OUKA
The University of Osaka
Institutional Knowledge Archive

Title	英文エッセイの「構造・論理分析ツール」の開発
Author(s)	染谷, 泰正
Citation	サイバーメディア・フォーラム. 2016, 17, p. 11-16
Version Type	VoR
URL	https://doi.org/10.18910/70406
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

英文エッセイの「構造・論理分析ツール」の開発

染谷 泰正（関西大学 外国語学部）

1 はじめに

学生の英語力の低下が叫ばれて久しい。その中でもとくにライティング力の低下はほとんど目をおおわんばかりの状況にある。もっとも「低下」というのは当たらない。日本人大学生の英語力が全国的な規模で高かったためしはかつてなかったからである。このことはとりあえずおいておくとして、英語教師の実感としては、学生の英語力の低迷は疑いようのない事実である。

こうした現状を受けて、関西大学外国語学部では、学生の英語力の実態をより正確に把握することを目的に、学生が書いた英文エッセイの収集を開始し、2013 年度からはこれを「関西大学バイリンガルエッセイコーパス」（略称 KU BE-Corpus, aka KUBEC）プロジェクトという名称のもと、複数の教員による科研費助成研究として実施している。プロジェクトの詳細は別の報告書（山西他 2013, 山西 2013, 山下 2014）を参照していただくとして、この稿では同プロジェクトの一環として開発した英文エッセイの構造・論理分析のためのツール（構造・論理タグエディタ）についてその概要を紹介する。

2 「構造・論理タグエディタ」の概要

2.1 そもそも何をするためのものか？

学生の書いた英文エッセイには、英語の言語的な問題のほかに、構造的問題（＝エッセイ／パラグラフがうまく組み立てられていない）や、論理的・修辞的な問題（＝論理的な整合性がない、文がつながらない、何をいいたいのかよくわからねーぞ！）が多く含まれている。この「構造・論理タグエディタ」は、こうした構造的および論理的・修辞的な問題点を、できるだけ簡単な方法で分析し、抽出することを目的としたツールである。具体的には、その名称が示すとおり、以下の２つのことを行う。

①**機能的構造分析**：学生の作成した英文エッイを構成する各パラグラフに、それぞれどのような機能的構成要素（i.e., Introduction/Opener, Thesis Statement, Organizer, Topic Sentence, Supporting Sentence, Extender, Transitional Sentence, Closing Sentence, Kicker などの義務的要素と選択的要素）があるか、またはどの要素が欠けているかを見る。(学生の書いた英文エッセイには必ずしもこうした標準的な機能には収まらない文要素＝したがってタグが付与できないもの＝が含まれていることがあるが、このツールを使うことで、そのような「逸脱」も適切に把握できるようになることが期待される)。本システムではこのような機能的構成要素に加えるタグを「構造・機能タグ」(Structural-Functional Tag; 略して「構造タグ」または S Tag）と呼ぶ。

②**論理関係および修辞構造分析**：各センテンス間の論理的・修辞的関係（i.e., ある文 (S1) が他の文 (S2) に対してどのような論理的・修辞的関係にあるか。たとえば、S1 は S2 に対して背景情報 (background information) という関係にあるとか、S2 は S1 の例示 (example) という関係にあるといった関係性) を見る。タグは、Rhetorical Structure Theory (Mann, W.C., & Thompson, S.A. 1988) をベースに、その他の研究成果も取り入れて独自に設定した。なお、前記①と同様に、学生の書いた英文には、必ずしもこうした既成の「タグ」で定義され得るような明瞭な論理関係がないもの＝したがってタグが付与できないもの＝が少なからずあるが、このツールを使うことで、そのような「逸脱」も適切に把握できるようになるものと思われる。本システムではこのようなタグを「論理・修辞タグ」(Logical-Rhetorical Tag; 略して L Tag) 呼ぶ。

図 1 は、KU BE-Corpus サイトの投稿管理画面（ログイン後のトップ画面）である。この画面の左下には、「エラータグエディタ」と「構造・論理タグエディタ」の 2 つのツールが用意されている。ここで後

者を選択すると、図 2 に示す「構造・論理タグエディタ」の初期画面が表示される。画面の左ペイン上部には “Load Essay File” と表示された機能バーがあり、これをクリックすると画面右にポップアップウィンドウが表示され、ここから作業対象とするデータファイルを指定する。この画面では見にくいが、このポップアップウィンドウには、すでに登録済みのエッセイデータを年度別またはクラス別に収録したデータのファイル名が表示されている。

作業対象ファイルを選択すると、図 3 のような画面になり、左側ペインにエッセイの作成者名（またはＩＤ）とエッセイタイトルの一覧が表示され、そのうちのいずれかのエッセイを選択する。図 3 は選択したエッセイデータが読み込まれた画面である。

1)　「構造タグ」（S Tag）の挿入

エッセイは、右側のペインに、パラグラフごとにセンテンス単位で自動分割されて表示される。ただし、各センテンスはデータ読み込み後に手動で再分割することも可能である。この図ではいささか見にくくなっているが、各センテンスの左側にはセンテンス番号が付与され、その後に空のボックスが用意されている。この空ボックスをマウスでクリックすると、図 4 に示すような形で S Tag の一覧がポップアップ表示され、この一覧から該当するタグを指定すると、前記の空ボックスに指定のタグが挿入されるとともに、当該パラグラフの末尾にもこれが順次コピーされるという仕組みである。図 4 には、第 1 パラグラフと第 2 パラグラフについて、それぞれ S Tag が付与された状態が表示されている。なお、一覧性を高めるために S Tag はそれぞれ異なったカラーで表示される（カラー指定はユーザが任意に設定することができる）。

2)　「論理・修辞タグ」（L Tag）の挿入

S Tag の挿入が終わったら、次に「論理・修辞タグ」を挿入する（図 5）。前述のとおり、L Tag は原則として隣接する 2 センテンス間の論理・修辞関係を明示的に表示するもので、本システムでは各パラグラフの末尾に横一列に表示されている構造タグ上にアサインする。例えば、[TS] と [SS] という 2 つのタグが並んでおり、後者は前者に対して ELBR

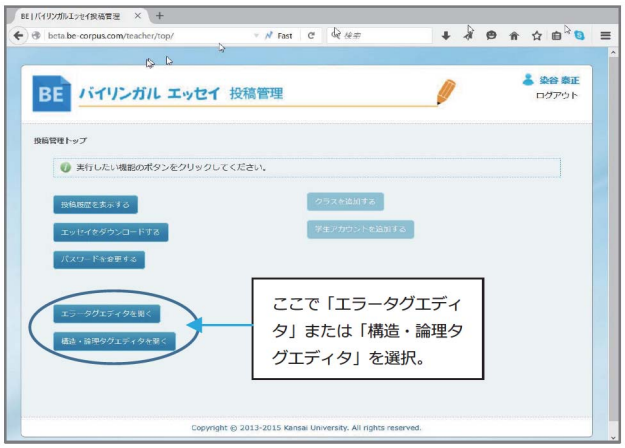

図 1　KU BE-Corpus の投稿管理画面（トップ画面）

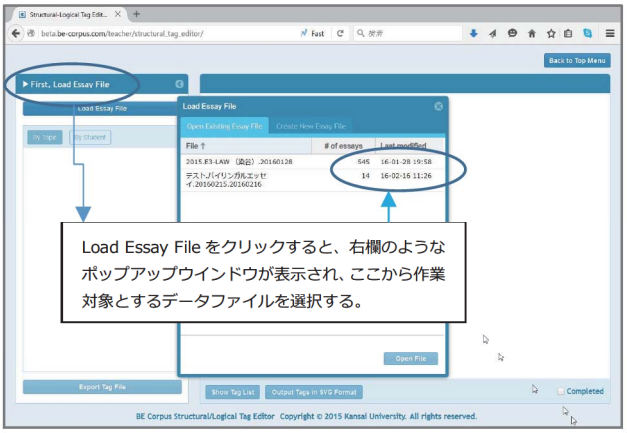

図 2　「構造・論理タグエディタ」の初期画面

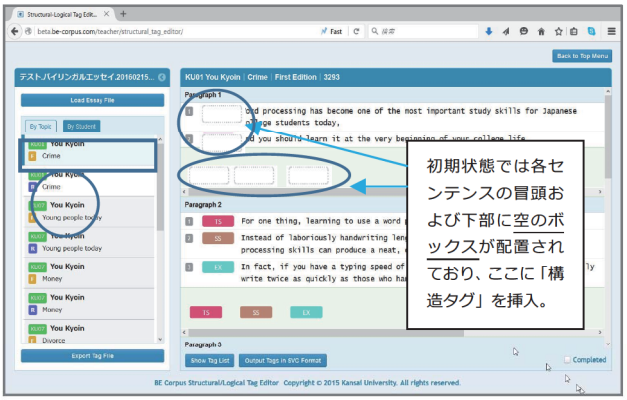

図 3　エッセイデータを読み込んだ画面

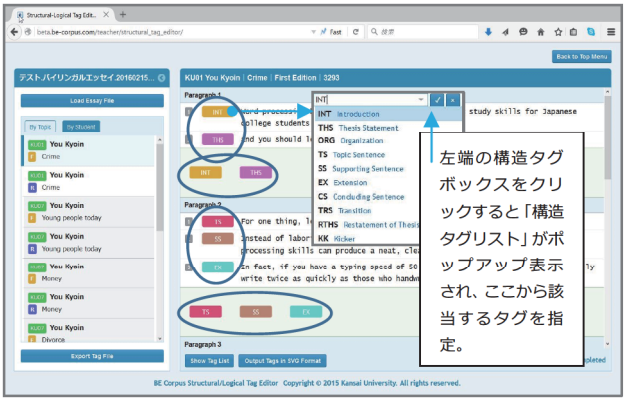

図 4　「構造タグ」(S Tag) の付与

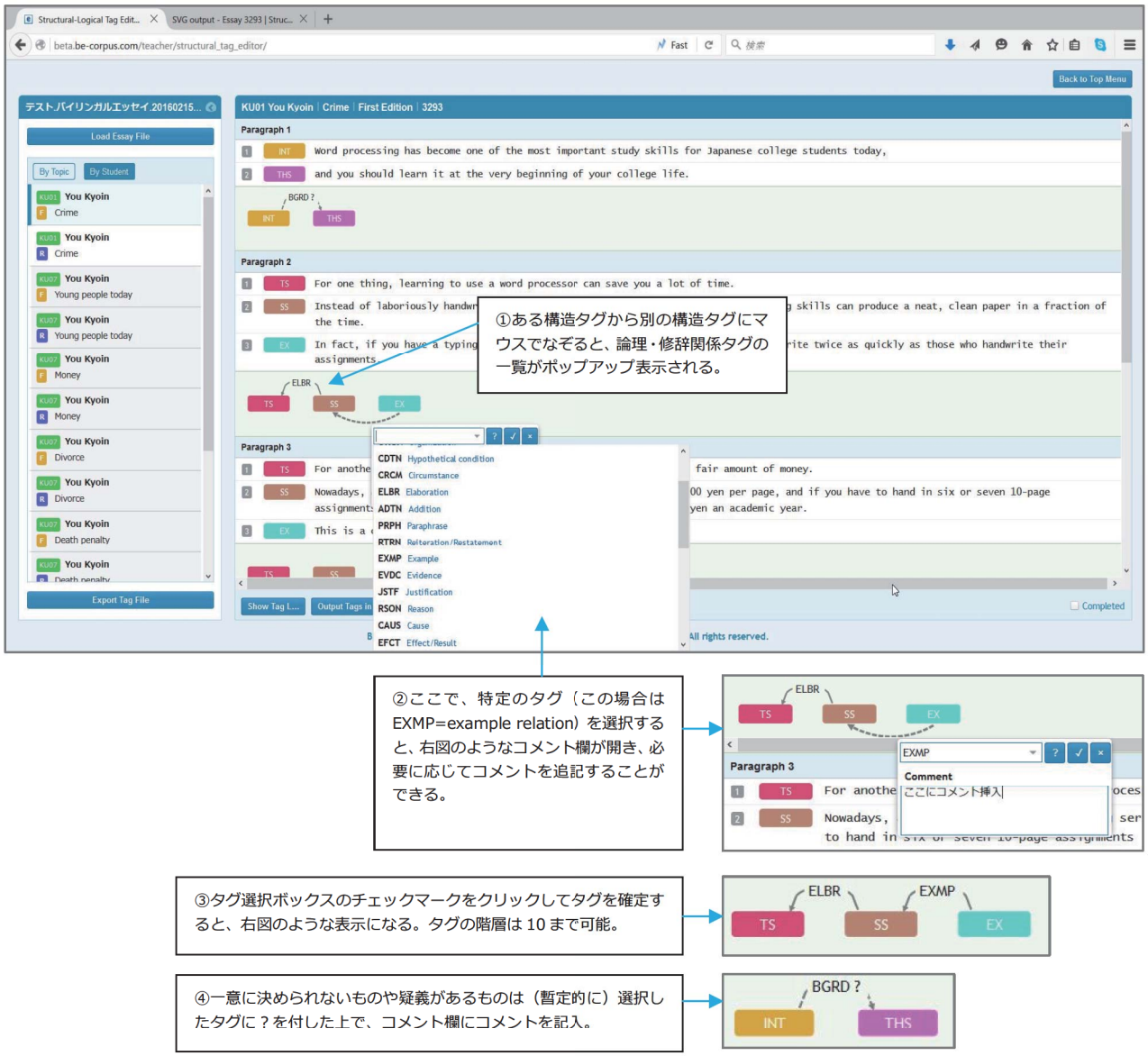

図 5　「論理・修辞タグ」(L Tag) の付与

(Elaboration) という関係にある場合、マウスを[SS]から [TS] に向かってなぞると、L Tag の一覧がポップアップ表示され、ここから ELBR を指定すると、この 2 つの S Tag の上部に矢印で方向が指定された L Tag が挿入される（タグ確定前は構造タグの下に点線で、確定後は上部に実線で表示)。なお、何らかの L Tag を指定すると自動的にコメント欄が開き、ここに任意のコメントを追加することができる（図 5 の解説②参照)。挿入したコメントは、タグ付け終了後にデータを出力する際、自動的に所定の位置に表示される（出力例は図 9 参照)。図 6 にすべてのタグ付け作業が完了した画面例を示す)。

3) データの出力１：テキストデータの出力

ひとつのエッセイのタグ付けが終了したら、次は結果の出力である（ただし、その時点で作業を終了してもよい)。タグ付け結果の出力には、①タグをテキストデータとして出力する方法と、②SVG イメージとして出力する方法の 2 つが用意されている。いずれも、図 6 の画面下部に表示されている [Show Tag List] と [Output Tags in SVG Format] という機能ボタンをクリックすると直ちに所定の出力結果が表示される。前者の場合、さらに S Tag だけを横一列に一括表示させる方法と、センテンス単位で縦一列に表示する方法、および S Tag と L Tag をセンテ

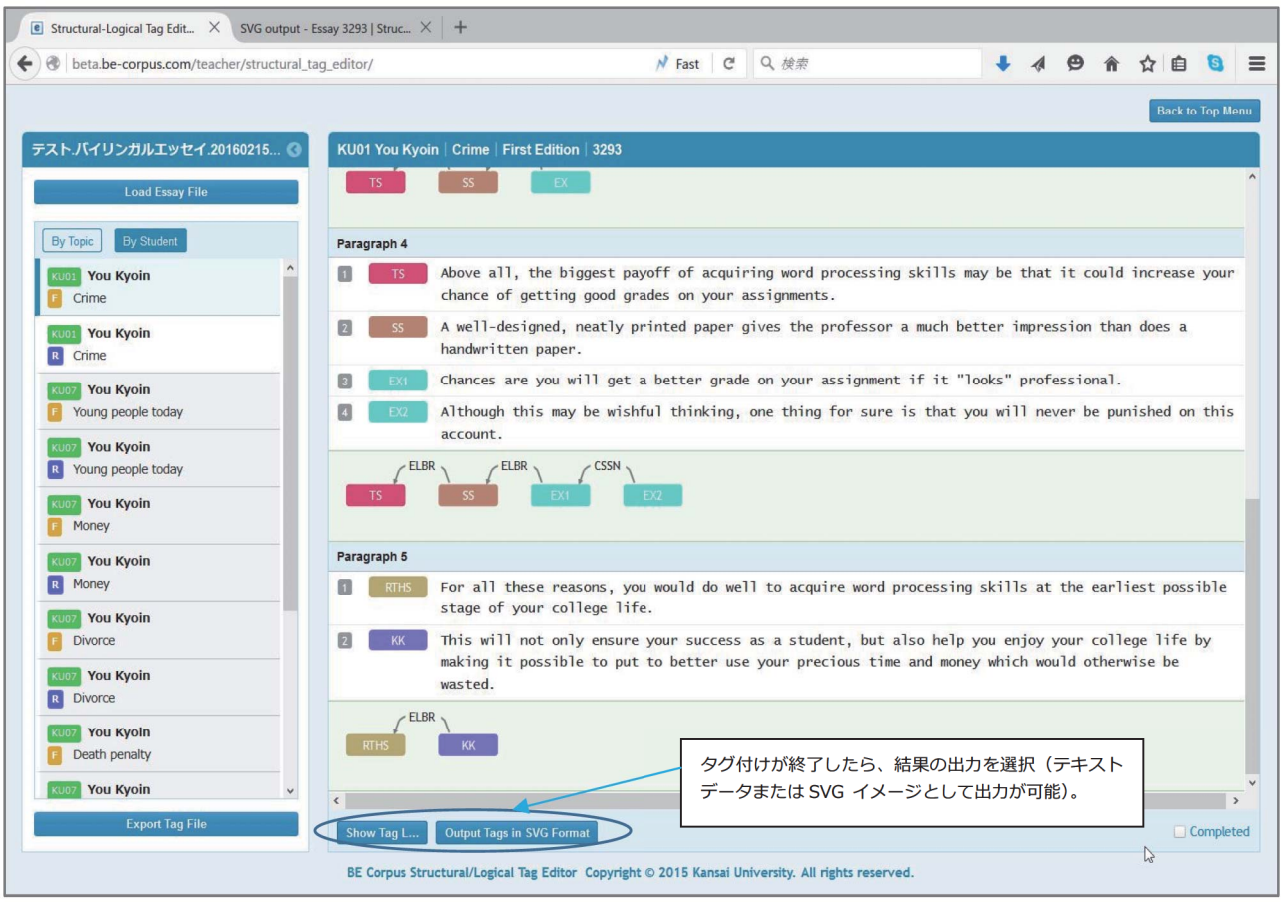

図 6　タグ付け完了画面

ンスごとに並列表示させる方法の 3 つのオプションが用意されており、それぞれ目的に応じて選択することができる。図 7 はこのうち 3 つ目の方法で、すべてのタグをテキストデータとして出力した例である。

```
P1    S1    <INT> <BGRD?: 1,2>
      S2    <THS>
P2    S1    <TS>
      S2    <SS> <ELBR: 2,1>
      S3    <EX> <EXMP: 3,2>
P3    S1    <TS>
      S2    <SS> <CNTR: 2,1>
      S3    <EX> <PRBL: 3,2>
P4    S1    <TS>
      S2    <SS> <ELBR: 2,1>
      S3    <EX> <ELBR: 3,2>
      S4    <EX> <CSSN: 4,3>
P5    S1    <RTHS>
      S2    <KK> <ELBR: 2,1>
```

図 7　タグをテキストデータとして出力した例

このうち、最初のカラムはパラグラフ番号、カラム 2 はセンテンス番号、カラム 3 は構造タグ (S Tag)、カラム 4 は論理・修辞タグ (L Tag) を示す。なお、L Tag の表示シンタクスは <TAG: α, β> という形式で、これは「α は β に対して TAG という関係にある」という意味である。より具体的には <BGRD: S1, S2> は「S1 は S2 に対して BGRD [= Background] の関係にある」という意味であり、同じく<ELBR: S2, S1> は「S2 は S1 に対して ELBR [= Elaboration] の関係にある」と解釈される。

4）データの解釈（例）

図 7 に示した出力データからどのようなことが読み取れるだろうか。もちろん、データの解釈はどこに焦点を当てるか、あるいは何を見ようとしているかによって異なってくるが、ごく概論的には、このデータから当該のエッセイについて次のような評価が可能であろう。すなわち、「このエッセイは P1 から P5 までの 5 つのパラグラフで構成され、各パラグラフ内のセンテンスは、それぞれカラム 3 に示し

たような機能的役割を担っている。また、各パラグラフ内の隣接センテンスは、カラム 4 に示したような論理的・修辞的関係で結ばれている。構造的および論理的に特に破綻していると思われる箇所は見られない」。

要するに、当該のエッセイの中身を詳細に検討する前に、ここまでのことが言えるということである。詳細な検討の対象となるのは、主としてこの段階で顕著な、あるいは面白い「逸脱」がみられるデータということになる。本プロジェクトのように数百万語におよぶ大量のデータを扱う場合、すべてのデータについてその構造的および論理的・修辞的な特徴や問題点について詳細に分析することは不可能であり、何等かの方法でデータのスクリーニングをする必要がある。前述のこのツールの目的からして、ここまでのことができれば、とりあえずほぼ十分であると考えられる。

5）手作業での構造明示化と再分析

図 8 は、前記のタグデータをワープロの図形編集機能を使ってより詳細に構造化（基本構造の可視化）した上で、このデータに特徴的な論理展開を赤字でハイライトし、さらに当初の分析では明らかになっていなかった推論（inference）に基づく非明示的な修辞関係を手動で明示化したものである。

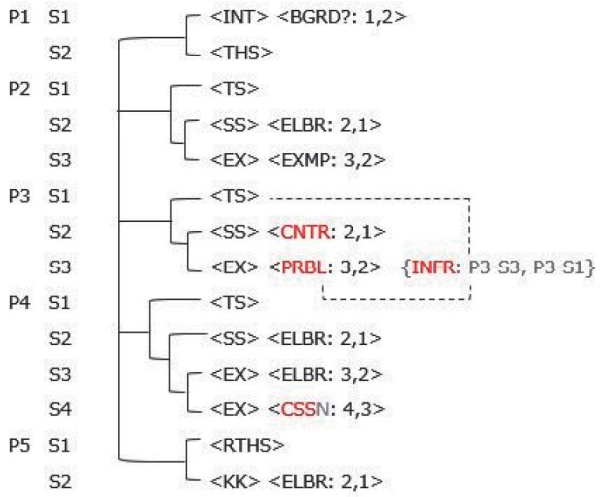

図 8　手作業での構造明示化と再分析

この分析図から、次のようなより高度な再解釈を導き出すことができる。すなわち、「当該のエッセイの論理構造を見ると、全体としてはほぼ標準的なものになっているが、P3 での <**CNTR**> (counter argument), <**PRBL**> (problem statement)、および P4 での <**CSSN**> (concession) の出現が特徴的である。このうち、前者は TS (P3-S1) でこのパラグラフのトピックを提示したあと、次の文でそれとは反対の議論を起こし、これを問題提起としている。この問題提起 (P3-S3) は、これ以降、明示的には敷衍されておらず、TS との論理関係は読み手の推論 (inference) に任せている（図 8 の {**INFR: P3-S3, P3-S1**} 参照）。ここは、いわば日本語的な発想がよく出ている箇所と言ってよいだろう（英文の規範からいえば、『だからワープロを自分で使うことで余計な出費を抑えるようにすべきである』のような文を置いて、このパラグラフの TS との論理関係を明示的に示すことが書き手の責任であるとされる）。P4-S4 の <CSSN> は、ここで筆者が主張していることが必ずしも常に真であるとは限らないが、それでもなお P4-S1 での主張には利点がある、という論法である」。

なお、紙面の都合上、ここでは本文データ (300 words) を示すことができないが、この再解釈・再分析は、当該のエッセイ本文を参照しながら行ったものである。タグデータからわかるのは、全体の傾向や特徴的なパターンのみであって、それ以上のことはやはり本文そのものに立ち戻って詳細に検討する必要があることは言うまでもない。

6）データの出力２：SVG イメージの出力

図 9 は、タグデータを SVG イメージとして出力した例である。前述のとおり、タグの指定画面で何らかのタグを選択すると自動的に「コメント欄」が開き、必要に応じて任意のコメントを追記することができるようになっている（図 5 の解説②参照）。ここで挿入したコメントは、図 9 に示したような形式で出力させることができる。たとえば、Paragraph 1 では [INT] (introduction) と [THS] (thesis statement) が [BGRD] (background information) という関係で結ばれているが、見方によってはこれは "Justification" という関係にあるということもできる。このように、何らかの疑義がある場合は、とりあえず選択

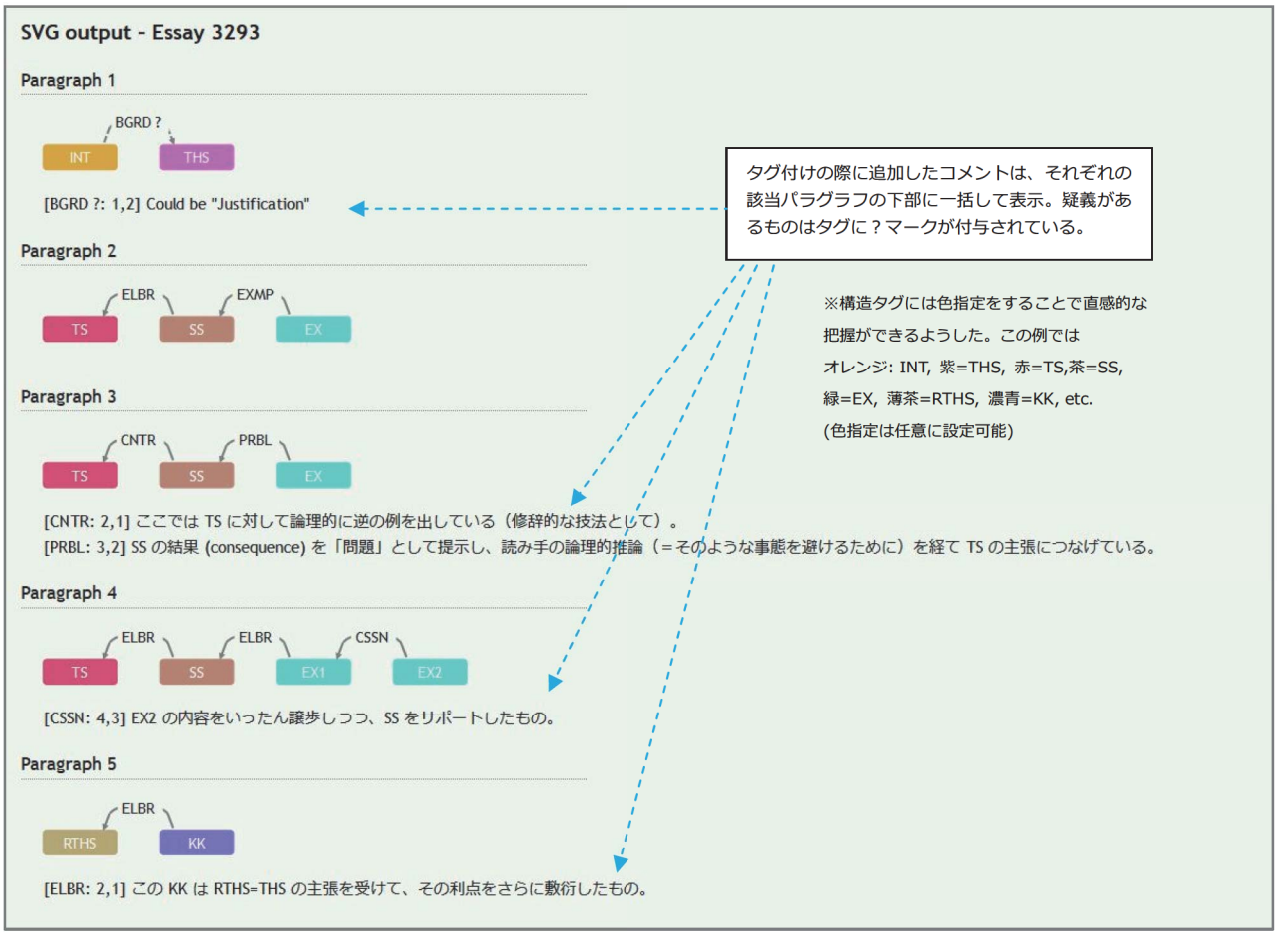

図 9　分析データを SVG イメージとして出力した例

したタグにクエスチョンマーク (?) を付した上で、コメント欄にコメントを追記し、作業終了後に、これを一括して出力させることができる。タグは必ずしも一意に決めることができない場合があり、この機能はタグ付け作業後の詳細分析の際に、大いに役に立つものである。

3　まとめと今後の課題

以上、 KU BE-Corpus プロジェクトの一環として筆者らの研究チームが開発した「構造・論理タグエディタ」の概要について紹介した。このツールはようやく開発が終わった段階であり、今後は、これをすでに収集済みの大量のエッセイデータに対して適用し、その有効性を検証するとともに、問題点や課題等を明らかにしていきたいと考えている。なお、このツールは、筆者らが開発したこの他の分析ツールとともに、適切な時期に一般公開する予定である。

※本研究は平成 26～28 年度科学研究費補助金の助成を受けて行われたものです（基盤研究（B）研究課題番号：26284085　研究代表者：山西博之（関西大学））。ただし、本稿の内容についての責任は筆者に帰します。

参考文献

山西博之・水本篤・染谷泰正（2013）「関西大学バイリンガルエッセイコーパスプロジェクト：その概要と教育研究への応用に関する展望」『関西大学外国語学部紀要』9, 117-139.

山西博之（2013）「バイリンガルライティング授業に対する学生の意識：「振り返りアンケート」のテキスト分析結果から」『JACET 関西支部ライティング指導研究会紀要』10,　57-62.

山下美朋（2014）「関西大学バイリンガルエッセイコーパスの構築と研究の可能性」『英語コーパス研究』22, 19-35.