

『日本語歴史コーパス 江戸時代編VI上方絵入狂言本』 Ver.1.0 形態論情報の概要

2025 年 11 月 20 日 久保 柊子

1. はじめに

『日本語歴史コーパス 江戸時代編VI上方絵入狂言本』 Ver.1.0（以下、『日本語歴史コーパス』を CHJ と称する）の形態論情報は、先行して公開された『CHJ 江戸時代編 I 洒落本』『CHJ II 人情本』『CHJ III 近松浄瑠璃』『CHJ IV 随筆・紀行』『CHJ V 注釈書』と同じ規程によっている。単位規程は原則的に「『日本語歴史コーパス 江戸時代編 I 洒落本』『日本語歴史コーパス 江戸時代編 II 人情本』形態論情報の概要」（村山 2019）、「『日本語歴史コーパス 江戸時代編 III 近松浄瑠璃』形態論情報の概要」（片山 2020）に準拠し、本文書では、『CHJ 江戸時代編VI上方絵入狂言本』 Ver.1.0 に独自な点について述べる。

2. 短単位データの作成

『CHJ 江戸時代編VI上方絵入狂言本』の短単位データの作成は、他の CHJ のサブコーパスと同様に自動形態素解析と人手修正によって行われている。

上方絵入狂言本は、主に舞台上の動きを記録するト書きと、歌舞伎役者による台詞によって構成されている。前者は文語体、後者は口語体によっており、それぞれの文体に対応するため、以下のように形態素解析辞書を区別している。

文の種別	解析辞書
ト書き部分	近世文語 UniDic
台詞部分	近世上方口語 UniDic

3. 活用型

『CHJ 江戸時代編 I 洒落本』では「会話」と「割書き」の本文種別、『CHJ 江戸時代編 V 注釈書』では「宣長訳」と「地の文・古今集」の本文種別によって「口語」と「文語」の区別を行っている。本コーパスでも、この本文種別による区別を踏襲し、原則として、ト書き部分を「文語」、台詞部分を「口語」とみなして処理を行なった。

台詞部分においても、文語活用でしか対応できないものについては、本文の形を優先して形態論情報の付与を行っている。

【例】

- (1) 七ちん万宝はくらにみち何にふそくもなふ。かたをならぶるものもない。
 …文語下二段-バ行 連体形-一般 (51-うか 1683_00101, 2320)
- (2) 誠にたがひにおやたちがいひ名付しておかしやつたれども
 …文語助動詞-タリ-完了 已然形-一般 (51-好色 1693_00111, 27220)
- (3) こふした事は人のしらぬ事なれば。くるしからぬ
 …文語形容詞-シク 未然形-補助 (51-白粉 1696_00101, 36190)
- (4) 松五郎様にはあす都へお上りなさるゝゆへおがみまする事はかなはぬ
 …文語下二段-ラ行 連体形-一般 (51-代々 1698_00101, 10560)

4. 指定辞の判別

指定辞について、『CHJ 室町時代編 I 狂言』では「なり」「じゃ」を、『CHJ 明治・大正編』では「だ」「じゃ」をそれぞれ併用しており、洒落本・人情本コーパスでは、指定の助動詞として「なり」「じゃ」「だ」の3種類をみとめ、処理を行っている。本コーパスでは、洒落本・人情本コーパスに従って上記の3種類を採用し、本文種別によって以下の使い分けを行った。

ト書き部分→「なり」
 台詞部分→「じゃ」「だ」

台詞部分であっても、ナリ活用で処理することが妥当な形式については、助動詞「なり」として処理する点、仮定条件を表すもの（「ナラ」、「ナラバ」）については便宜上すべて助動詞「だ」に倒す点についても洒落本・人情本（村山 2019）に従った。

5. 欠損箇所

本コーパスには、底本の状態により文字の判別が困難な箇所が多々含まれる。このような欠損字は「□」で表した。また、短単位の分割位置の推測が可能な箇所は一短単位に分割したのち、品詞「欠損」を付与している。

【例】

- (5) 兩人共にく□□□承り。弟めが所は私の手にかけてせう。
 (51-白粉 1696_00102, 44310)
- (6) いへ私はひ□り身でござりまする (51-代々 1698_00101, 6910)
- (7) ひとりならばさびしうござりませふが是はどふや□□□ひ事
 (51-代々 1698_00101, 7450)

【参考文献】

- 村山実和子（2019）『『日本語歴史コーパス江戸時代編Ⅰ洒落本・Ⅱ人情本』形態論情報の概要』 <https://clrd.ninjal.ac.jp/chj/doc/morph-edo-2019.pdf>
- 片山久留美（2020）『『日本語歴史コーパス江戸時代編Ⅲ近松浄瑠璃』形態論情報の概要』 <https://clrd.ninjal.ac.jp/chj/doc/morph-chikamatsu-2020.pdf>
- 松崎安子（2023）『日本語歴史コーパス江戸時代編Ⅳ随筆・紀行』Ver.0.8 形態論情報の概要 <https://clrd.ninjal.ac.jp/chj/doc/morph-zuihitsu-202303.pdf>
- 久保栞子（2025）『『日本語歴史コーパス 江戸時代編Ⅴ注釈書』形態論情報の概要』 <https://clrd.ninjal.ac.jp/chj/doc/morph-chushaku-202503.pdf>