

KNB コーパス：アノテーション仕様書

橋本 力 黒橋 禎夫 河原 大輔 新里 圭司

平成 21 年 9 月 25 日

目 次

1	はじめに	2
2	文区切りアノテーション仕様	2
3	括弧抽出アノテーション仕様	3
4	形態素アノテーション仕様	5
4.1	誤変換、脱字、衍字	5
4.2	口語的表現	6
4.2.1	方言等	6
4.2.2	外国語	7
4.2.3	擬音語、擬態語	8
4.2.4	意図的な言い淀み／言い直し	8
4.3	創造的表現	8
4.4	その他の問題：個別事例	9
4.4.1	形態素／タグ単位の区切り	9
4.4.2	その他	9
5	係り受けアノテーション仕様	9
5.1	適切な係り先が存在しない場合の対応	10
5.1.1	括弧外の文節の係り先が誤って括弧内に入っている場合	10
5.1.2	その他	10
6	格・照応・省略アノテーション仕様	10
7	固有表現アノテーション仕様	11
7.1	個別事例	11
8	評判アノテーション仕様	12
8.1	評判コーパスファイルのデータ構造	12

1 はじめに

近年、ブログを対象とした情報アクセス・情報分析技術が盛んに研究されている。我々は、この種の研究の基礎データの提供を目的とし、249 記事、約 4,000 文からなる、人手による解析済みブログコーパス（以下、KNB（**K**yoto **U**niversity and **N**TT **B**log）コーパス）を構築した。KNB コーパスの主な特長は次の 4 点である。

- (1) a. 文区切りアノテーション
- b. 京都大学テキストコーパス（以下、京大コーパス）互換の、形態素、係り受け、格・省略・照応、固有表現アノテーション
- c. 評判表現アノテーション
- d. アノテーションを可視化した HTML ファイルの提供

本仕様書では、以下に挙げる、京大コーパスのアノテーション仕様にはない独自のものを中心に、KNB コーパスのアノテーション仕様を説明する。

- (2) a. 文区切りアノテーション仕様
- b. 括弧抽出アノテーション仕様
- c. ブログの特徴に対応するための形態素、係り受け、格・省略・照応、固有表現アノテーション仕様
- d. 評判アノテーション仕様

以下の内容は、京大コーパスに付属のマニュアル [2, 1] (3) を前提としている。本仕様書に記述のないアノテーション仕様に関しては、これらのマニュアルを参照されたい。

- (3) a. 形態素・構文情報のタグ付け基準マニュアル
http://nlp.kuee.kyoto-u.ac.jp/nl-resource/corpus/KyotoCorpus4.0/doc/syn_guideline.pdf
- b. 格関係、照応・省略関係、共参照情報のタグ付け基準マニュアル
http://nlp.kuee.kyoto-u.ac.jp/nl-resource/corpus/KyotoCorpus4.0/doc/rel_guideline.pdf

2 文区切りアノテーション仕様

Web 上の文書では、(4) のような区切りが明確な文だけでなく、(5) にあるような区切りが不明確な文もある（[EOS] はアノテートされた文区切りを、↓ は元の記事における改行を表す）。

- (4) a. 私はプリペイド携帯をずっと使っている。[EOS]
 - b. ヒマな大学生の人はチケット買って西京極でボクと握手！↓ [EOS]
- (5) a. なぜか清水寺に着きました笑 [EOS]

- b. これに決めた !! ↓ [EOS]
 と思ったら… ↓ [EOS]
 なんと品切れ。 [EOS]

- c. 京都のほうだったような … [EOS] まあ観光スポット多いですね?京都は。 [EOS]

原則的には、母語話者の直観に基づき、一文として最も適切だと思われる個所で区切る。それ以外に、(6) に挙げる個別的な方針を導入した。

- (6) a. 日付だけからなる行も一文とする

- 「2006 年 10 月 09 日 ↓ [EOS]」

- b. URL だけからなる行も一文とする

- 「<http://www.shigureden.com/> ↓ [EOS]」

- c. 簡条書きの一行一行をそれぞれ一文とする

- 「・ 藤井大丸 ↓ [EOS]
 セレクトショップが多いです。 [EOS] 大学生だと UNITEDARROWS や… ↓ [EOS]
 ・ 紀伊國屋書店 ↓ [EOS]
 新京極通三条下ルらしいです。 [EOS] 前は河原町にあったブック… ↓ [EOS]」

- d. 途中に URL が含まれていても、全体を一文とする

- 「<http://url/>とでも入力すると、 ↓ <http://url/> ↓ と出ます。 [EOS]」

- e. 途中に文末記号があっても、明らかな一文である場合は区切らない

- 「散歩 ?? かな。 ↓ [EOS]」

- f. 文頭、文末の記号は一文に含める

- 「そんな日本語ないか。笑 ↓ [EOS]」
- 「脱力。 ORZ ↓ [EOS]」
- 「P.S. 数年前、電車の… [EOS]」

3 括弧抽出アノテーション仕様

京大コーパスでは括弧表現は削除されていたが、ブログ記事の括弧表現は、新聞記事と比べると、本文と密接不可分な内容のものが多く無視できない。(7) は KNB コーパスの括弧表現の例である。

- (7) a. ここの F o o d (パンメインでカレーとか) の量は少なかったなー
b. 貴重な (まあどのへんが貴重なかはわからないけど) 時間を無駄にしよう。
c. どでかい神楽松明 (激しく燃えている!) を担いで、狭い鞍馬街道をどこからともなく練り歩き出す。

一方で、ブログ記事の括弧表現は多種多様で、文内に埋め込まれたままだと、係り受け等のアノテーションが困難になる。そこで、括弧表現を文中から取り出して、一つの独立した文とした。(9)は(8)から括弧表現を抽出したものである。

(8) # S-ID:KN012_Gourmet_6-1-23

ここも F o o d (パンメインでカレーとか) の量は少なかったなー

(9) a. # S-ID:KN012_Gourmet_6-1-23-01

ここも F o o d の量は少なかったなー

b. # S-ID:KN012_Gourmet_6-1-23-02 括弧タイプ:例示 括弧位置:7 括弧始: (括弧終:)
パンメインでカレーとか

「#」の行は文 ID を表す。抽出され一文になった括弧表現は、元の文の直後に置かれ、新たに文 ID が与えられる。具体的には、元の文の ID 末尾に「-01」を付与し、抽出されて別の文となった文の ID には「-02」「-03」「-04」等を付与する。例えば(8)は ID が「KN012_Gourmet_6-1-23」だが、括弧抽出後は、「KN012_Gourmet_6-1-23-01」(9a)と「KN012_Gourmet_6-1-23-02」(9b)の二文になる。

なお、括弧表現の元の文における位置情報が記録されており、復元も可能である。(9b)における「括弧位置:7」がその情報にあたる。

さらに、抽出された括弧文には、年齢、読み、日付、金額、場所、所属、同義、文、その他のいずれかの括弧タイプが与えられる。(10)に、新聞記事での括弧タイプの例を挙げる。

(10) **年齢**：共産党は一日付の機関紙「赤旗」で、宮本顕治議長(86)のインタビューを掲載した。

読み：こんな単純な発想にあやうさ、脆(もろ)さを感じる人は多いでしょうが、混迷の転換期を乗り切るため「日本は変わった」ことの証(あかし)であり、メッセージになるはずです。

日付：日本の選挙制度は一八八九(明治二十二)年、小選挙区制中心でスタートしたのだが、現行の制度は、この男子普選法が原型となっている。

金額：昨年十月には総額六百万カナダドル(約四億四千万円)を投じて、鉄筋三階建ての編集・印刷センターを建設した。

場所：2年前の世界選手権(カナダ・ハミルトン)で、男子3、女子1の金メダルを獲得した開催国・日本がお家芸の座を守れるか。

所属：女子では、48キロ級で61連勝中の田村亮子(帝京大)に連覇の期待がかかる。

同義：アジア・太平洋経済協力会議(APEC)は、今年五月にもソウルで情報・通信担当閣僚を集めた「APEC情報サミット」を開催する。

文：秋山和慶の音楽監督就任三十周年、そして第四百回定期という記念すべき機会に演奏された、シェーンベルクの未完のオペラ《モーゼとアロン》のサントリーホールにおける演奏会形式による(しかし作曲者の指示を守り、可能なかぎりの演出、装置や照明等を備えた)公演は、この作曲家が生涯の課題とした宗教的命題と芸術表現の問題を一气呵成(かせい)、劇的に演じきるものであり、「旧約」の成立の場を感動的に表現するものであった。

その他：地元最大の中国語日刊新聞「星島日報」(本社・香港、四万八千部、古偉凱編集局長)が、その実態を語る。

(11) のような列挙のための数字付きの括弧は例外とし、文中に括弧を入れたままとする。

- (11) 改革の柱として (1) 財界・大企業本位の政治にメスを入れる (2) 日米安保と在日米軍の横暴に取り組む――の二点を挙げた。

4 形態素アノテーション仕様

KNB コーパスの形態素アノテーションは、品詞・活用体系、フォーマットともに京大コーパスの仕様に準拠しているが、次のようなブログの特徴に対応するため仕様を拡張した。なお、以下で「/」は形態素区切りを表す。

- (12) a. 誤変換、脱字、衍字
- b. 口語的表現（方言、外国語、擬音／擬態語、意図的な言い淀み／言い直し）
- c. 創造的表現（記号、Web で頻出のスラング）

また §4.4 で、KNB コーパス形態素アノテーションで議論になった上記以外の問題の個別事例を挙げる。

4.1 誤変換、脱字、衍字

(13) に例を挙げる（矢印 → の右側が誤変換、脱字、衍字の例）。

- (13) a. 誤変換：「通信機能が 内蔵 されたもの」→「通信機能が 内臓 されたもの」
- b. 脱字：
- 「早めに 行かないと」→「早めに かないと」
 - 「属する ようになり」→「属する よう なり」
- c. 衍字：
- 「高級な 料亭 や、焼肉屋、…ダイニングも多い。」→「高級な り料亭 や、焼肉屋、…ダイニングも多い。」
 - 「何が安いのか、考えて 買って いきます。」→「何が安いのか、考えて 買いつて いきます。」

誤変換と脱字には (14) のように対応した。

- (14) a. （元の表現を残しつつ）正式な書き方・表現に基づいてアノテーションする。例えば「早めにかないと」なら「早めに行かないと」としてアノテーションする。
- b. メモ欄に、「ER:(正しい書き方)」のように、誤りフラグ ER と正しい書き方を記載。

衍字には (15) のように対応した。

- (15) a. 正しい語の前または後ろにある衍字の場合（「高級な り 料亭」など）

- 形態素は「り」と「料亭」に分け、全体で1タグ単位とする。衍字「り」の品詞は未定義語とする。また、メモ欄に「ER:料亭」と記載する。
- b. 一語の中に衍字が入っている場合（「買 いってきます」など）
- 元の表現を残したまま、正しい表現の品詞情報を与える。また、メモ欄に「ER:買って」と記載する。

4.2 口語的表現

方言や外国語、擬音／擬態語、意図的な言い淀み／言い直し等がこれに該当する。

4.2.1 方言等

まず、方言に対して、メモ欄に「DI」(dialect)と書いた。

また、方言では活用をどう記述するかが問題となる。我々は、京大コーパスとの互換性と文法記述の正確性を最大限確保するため、形態素解析器 JUMAN¹ の活用記述法に準拠して、既存の活用に該当しない方言に対して、新たな活用を定義した。(16)に例を挙げる。

(16) a. 関西方言の「や」等

- 「面倒やん」…ナ形容詞(面倒だ)や列基本形
 - 「ちいさいもんや」…判定詞(だ)や列基本形
 - 「楽しいやろ?」…助動詞(だろう)や列基本省略推量形
 - 「市バスやなくて」…判定詞(だ)や列基本未然形
- b. 「きれいジャン」は「きれいジャ/ン」と解析する。「ン」は述語接尾辞「ない」の音便基本形

その他、KNB コーパス形態素アノテーションでの個別事例を以下に挙げる。

(17) a. 「いい/んじゃ/け」…1タグ単位

- b. 「剣道/し/よる」…「よる」動詞性接尾辞
- c. 「化石/化/し/とる」…「とる」動詞性接尾辞
- d. 「頑張り/よる」…「よる」動詞性接尾辞
- e. 「持って/へん/わあ」…「へん」助動詞無活用型。「わあ」は終助詞
- f. 「チェック/せ/な/あき/ま/へん」…「チェック」サ変。「せ」は「する」の文語未然形。「な」接尾辞「ない」の語幹。「あき」は「あく(明く)」子音動詞カ行の基本連用形。「ま」は動詞接尾辞「ます」の未然形。
- g. 「食べ/ない/つしよ」…「つしよ」は「助動詞だろう型促音便形デス列基本省略推量形」(i.e. でしょ)

¹<http://nlp.kuee.kyoto-u.ac.jp/nl-resource/juman.html>

- h. 「まあ/ええ/か」...3 形態素。「ええ」は「えい」の音便基本形
- i. 「取れ/んく/なる」...「取れ」:「取れる」の未然形。「んく」: 接尾辞「ない」の音便連用形。
- j. 「帰れ/つつわ/れて/も」...「つつう」の未然+「れて」(「れる」のタ系連用て形)+「も」副助詞
- k. 「身/が/もた/ねー」(「身が持たない」の意)...「もつ」未然形+「ない」音便基本形
- l. 「そっ/から/また」(「そこからまた」の意)...「そっ」指示詞(新たな指示詞)
- m. 「~/つつゅう/こと/かい/な」...「つつゅう」は「つつう」の音便基本形。「かい」「な」終助詞。
- n. 「もの/大切に/しい/や」...「しい」は「する」の音便省略命令形。「や」終助詞
- o. 「もの/大切に/し」...「し」は「する」の省略命令形(「しろ」の省略)。
- p. 「連れて/こー/っと」...「こー」は「く」の意志形異表記。「っと」終助詞
- q. 「足ん/ない」...「足り(/ない)」の音便未然形
- r. 「売って/ちょ」...「ちょ」終助詞
- s. 「バイキング/形式/やって/ん/けど」...「やって」は「だ」のヤ列タ系連用テ形。「ん」は「んだ」の語幹。「けど」接続助詞

4.2.2 外国語

外国語とは、例えば「今日は/3限を/さばらせて/友達を/連れて/祇園|界限へ/G O !!」における「G O」である。外国語に対しては次のように対応した。

- (18) a. 基本的にサ変名詞か形容詞としてとらえる。
 - 「祇園界限へG O !!」...「G O」はサ変名詞
 - 「GREAT/□/VIEW」...3 形態素。1 タグ単位。「GREAT」がナ形容詞。「VIEW」がサ変名詞。
 - 「LOVE だ」「HAPPY だ」...ナ形容詞。
- b. サ変と形容詞以外は普通名詞としてとらえる。
- c. わからない場合は「記号」とする。
- d. 「ヨミ」は英単語アルファベットそのままとする。
- e. 外国語のフレーズは、一単語一形態素とし、フレーズ全体で1 タグ単位とする。
 - 「with/□/my/□/friend」...5 形態素。1 タグ単位。

4.2.3 擬音語、擬態語

KNB コーパスでは、JUMAN 辞書に登録されていない、『ピリリリリリ』『ポフュー』等の擬音語・擬態語が頻出した。

- (19) a. 「携帯『ピリリリリリ』」
b. 「『ポフュー』って粉が出て」

これらは全て副詞とした。

4.2.4 意図的な言い淀み／言い直し

(20) の下線部は意図的な言い淀みの例である。

- (20) 牛乳を入れて… ぎゅうにゆ… にゆ…

これらは未定義語として、メモ欄に「言い淀み」あるいは「言い直し」と記載する。

4.3 創造的表現

顔文字等の記号や、「サーバ」を意味する「鯖」等の Web 上で多用されるスラングは、ブログ特有の表現といえるが、これらを創造的表現と呼ぶことにする。

(21) に記号の扱いを述べる。

- (21) a. 「…」 「。。。」「～」「ー」等の扱い
- 「…」 「。。。」：直前の形態素に含めずに記号として独立させる。
例) 「疑問/です/が/…」 「とっても/おいしく/て/。。。」
 - 「～」「ー」：直前の形態素に含める（独立した記号とはしない）。
例) 「一緒に/いき/ませ/ん/か～」 「いい/よー」
 - 「楽しーけどさ」「やめまーす」「つれて行こーっと」等、活用する語の一部の場合、活用形は「～異表記」とする。例えば、「見つけたっ」の場合は「母音動詞タ形異表記」とする。
- b. 同じ記号の連続：1 形態素にする。
- 例) 「一緒に/いこ/?/!!」 「買う/人/も/いる/ん/です/けど/ねー/……」
- c. 顔文字：1 形態素にする。
- 例) 「(* ^^ *)」 「一人/で/味わい/まし/た/orz」
- d. 「♪」「→」は記号にする。
- 例) 「これ/は/楽しい/♪」 「貴船/→/鞍馬」

KNB コーパスにおけるスラングは、「サーバ」を意味する「鯖」と、「マスコミ」を意味する「マスゴミ」、「終わった」を意味する「オワタ」とがあった。スラングも、他の表現と同様に京大コーパスとの互換性を最大限保つよう配慮する。例えば上記の「鯖」は、普通名詞として扱う。なお、スラングには、メモ欄に「スラング：(正式/一般的な表記)」を付記する。

4.4 その他の問題：個別事例

4.4.1 形態素／タグ単位の区切り

(22) a. 複合語

- 「すっ飛ばして」は、「すっ」は動詞接頭辞。「かつ飛ばす」と同様。
- 「小/一/時間」(3 形態素)の「小」は名詞接頭辞
- 「進/々/亭」の「亭」のように他でも使われる、生産的な形態素なら独立した1 形態素にする。「進/々」は2 形態素1 タグ単位で、「普通名詞/記号」。

b. 連濁する形態素列

- 「レジ/際/に」は3 形態素。表記:際。ヨミ:ざわ。原形:際
- 「三/羽/ガラス」は3 形態素。表記:ガラス。ヨミ:がらす。原形:カラス (cf. さしみ じょうゆ)

c. 副詞

- 「ふらっと」「ぼーっと」など:1 形態素

4.4.2 その他

(23) a. 「かしらん」全体で終助詞

b. 「たむろって」の活用形の記述法は「タ形連用テ形促音便形（って）」

c. 「てか」は接続詞

d. 「コミュ」は普通名詞

e. 「変」だけでサ変名詞

f. 「なんでもかんでも」は1 形態素の副詞

g. 「如かず」の「ず」は助動詞「ぬ」の基本連用形。「書かず」などと同様。「如く」の子音動詞カ行の未然形

h. 「思い/きや」は「動詞/接続助詞」

5 係り受けアノテーション仕様

形態素のアノテーションと同様、係り受けのアノテーションも京大コーパスの仕様に準拠している。形態素アノテーションではブログ特有の現象が多く存在したため、京大コーパスの仕様の拡張がいくつか必要になったが、係り受けアノテーションでは、§5.1.1 の場合をのぞいて、仕様の拡張はほとんど必要なかった。

5.1 適切な係り先が存在しない場合の対応

5.1.1 括弧外の文節の係り先が誤って括弧内に入っている場合

筆者の誤りによって係り先の語が括弧に入れられ、括弧抽出処理によって別の文になっている場合がある。その場合、抽出されてしまった係り先文節の係り先文節を係り先とする。

次の例（KN001_Keitai_1-1-7の一部）では、「壊れる」の係り先は括弧に括られている「前に」である。つまり、閉じ括弧の位置が間違っている（「見えなくなる」と「前に」の間にくるべきである）。

KN001_Keitai_1-1-7

01 …壊れる（画面が見えなくなる前に）他の携帯を手に入れようと…

上記の例文は括弧抽出処理を経て次のように分けられる。

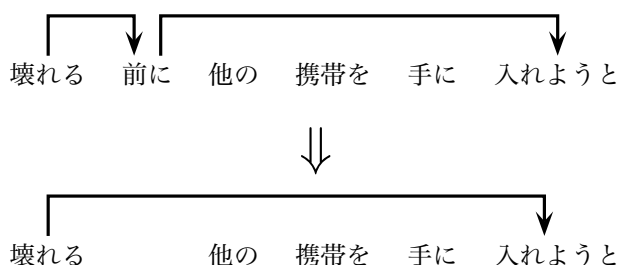
KN001_Keitai_1-1-7

01 …壊れる他の携帯を手に入れようと…

02 画面が見えなくなる前に

つまり、閉じ括弧位置の間違いのため、「壊れる」の係り先「前に」がなくなってしまう。

この場合、「壊れる」の係り先を、「壊れる」の本来の係り先である「前に」の係り先「入れようと」にする。



5.1.2 その他

話し言葉では「言いさし」や「ねじれ」と呼ばれる現象があり、その場合、適切な係り先が存在しない文節がある。KNB コーパスではそのようなケースは存在しないが、今後コーパスを拡張していく上でそういう現象に遭遇した場合は、京大コーパス仕様にに基づき、次のように対応する。

- 適切な係り先のない文節は、左から右へ方向へ、最も適切と思われる文節にかかるようにアノテーションする。

6 格・照応・省略アノテーション仕様

京大コーパスに準拠する形式で、格・省略・照応のアノテーションを付与した。(24)、(25)、(26)に、格、省略、照応の例を挙げる。

(24) Food は量が少ない。(ガ格：量, 外の関係：Food)

(25) 4, 5 回しか 行ったことないけど。(ガ格：一人称)

(26) a. 父と野球。

b. 父 は野球が好きだった。(父 = 一文前)

7 固有表現アノテーション仕様

IREX'99 準拠の固有表現アノテーションを付与した。付与した情報は次の通りである。

(27) a. ORGANIZATION … 組織名

b. LOCATION … 地名

c. PERSON … 人名

d. ARTIFACT … 固有物名

e. PERCENT … 割合表現

f. MONEY … 金額表現

g. DATE … 日付表現

h. TIME … 時間表現

i. OPTIONAL …

(28) に例を挙げる。

(28) a. 京都 を回ってみようと思います☆ (京都：LOCATION)

b. 近鉄 ファンだった。(近鉄：ORGANIZATION)

7.1 個別事例

(29) a. 「英/国籍」の「英」、「京」など、「英国」「京都」を指す場合は品詞を名詞：地名とし、NE：LOCATION を付与する。

b. 「金閣寺というものはない」のように「金閣寺」という名前を問題にしているときは NE を OPTIONAL とする (LOCATION や ORGANIZATION としない)。

c. 「KUNTT Kyoto-Keitai」「sound of music」などのタイトルは 全体で 1 タグ単位とし、NE：ARTIFACT を付与する。

d. 「MNP」「Java」は品詞：固有名詞とし、NE：ARTIFACT を付与する。

e. 「マ〇ー」(注:「ママー(スパゲッティ)」のこと)のように〇で隠している場合は OPTIONAL とする。メモ欄に「ボカシ」と記載する。

f. LOCATION と ORGANIZATION のどちらを付与すべきか迷った場合は、OPTIONAL とする。

8 評判アノテーション仕様

評判アノテーションは、何らかの評判表現を含む文に対して次の情報を付与することで行った。

- (30) a. 評判保持者：評判の保持者。文中に存在しない保持者は [...] でマークされる。
b. 評判表現：文中での評判を表している部分。
c. 評判タイプ：評判の種類と評判の極性。

当為：提言、助言、対策。「～すべきだ」「～しましょう」

要望：希望、要求。「～してほしい」「～を求める」

感情 +/-：気持ち。「好き」「悲しい」

批評 +/-：賛成と反対、称賛と批判。「素晴らしい」「納得できない」

メリット +/-：利点と欠点。「効果がない」「うるさい」

採否 +/-：積極的利用、推進。「利用する」「導入する」「採用する」

出来事 +/-：良い／悪い出来事や状態。「壊れた」「受賞した」

- d. 評判対象：評判の対象。文中に存在しない保持者は [...] でマークされる。

(31)(32) に例を挙げる。

- (31) おかきやせんべいの店なのだが、これがオイシイ。

- a. 評判保持者：[著者]
b. 評判表現：オイシイ
c. 評判タイプ：批評 +
d. 評判対象：おかきやせんべい

- (32) 貧乏人臭い、なんか怪しげな人っぽいといった類のものです。

- a. 評判保持者：[不定]
b. 評判表現：貧乏人臭い、なんか怪しげな人っぽいといった類のものです。
c. 評判タイプ：批評 -
d. 評判対象：[プリペイドユーザ]

8.1 評判コーパスファイルのデータ構造

評判コーパスファイルは、評判情報の有無に関らず KNB コーパスの全文を含んでいる。一行が一文の情報を保持し、タブで区切られた次のフィールドから成る。

- (33) a. 文 ID
b. 文

- c. 評判保持者
- d. 評判表現
- e. 評判タイプ
- f. 評判対象

一文に複数の評判表現がある場合、それらの「評判保持者」「評判表現」「評判タイプ」「評判対象」はそれぞれ、「\n」で区切られて与えられる。次の例では、「しょぼかった」「画面も小さかった」「音も3和音とかやった」の3つの評判表現が一文に含まれている。

(34) 文 ID: KN027_Keitai_1-1-6-01 \t

文: 最初はカメラもしょぼかったし、画面も小さかったし、音も3和音とかやった。 \t

評判保持者: [著者] \n [著者] \n [著者] \t

評判表現: しょぼかった \n 画面も小さかった \n 音も3和音とかやった \t

評判タイプ: 批評ー \n メリットー \n メリットー \t

評判対象: カメラ \n 画面 \n 音 \t

「\n」で区切られたフィールドの順番は対応している。つまり、評判表現「しょぼかった」の評判保持者は「[著者]」、評判タイプは「批評ー」、評判対象は「カメラ」である。

参考文献

- [1] 河原大輔, 笹野遼平, 黒橋禎夫, 橋田浩一. 格・省略・共参照タグ付けの基準, 2005. http://nlp.kuee.kyoto-u.ac.jp/nl-resource/corpus/KyotoCorpus4.0/doc/rel_guideline.pdf.
- [2] 黒橋禎夫, 居蔵由衣子, 坂口昌子. 形態素・構文タグ付きコーパス作成の作業基準 version 1.8, 2000. http://nlp.kuee.kyoto-u.ac.jp/nl-resource/corpus/KyotoCorpus4.0/doc/syn_guideline.pdf.