	0.0019	65	0.0030	-0.0021
p/-j, +/a	62	0.0023	45	0.0035	-0.0018	a/-j, +/a	206	0.0067	170	0.0079	-0.0022
a/+, +j, +/a	41	0.0015	36	0.0028	-0.0020	p/-j/a	851	0.0275	642	0.0298	-0.0032
a/+, -/a	200	0.0074	118	0.0093	-0.0022	a/+j/sel	366	0.0118	293	0.0136	-0.0033
a/-, -j/a	110	0.0041	73	0.0058	-0.0024	p/+, -/p	167	0.0054	160	0.0074	-0.0044
a/+, -j/a	200	0.0074	123	0.0097	-0.0032	sel/-j/a	288	0.0093	247	0.0115	-0.0046
a/+, +j/a	212	0.0079	140	0.0110	-0.0054	a/+, -/a	254	0.0082	225	0.0104	-0.0049
a/+/a	4334	0.1613	2242	0.1768	-0.0060	a/+j/p	801	0.0259	626	0.0291	-0.0062
a/+j, +/a	307	0.0114	192	0.0151	-0.0062	sel/-/a	203	0.0066	256	0.0119	-0.0135
a/+j, -j/a	195	0.0073	147	0.0116	-0.0085	p/-/sel	2	0.0001	139	0.0065	-0.0138
a/-j/a	1882	0.0700	1191	0.0939	-0.0596	a/+sel	474	0.0153	473	0.0220	-0.0178
a/+j/a	2153	0.0801	1336	0.1054	-0.0624	sel/+p	1	0.0000	182	0.0084	-0.0191

正誤表

7 ページ右カラム

誤：

$$S(p_i) = \log(C(p_i)) \frac{C(p_i)}{N^+} - \log(C(p_i)) \frac{C(p_i)}{N^-} \quad (1)$$

ここで， $C(p_i)$ 視線パターン p_i の頻度， N^+ と N^- はそれぞれ正解と不正解の事例に出現した視線パタンの総数である．

正：

$$S(p_i) = \log(C^+(p_i)) \frac{C^+(p_i)}{N^+} - \log(C^-(p_i)) \frac{C^-(p_i)}{N^-} \quad (1)$$

ここで， $C^+(p_i)$ と $C^-(p_i)$ はそれぞれ，正解と不正解の事例に出現した視線パターン p_i の頻度， N^+ と N^- はそれぞれ正解と不正解の事例に出現した視線パタンの総数である．