

# 談話レベルの推敲支援のための人手修正基準

飯田 龍                      徳永 健伸  
東京工業大学 大学院情報理工学研究科  
{ryu-i,take}@cl.cs.titech.ac.jp

## 1 はじめに

近年、文章作成・推敲支援などの教育的応用を指向した自然言語処理の研究が注目を集めている。これらの研究課題は機械学習アルゴリズムの発展により、語や語の並びなどの表層的な手がかりを用いて小論文の自動評価や助詞など単語レベルの誤用検出などの研究が進められている [2]。これらの技術の進歩に対し、大学(院)生の卒業論文・修士論文などを実際に推敲する立場で分析してみると、書くべき内容は充足しているものの、文と文のつながりや、文内の構成素の配置が適切ではないため、結果的に可読性の低い文章となっていることが多い。これは日本では中学・高校のカリキュラムに文章作成の授業が無いこと、また特に理工学系の学生は一般的に大学入試に小論文が含まれないために、文章作成に関する指導を受ける機会が少ないことに起因する。また、留学生が作成した日本語の文章を見ると、主語が過剰に省略されるなど、参照表現生成における問題も存在する。しかし、既存の文章作成・推敲支援のシステムでは前述の通り表層的な手がかりを主に利用しているため、これらの問題を改善することは難しい。このような背景から、本研究では、自然言語処理技術、特に談話処理技術の応用として、人間が記述した小論文などの文章を談話的な観点から推敲する技術の開発を目指す。

談話レベルの自動推敲を実現するために、自然言語処理が取り組む方向性として以下の二通りが考えられる。

1. 言語の理論的な立場、特に談話に関する理論的な見地から、首尾一貫性のある文章を生成するための理論を構築する。
  2. 実際に人手で推敲された結果を分析し、経験的に自動推敲に必要な知見を得て、それをモデル化する。
- 前者については、言語生成の研究が特に関連する研究分野となり、参照表現の生成、談話関係解析、首尾一貫性の推定など、細分化された研究分野で得られた知見をどのように実際の推敲・校正の問題に適合するかを考える問題となる。我々はこのような談話的な特徴のうち、参照表現の生成、首尾一貫性の推定、節の受動化についてどのような観点で文章を生成すればよいかの調査を続けている [12, 4, 5]。

一方、後者の経験的なアプローチとして推敲の問題を考える場合、推敲前のデータ(例えば、著者が実際に作成した文章)とそれを専門家が修正した推敲後のデータの両方が必要となる。このため、いかにして推敲前のデータを収集するか、また収集したデータをどのように修正して推敲後のデータを作成するか問題となる。推敲前のデータ収集については、単純に品質の低い文章(日本語学習者が書いた不慣れな文章など)を収集した場合、

文法的な誤りなど、談話レベルの問題以外に修正する項目が多く含まれるため、談話的な特徴を分析する前に人手による文法的な修正が必要となり、原文を対象とした推敲の処理が想定しづらいという問題が起こる。このため、文法的には妥当だが、談話的な観点から推敲することで改善が可能な文章を扱う必要がある。また、作業者に談話レベルの修正を行わせるために、どのような指示をすべきかについても検討する必要がある。これは、単純に専門家に推敲を依頼した場合、助詞や細かい言い回しなど局所的な修正に留まる傾向があり、本研究で対象としたい談話レベルの推敲のためのデータを収集できないという問題が起こるためである。

このような背景から、本研究では談話レベルの推敲のためのデータ作成について考える。まず、2節で談話レベルの修正を作業者に指示する方針について考察し、3節ではその方針にしたがってデータ作成を行った結果と、作成されたデータの特徴について述べる。次に、4節で作成したデータの性質を見るために、既存の首尾一貫性を推定するモデルを用い、推敲前と推敲後の文章の首尾一貫性を見積り、首尾一貫性モデルが適切に推敲後の文章を選択できるかを調査する。最後に5節でまとめと今後の課題について述べる。

## 2 データ作成のための人手による文章修正

前節で指摘したように、校正の専門家に何の制約もなく小論文の修正を依頼すると、助詞や内容語の選択など局所的な修正に終わる場合が少なくない。これは、小論文の筆者の主張の流れを変更しないように書き換えるための技巧だと言えるが、これでは談話の構造まで考えた修正が行われているとは言えない。一方で、大幅に編集される場合は、元の文章を記述した筆者が想定していない内容レベルの追記・修正が行われることがあるが、それは内容レベルの妥当性を吟味する必要があるため、言語処理として修正を行うには敷居が高い問題となる。

このような理由により、作業者が行う作業を談話単位の集合に対するラベル付けとその再配置、またそれにとりも参照表現・接続表現などの局所的な修正に限定することで、内容レベルの書き換えをできるだけ行わないように制約をかけた。具体的には、1)の談話単位分割の前処理を施した後、作業者は2)の言説構成素のまとめあげと3)の局所的な修正を行う。さらに、4)に示すように、作業者は構成の結果に基づいて元の文章を採点する。

**1) 談話単位への分割**：まず、文章を構成する最小の単位となる談話単位に文章を分割する。文を節相当に分割するとともに、主題化された談話要素(「は」をともなう名詞句)を非主題化する。例えば、係助詞「は」「が」「を」「 $\phi$ 」に置き換えることで文中でどの談話要素が主

表 1: 採点項目

1.	適切に表現を省略できているか・省略しすぎていないか
2.	適切に参照表現 (e.g. この, その, これ, それ) を使っているか
3.	適切に接続表現を利用できているか
4.	談話単位の順序は適切か (適切な位置で主張が述べられているか, など)
5.	不要な主題化を行っていないか (e.g. ~について, ~に関して, などが冗長に書かれていないか)
6.	段落の構成は適切か
7.	1 文内に過剰に内容を詰め込みすぎているか (一文一義である必要はないが, 元の小論文に対し文の構成を変更すべきであると判断した場合は点数を下げる)

題化されていたかという情報を削除する。さらに, 分割した談話単位の末尾を正規化する (終止形にする)。

**2) 言説構成素へのまとめあげと並び換え:** 作業者は単独に書かれた内容にしたがって (局所的に) 修正を行うのではなく, 修正のためのトリガーとして文章の構造を明示的にアノテーションし, その構造が典型的な論説の構造から逸脱していることを顕在化させることで, 談話的な観点から適切な談話単位の配置となるように並び換える。ここで, 著者の主な主張, 主張の主な理由, 結論など, 議論的で論理的な文章を構成する談話単位の集合を**言説構成素**と呼び, 作業者は 1) で分割された談話単位のうち, どの箇所がそれぞれの役割に該当するかを吟味し, あらかじめ定義された言説構成素のラベルを付与する。さらに, 各言説構成素が文章中のどの位置で述べられるべきかについても検討し, 必要に応じて言説構成素の順序を入れ替える。

**3) 参照表現・接続表現などの局所的な修正:** 言説構成素の並び換えを行うことで, 文章中の接続関係が不明瞭になったり, また参照表現が冗長になるなど, 結束性が低い文章となる場合がある。そこで, 修正後に結束性の高い文章となるように, 必要に応じて談話単位の系列を文として構成し, 接続表現の追加・削除を行う。さらに, 参照表現も必要に応じて (ゼロ) 代名詞・名詞句に修正する。修正前にどのような表現で記述されていたかの履歴も残すために, 修正時には「〈修正前 | 修正後〉」のような記法で編集を行う。

**4) 文章の校正に関する元の文章の採点:** さらに, 校正を行う作業者に談話的な特徴を意識して修正を行わせるために, 明示的に談話的な特徴について筆者が作成した元の文章を表 1 に示す採点項目で採点を行った結果についても作業内容に含める。採点の際, 各項目について 1~5 点のいずれかの点を付与する。この結果を用いて, 例えば, 作業者ごとに採点された値が低い事例のみを分析することが可能になる。

以降で, 特に 1) の談話単位分割と 2) の言説構成素へのまとめあげ・並び換えについて, それぞれ 2.1 と 2.2 について説明する。

### 2.1 談話単位分割

文を命題相当の単位である談話単位に分割する処理については, 最終的には 2 節で示した 1) から 3) の手順を自動化することを想定して, 機械的に分割する手法を考える。この処理は基本的には日本語の節に相当する単位に分割する処理と考えてよい。文節を単位に述語が含まれる文節の後で分割すれば達成できるように思われる。しかし, 節が連体修飾の関係にあり名詞句を修飾す

表 2: 談話単位分割の実験結果

	再現率	精度	F 値
ベースライン	0.817	0.191	0.309
提案モデル	0.729	0.807	0.766

る場合, 「(～する) 可能性がある」といった英語では助動詞で表現される表現が名詞句となるため見かけ上連体修飾の関係にある場合, 「(～する) 前に」のような接続表現も名詞をとともうため連体修飾とみなされる場合など, 多様な表現が含まれるため, ヒューリスティックに命題として扱いたい単位に分割することが難しい。談話単位を構成する単位を人手でアノテーションし, その結果を用いて談話単位分割器を構築することが考えられるが, この単位の末尾に出現する表現は文章のスタイルによって異なるため, あるドメインの文章のみにアノテーションを行っても, その結果得られる談話単位分割器は他のドメインで利用できなくなる可能性があるため, できるだけ多様な表現が含まれるコーパスを対象にアノテーションをすることが望ましい。そこで, 本研究では現代日本語書き言葉均衡コーパス<sup>1</sup>の各サブコーパス (書籍コーパスなど) から均質な分布になるよう文の集合を抽出し, 全体で 156 記事 (5,815 文) に談話単位末尾の情報をアノテーションした。

このデータを用い, 談話単位分割の分類器を構築する。学習には最大エントロピーモデル<sup>2</sup>を利用し, 素性には分類対象となる文節とその前後の文節の主辞の形態素の基本形・品詞, その文節に含まれる機能語の基本形・品詞, 分類対象の文節から文末までのパス中に出現する名詞の基本形・品詞などを利用した。この分類器の性能を調査するために 10 分割交差検定で評価を行った。比較対象としては, ベースラインモデルとしてすべての述語を含む文節を談話単位末とするモデルを採用した。実験結果を表 2 に示す。この結果より, 談話単位分割を自動で行った場合, F 値で 0.7 を越える高い性能で分割できることがわかる。そこで, この自動分割器を用い, 小論文のデータの談話単位分割を行った。

### 2.2 言説構成素のまとめあげ・並び換え

作業者が文章を談話的な観点で構造化する場合, 修辭構造理論 [8] のようにボトムアップに構造を組み上げるような観点で文章を考えるのか, Teufel [9] の研究に代表されるような議論のセグメンテーション (argumentative zoning; AZ) を考え, 文章中の談話的な構成素がそれぞれどのような内容を表しているのかセグメンテーションするという二通りの方向性が考えられる。本研究で対象とする小論文のような議論的で論理的な文章を考えた場合, 文章に含まれる言説構成素は著者の主たる主張, その理由, 想定される反論やその解決策, 結論など限定的であると考えられるため, 我々は後者の AZ のような談話的な構成素へのセグメンテーションに基づく文章の構造化を考える。具体的には, 表 3 に示す Lin ら [7] が提案した構成素の定義を用い, 文章中の言説構成素 (談話単位の集合) に対して表 3 中のラベルのうちいずれかをアノテーションする。ただし, これらのラベルを付与するだけでは, 言説構成素単位の対応関係が理解できない。例えば, ある言説構成素に筆者の主張の主な理由を表す

<sup>1</sup><http://www.tokuteicorpus.jp/>

<sup>2</sup><http://www.cs.utah.edu/~hal/megam/>

表 3: 言説構成素のラベル

ラベル	名前	説明
pro	prompt	著者に提示された問題を繰り返しているだけの内容
tran	transition	次の話題に移るための言い回しのみを含む (e.g. 「この理由として 2 つの理由があげられる」など)
the	thesis	著者の与えられたに関する立場を表明する内容 (e.g. 「早期教育に賛成である」など)
main	main idea	thesis を支持する主な理由や議論を述べた内容
elab	elaboration	main idea に関する説明や考えなどを含む内容。ただし、具体例は除く
supp	support	main idea や elaboration に関連する具体例
conc	conclusion	小論文の要約や結論に関連する内容
reb	rebuttal	thesis や main idea に反対する意見や具体例などを含む
solu	solution	rebuttal で示された疑問・問題を解消する内容
sugg	suggestion	rebuttal の解決方法を示す内容
back	background	main idea などの背景に相当する記述

main タグがアノテーションされ、一方、他の言説構成素に具体例を表す supp を付与した場合、それらの間に対応関係 (つまり、supp の部分が main の具体例を表していること) が理解できない。そこで、作業者は必要に応じて関係のラベルに同一の添字を付与することでこの対応関係もアノテーションする。

このラベルのアノテーションに加え、小論文の構成として不適切な配置の場合には談話単位の位置を修正することで、作業者が想定する理想的な小論文の構成となるように再配置する。

### 3 人手修正作業によるデータ作成とその分析

1 節で述べたように、推敲の対象とする文章の選別自体も談話レベルの推敲・校正という問題を考える上で重要となるが、この点を吟味した結果、本研究では記述される内容が議論的で、かつ論理的に記述されることが期待される内容であり、かつ母語が日本語である人物が文法的な誤りを (ほとんど) 含まない文章として日本人の高校生があるテーマについて記述した小論文のデータを利用した。具体的には、宇佐美 [10] が収集した小論文のうち、「小学校の授業における英語の早期教育は必要であるか否かについて、あなたの意見とその根拠が明確になるように、400(800) 字以内で論述しなさい」という課題について高校生が作成した小論文のうち 120 記事に対し、2 節で示した校正のやり方で人手修正データを作成する。作業者は日本語教師や小論採点など何らかの形で日本語の指導に携わったことのある人材を雇用し、各 3 名ずつ 60 文章を修正し、合計 120 の文章について修正結果を得る予定である。人手修正作業は現在進行中であり、以降では作業が完了した 118 記事を対象に修正作業の特徴を分析した結果について説明する。

まず、言説構成素がどのように遷移するかを調べるために、アノテーションされた結果中の隣接する言説構成素のラベル対の頻度を調査した結果を表 4 にまとめる<sup>3</sup>。この表より、例えば、BOT→the の頻度が高いことから、文章の最初には著者の立場を述べる傾向にある、また主な理由を述べた後、その詳細や具体例を述べる傾向にある (main→elab or supp)、さらに最後は結論を述べて文

<sup>3</sup>この際、アノテーションされたラベルの添字 (例えば、main1 の 1 など) は捨象して頻度を求めた。

表 6: 実験結果: 首尾一貫性の評価

モデル	精度 (%)	参考結果
ベースライン	50.0	50.0
entity grid モデル [1]	25.7	70.7
共参照解析を用いたモデル [4]	38.6	76.1

章を終わる (conc→EOT) という傾向が分かる。これに対し、人手で修正されることによって、文章の最初に背景情報を記述している場合 (BOT→back)、それを回避するように修正されたり、また主張に関する主な理由を述べた直後に反論を書かずに、間に具体例を挿入することを修正者は好む傾向にあることがわかる (main→reb: 減, main→supp: 増)。このように言説構成素の遷移に関する情報は機械的に推敲を達成しようとする際に重要となる手がかりであり、今後修正前の文章に関してこれらのラベルを自動的に同定できるか、そのラベルに基づいて並び換えを行うことで文章の構成を良くすることができるかなどを調査する必要がある。

次に、実際に作業者がどのように修正したのかを紹介するために、人手で修正された小論文の具体例を表 5 に示す。この例では、修正前は (0) の背景に相当する記述から小論文がはじまるという文章の構成であったが、専門家が修正することで、文章頭に書き手の主張を配置し、さらにそれに付随してその主張の理由を次に配置するように修正していることがわかる。このような個別の事例を分析することで、どのような場合にどの種類の修正がなされているのか、またどのような場合は修正されないのかを付与されたラベルに基づいて分析可能になると考えられる。

### 4 首尾一貫性の評価実験

次に、既存の首尾一貫性のモデルを用いて、修正前と修正後の文章対を入力とし、修正後を同定できるかを調査した。この実験のために、談話要素の文間の遷移情報を利用する Barzilay らの entity grid モデル [1] と、我々が以前提案した共参照解析の出力を用いた首尾一貫性のモデルを利用する。entity grid モデルの構築の際は、文献 [4] で採用した手法と同様に、「は」「が」「に/を」「それ以外」の 4 種類のラベルについて遷移の個数を求め、RankingSVM[6] の素性として利用した。また、共参照解析のモデルには Iida ら [3] のモデルを利用した。entity grid モデルと共参照解析モデルの学習には新聞記事に共参照関係がアノテーションされた NAIST テキストコーパス [11] を利用した。実験設定の詳細については文献 [4] を参照されたい。

評価には、評価対象として作業済みの 118 記事のうち、修正前と修正後で談話単位の順序が変更されている 70 記事を利用した。このうち、どのくらい正しく修正後の文章を選択できるかの精度で各モデルの性能を評価する。ベースラインとして、修正前と修正後の文章をランダムに選択するモデルを採用し、結果を比較する。

実験結果を表 6 に示す。また、比較対象とすべき結果として各モデルを NAIST テキストコーパス (新聞記事) に適用した結果も参考結果として掲載する。この結果より、まず新聞記事を対象とした場合それぞれの首尾一貫性モデルはベースラインを 20% 以上上回っていることがわかる。一方、今回作成した修正前・修正後の小論文データを対象とした場合、どちらのモデルもベースライ

表 4: 言説構成素ラベルの局所的な遷移

cur \ next	BOT	EOT	main	elab	supp	the	back	conc	reb	solu	sugg	tran	pro	total
BOT	0(0)	0(0)	1(-4)	0(-2)	0(-3)	85(24)	7(-16)	0(0)	0(0)	0(0)	0(0)	0(0)	4(1)	97
EOT	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0
main	0(0)	10(-7)	12(-1)	87(8)	45(3)	3(-7)	35(10)	26(5)	13(-15)	0(0)	1(0)	10(3)	1(0)	243
elab	0(0)	5(0)	55(-1)	27(0)	68(3)	4(-3)	6(-4)	13(-4)	24(9)	0(0)	1(-1)	7(1)	0(0)	210
supp	0(0)	0(-2)	57(5)	42(-5)	15(2)	0(-2)	9(-1)	9(-5)	17(10)	0(0)	0(-1)	6(-1)	0(0)	155
the	0(0)	4(-3)	40(3)	11(3)	4(-7)	1(0)	34(10)	1(-2)	4(-3)	2(0)	0(0)	5(-2)	0(0)	106
back	0(0)	0(0)	37(1)	30(1)	13(2)	5(-6)	5(4)	2(0)	7(-1)	1(0)	1(0)	3(-1)	0(-1)	104
conc	0(0)	75(18)	2(-4)	1(-2)	0(-2)	1(-1)	0(-1)	0(0)	0(-3)	0(-1)	0(-3)	0(-1)	0(0)	79
reb	0(0)	0(0)	7(-1)	3(-1)	1(0)	1(-4)	1(0)	0(-1)	8(4)	36(2)	20(-2)	2(1)	0(0)	79
solu	0(0)	1(-3)	9(1)	0(-3)	2(0)	0(0)	0(-2)	12(3)	2(0)	5(1)	18(2)	2(1)	0(0)	51
sugg	0(0)	2(-3)	6(1)	4(1)	4(1)	1(-2)	2(-1)	13(4)	0(-2)	7(-2)	5(3)	2(0)	0(0)	46
tran	0(0)	0(0)	17(0)	5(0)	3(0)	1(0)	5(0)	3(0)	3(1)	0(0)	0(0)	1(-1)	0(0)	38
pro	0(0)	0(0)	0(0)	0(0)	0(0)	4(0)	0(0)	0(0)	1(0)	0(0)	0(0)	0(0)	0(0)	5

括弧内の数は談話単位を元の順序に記事を並べた場合との増分を示しており、正の数が増えた数を表す。BOT と EOT はそれぞれ文章頭と文章末尾を表す。

表 5: 人手修正の具体例

ラベル	談話単位 ID	修正結果
0	the (7)	〈ただし〉小学校における英語の早期教育〈が は〉必要である〈という〉。← <b>主張といえる文がこれしかないので、ここへ移動</b>
1	main1 (5)	〈そのような意味では、〉中学校や高校で英語を学習して、「英語が難しい」と苦手意識を持ってしまいう前に、小学校で「英語が楽しい」と思えるような教育をするの〈が は〉、むしろ必要なこと〈 と思う〉。← <b>主張を裏付ける理由をここへ移動</b>
	(6)	日本語と同様に、相手とのコミュニケーションをとる手段として早期から英語に触れていれば、後になってから苦勞して学ぶということに〈 は〉ならない。
2	back1 (0)	私〈が は〉小学校中学年、高学年のときに英語に触れる機会があった。
3	elab1 (1)	それ〈が は〉中学校や高校における「英語の授業」という。〈 よりは、〉
	(2)	遊びの感覚で楽しめる。〈。 〉
	(3)	ものであった。
4	supp1 (4)	例えば、英語の歌を歌ったり、朝の健康観察のときに先生と英語であいさつをしたり、英語を身近に感じることができるものであった〈 と思う〉。
5	main2 (8)	〈前衛の通り ただし、〉子どもたちが「楽しい」と思えるよう〈だ。 な〉
	(9)	ものでない。〈。 と〉
	(10)	〈 英語の早期教育は〉全く意味がなく、むしろ逆効果になってしまう。
6	supp2 (11)	私〈が は〉幸運なことに、小学校だけでなく中学校で〈 も〉英語の授業が「楽しい」ものであると思うことができた。
	(12)	それ〈が は〉何よりも、当時の ALT の先生のおかげであったかもしれない。
	(13)	というのも、彼女と私に〈 は〉共通の趣味があり、よくそのことについて話したり、英語の授業で洋楽を聴いたり、英語で書かれたレシビを見ながらクッキーを作ったりと、私だけでなく誰か楽しめるような時間にしてもらった〈 からである〉。
7	reb (18)	「そんなに早くから子どもに英語を学ばせる必要〈が は〉ない」という人〈が も〉いるであろう。〈。 が、〉← <b>反対意見とそれについての著者の見解なので、起承転結の「転」としてここへ移動</b>
8	solu (19)	その早い時期に子どもが楽しんで英語に触れることができれば、それが子どもの可能性を広げるということに繋がるのではないだろうか。
9	conc (14)	英語に限ったことではない。〈。 が、〉
	(15)	何かを学ぶということにおいて、一番大切な〈が は〉、学ぶ本人がどれぐらい意欲を持〈つ。 つて〉
	(16)	学べるかということである。
	(17)	「もっと知りたい」という意欲が何より本人の力を伸ばすのである。

太字で書かれたコメントが修正を行った作業者のコメントを表す。

ンより結果が悪いことがわかる。この理由として、新聞記事と小論文で捉えるべき談話の特徴が異なることがあげられる。これは、新聞記事などの文章では典型的にはイベントとそれに付随する情報を首尾一貫して記述しているのに対し、小論文では「英語」「小学生」などのキーワードを文間で共有するが、それらが必ずしも主題化され、以降で(局所的に)遷移しているわけではないので、既存の首尾一貫性のモデルで修正後の良さを見積ることができなかつたと考えられる。

## 5 おわりに

本稿では、議論的で論理的な文章を談話的な観点から自動推敲することを目的とし、特に経験的なアプローチを考えた場合に必要となる談話レベルの人手修正データを作成するために、作業者にどのような指示を出すべきかを検討し、その方針を定めた。具体的には、作業者の作業を言説構成素の再配置とラベル付け、またそれにとともなう参照表現・接続表現などの局所的な修正に限定することで、内容レベルの書き換えをできるだけ行わないように制約をかけた。さらに、現在人手作業が完了した作業結果を分析することで、言説構成素のラベルに基づく自動修正の可能性を示した。今後は、このデータを対象にさらに分析を進めることで、専門家が書かれた文章のどのような観点に着目し、どのような場合に修正を行っているのかを明らかにしたい。

## 謝辞

本研究は科研費若手研究 (A) 「談話解析技術に基づいた文章推敲支援」(課題番号: 23680014) の支援を受けた。記して謝意を表す。

## 参考文献

- [1] R. Barzilay and M. Lapata. Modeling local coherence: An entity-based approach. *Computational Linguistics*, Vol. 34, No. 1, pp. 1-34, 2008.
- [2] R. Dale and A. Kilgarriff. Helping our own: Text massaging for computational linguistics as a new shared task. In *Proceedings of the 6th International Natural Language Generation Conference*, pp. 261-265, 2010.
- [3] R. Iida and M. Poesio. A cross-lingual ILP solution to zero anaphora resolution. In *Proceedings of ACL-HLT 2011*, pp. 804-813, 2011.
- [4] R. Iida and T. Tokunaga. A metric for evaluating discourse coherence based on coreference resolution. In *Proceedings of COLING 2012*, pp. 483-494, 2012.
- [5] R. Iida and T. Tokunaga. 談話的な手がかりを利用した日本語の節の受動化. 言語処理学会第 17 回年次大会, 2013.
- [6] T. Joachims. Optimizing search engines using clickthrough data. In *Proceedings of the ACM Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 133-142, 2002.
- [7] Z. Lin, H. T. Ng, and M-Y. Kan. Automatically evaluating text coherence using discourse relations. In *Proceeding of ACL-HLT 2011*, pp. 997-1006, 2011.
- [8] W. C. Mann and S. A. Thompson. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, Vol. 8, No. 3, pp. 243-281, 1988.
- [9] S. Teufel. *The Structure of Scientific Articles: Applications to Citation Indexing and Summarization*. CSLI Publications, 2010.
- [10] 宇佐美慧. 小論文試験による評価データの心理計量学的性質の検討-制限字数の影響に焦点を当てて-. Master's thesis, 東京大学, 2009.
- [11] 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.
- [12] 飯田龍, 徳永健伸. 日本語書き言葉を対象とした参照表現の自動省略-人間と機械処理の省略傾向の比較-. 情報処理学会自然言語処理研究会予稿集, NL-206-15, 2012.