

『日本語歴史コーパス江戸時代編 I 洒落本』形態論情報の概要

2018年3月30日 村山実和子

はじめに

本コーパスにおける形態論情報は、検索の便宜・処理上の便宜等を考慮して付与されているため、学術上の通説、あるいは既存の索引類等と異なる尺度で付されたものも存在し、必ずしも「学術的な正しさ」を企図して付与されたものではない。そのため、場合によっては目的の語がヒットしなかったり、利用者各位の研究目的とは合致しない分類がなされていたりするおそれがある。また「で」（助詞／助動詞）「又」（副詞／接続詞）など、時に品詞分類等が困難なケースも存在する。研究利用に当たっては、この点を留意の上、目的のものはすべて表示されているかどうか、また付与された情報が研究目的に適うものかどうか、文字列検索・語彙素検索の結果と照合したり、各位において再分類したりする等、多角的に確認・検討することを推奨する。

1. 言語単位

『日本語歴史コーパス (CHJ)』では、用例収集を目的とした「**短単位**」・言語的特徴の解明を目的とした「**長単位**」の2種類の言語単位を採用している。これは、『現代日本語書き言葉均衡コーパス (BCCWJ)』で採用した単位を基に設計したものである。基となっているBCCWJの言語単位は『日本語話し言葉コーパス (CSJ)』との互換性の保持を図り、国立国語研究所が行った語彙調査の単位を基に設計された。これまでに国立国語研究所が実施してきた語彙調査における言語単位のうち、短い単位の系列に属するものが「短単位」、長い単位の系列に属するものが「長単位」である。

『日本語歴史コーパス江戸時代編 I 洒落本』の言語単位は、通時的な日本語研究で利用するために、現代語や別の時代のコーパスとの互換性の保持を図っている（ただし**2018年3月時点では、短単位データのみ公開**）。その一方で、BCCWJや、他の歴史コーパスの規程をそのまま用いるのではなく、本コーパス用に単位認定規程の修正・拡張を行った。短単位には、代表形（語彙素読み）・代表表記（語彙素）・品詞・活用型・活用形を与える。代表形は国語辞典の見出しに、代表表記はその見出しに与えられた漢字等の表記に相当するものである。

2. 短単位の概要

短単位は、言語の形態的側面に着目して規定した言語単位である。短単位の認定にあたっては、まず意味を持つ最小の単位（最小単位）を規定し、その最小単位を文節の範囲内

で短単位認定規程に基づいて結合させる（もしくは結合させない）ことで認定する。

2.1 最小単位

●最小単位は現代語において意味を持つ最小の単位である。本コーパスにおける最小単位については、現代語との関連を重視して、原則として現代語を対象とした最小単位認定を行うが、必要に応じて、使用実態や室町時代編や明治・大正編の状況に基づき個別の判断をすることがある。語種等により、次のように認定する。

※「/」は最小単位の分割位置を表す。

和語：/お/そろひ/なされ/て/よふ/お/こし/あそばし/まし/た

漢語：/雪/月/花/ /三/味/線/ /奉/公/人/ /一/蓮/託/生/

外来語：/きせる/ /しやぼん/ /旦那/ /びろうど/

記号：/。/ / / \ /

人名：/宝井/其角/ /歌川/国貞/ /白/楽天/ /わる井/しあん/

地名：/い勢/の/津/ /馬喰/町/

●上記のように認定した最小単位を、短単位認定のために下表のとおり分類する。

表1 最小単位の分類

分類		例
一般		和語：花 ほど 賑やか 面白い 笑う…
		漢語：大 尽 勿 体 遠慮
		外来語：たばこ しやぼん 襦袢
付属要素		接頭的要素：御（お、ご、み） 大（おお） 不（ぶ） …
		接尾的要素：様（さま） 気（げ） がましい 兼ねる …
その他	記号	， 。 「 」 / \ …
	数	一 二 十 百 千…幾 数 何…
	固有名	人名：山東 京伝 海老蔵 おゆき…
		地名：大坂 日本橋 しまばら 阿蘭陀…
助詞・助動詞	の を が は こそ まで る・らる ず ごとし やる なり だ じゃ…	

2.2 短単位

- 短単位データの作成は自動形態素解析と人手修正によって行われている。形態素解析処理は形態素解析器に「MeCab」、解析用辞書に「近世口語 UniDic」を使用している。
- 短単位の認定規定は、上表の分類ごとに適用すべき規定が定められる。その規定に基づき、最小単位を結合させる（又は結合させない）ことによって、短単位を認定する。以下、「一般」・「数」・「その他」に分けて、短単位認定規定の概要を示す。「|」は短単位

の分割位置を、「=」は短単位を切らないことを示す。

[1] 一般

《和語・漢語》

最小単位2つの結合までを1短単位とする。

【例】 | 親=子 | | 差し=支え | | 顔=見世 | | 悪=態 | | 千=変 | 万=化 |
| 青=楼 | 客 |

例外：切る位置が明確でないもの、あるいは切った場合と一まとめにした場合とで意味にずれがあるものは、3最小単位以上の結合であっても1短単位とする。

【例】 | 御の字 | | 身の程 | | 棒手振り |

例外：最小単位が3つ以上並列した場合、それぞれの最小単位を1短単位とする。

【例】 | 雪 | 月 | 花 | | 土 | 農 | 工 | 商 |

《外来語》

1 最小単位を1短単位とする。

【例】 | たばこ | 盆 | | はだ | じゅばん |

[2] 数

「数」以外の最小単位と結合させない。「数」どうしの結合は、一・十・百・千の桁ごとに1短単位とする。「万」「億」等は、単独で1短単位とする。

【例】 | 卅 | 八 | 匁 | 五 | 分 | | 廿 | 七八 | 日 | | 千 | 遍 | 万 | 遍 |

[3] その他

1 最小単位を1短単位とする。

付属要素 | 御 | かた | | 女郎 | 衆 | | きうくつ | らしい |

助詞・助動詞 | よそ | から | 我身 | に | しら | せ | まし | た | ほど | に |

人名 | 並木 | 五瓶 | | 白 | 楽天 | | 紀国屋 | 文右丞門 |

地名 | 下総 | の | くに | | あたご | 山 |

3. 他のコーパスと異なる処理・特殊な処理

本コーパスのデータの作成にあたっては、原則、現代語のコーパスおよび明治・大正編、また室町時代編における処理を踏襲して行った。ただし、古代語から近現代語への過渡的様相を示す近世語の資料においては、現代・中世いずれの規程に拠っても一律に処理できないケースがしばしば現れる。そのため、可能な限り既存の枠組みを尊重しつつ、独自の処理、あるいは特殊な処理を行った箇所がある。以下にはそのうち、全体に関わる特に注意すべきものを挙げる。

3.1 活用型

[1] 文語活用と口語活用

現代語のコーパスおよび『日本語歴史コーパス』では、活用語について「文語」「口語（明示なし）」の二大別を行っている。しかし、洒落本は前述のとおり古代語から近現代語への過渡的様相を示す資料であり、いずれに拠っても処理が困難な事例が現れる。同様に過渡期である『室町時代編』では、品詞・語によって方針を立てたが、『江戸時代編 I 洒落本』では、**本文種別により大別する**ことで対処している。

●原則として、**会話相当の箇所を「口語」活用、それ以外の箇所（地の文、割書等）を「文語」活用とみなし、処理を行なった。**

【例】

本文種別「会話」

(1) |八|よい|から|いふ|通り|。|どふぞ|帯|を|とい|て|ね|て|お|くれ|。|

(箱まくら／動詞-一般，下一段-ナ行) …**口語**で処理

本文種別「割書き」

(2) |よふ|ね|入ら|ぬ|と|帯|を|とき|上着|の|まま|市村|が|ね|て|ある|ところ|へ|は|いる|

(箱まくら／動詞-一般，文語下二段-ナ行) …**文語**で処理

会話内であっても、韻文相当または手紙の読み上げについては、文語活用で処理している。

本文種別「会話-韻文」

(3) |小太夫|狂歌|よむ|かほ|みれ|ば|ちぢ|に|もの|こそ|うれしけれ|我|身|ひとり|の|君|
に|あら|ね|ど|

(原柳巷花語／動詞-非自立可能，文語上一段-マ行，已然形-一般) …**文語**で処理

●また、会話内であっても文語活用でしか対応できないもの、会話以外であっても口語活用でしか対応できないものについては、形を優先して、形態論情報の付与を行っている。

【例】

本文種別「会話」

(4) |いや|も|朝|から|ばん|まで|さけ|の|たゆる|ま|が|ござり|ませ|ぬ|
(聖遊廓／「絶える」動詞-一般，文語下二段-ヤ行，連体形-一般) …**文語**で処理

(5) |情|に|貴賤|なき|物|と|高尾|が|の|こす|詞|も|あり|
(無論里問答／「有る」動詞-非自立可能，文語ラ行変格，終止形-一般) …**文語**で処理

本文種別「割書き」

(6) |色事|は|しら|ぬ|の|か|どれ|をしへ|て|やらふ|と|引よせる|
(箱まくら／「引き寄せる」動詞-一般，下一段-サ行，終止形-一般) …**口語**で処理

本文種別「地の文」

(7) |娘|が|大|恩|を|うけ|し|。|一|ち|礼|を|述べ^{のべ}れば。|

(花街鑑／「述べる」動詞-一般，下一段-バ行，仮定形-一般) …**口語**で処理

●形容詞型活用は，室町時代編の原則を継承し，**基本的に口語活用**とみなした。文語活用でなければ対応できないものを「文語」とした。

【例】

本文種別「会話」

(8) |そんな|無理|な|事|ばかり|。|おつせへす|と|しん|に|哀しく|なり|
いす| (花街鑑／形容詞-一般，形容詞，連用形-一般)

本文種別「地の文」

(9) |両親|の|心|は|。|さぞ|や|悲しく|あんじわづらひ|。|
(花街鑑／形容詞-一般，形容詞，連用形-一般)

本文種別「引用-手紙」

(10) |きて| / \ |かなしき|は|御|名残|に|御座|候|
(玉菊全伝花街鑑／形容詞-一般，**文語形容詞-シク**，連体形-一般)

[2] 品詞「動詞」で活用型「助動詞-」のものについて

●近世語に特有の助動詞・動詞を辞書に登録するにあたり，表2のように処理を行った。これらの活用型の区別については，出自や活用パターンを同じくするものもあるが，形態に着目し，便宜上「-ンス」「-マス」「-イス (エス)」と分けている。

●また，品詞は動詞でありながら，活用型を「助動詞-」とするものがある。これらは，もともと助動詞を含む形式が転じたものと見られ，形式上切り離すことが困難であるため，例外的に活用型「助動詞-」としている。

表2 最小単位の分類

活用型	語彙素	品詞
助動詞-ンス	さしゃんす	助動詞
	しゃんす	助動詞
	やさんす	助動詞
	やんす	助動詞
	さんす	助動詞
	ざんす	助動詞
	さんす	動詞
	あんす	動詞
	おいなんす	動詞
	くだんす	動詞
	ごんす	動詞
	なさんす	動詞
	おしゃんす	動詞
	なんす	動詞
おざんす	動詞	
助動詞-イス	おっせえす	動詞
	なはえす	動詞
助動詞-マス	なます	動詞
	ごます	動詞

[3] 指定辞の判別について

●指定辞について、『室町時代編 I 狂言』では「なり」「じゃ」を、『明治・大正編』では「だ」「じゃ」をそれぞれ併用している。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、処理を行った。

【例】

(11) |さみ|せん|の|うら|を|たたく|きやく|は|。|を|とこ|ぶり|よく|て|も|を|をく|きは|れる|もの|なり|

(箱まくら／「なり」, 文語助動詞-ナリ, 終止形-一般)

(12) |神カミ|を|申マウくだし|よふ|も|ばち|が|あたら|ぬ|事|じゃ|

(原柳巷花語／「じゃ」, 助動詞-ジャ, 終止形-一般)

(13) |物|の|入イつ|て|ある|長持|と|間夫|の|ない|女郎|と|は|ない|もの|だ|

(郭中奇譚／「だ」, 助動詞-ダ, 終止形-一般)

●このうち、「なり」「だ」については、活用形により形式が共通する場合がある（連体形ナ、連用形ニ、連用形-融合ジャ・デ）。そこで本コーパスでは以下のように区別している。

==

会話以外 → 助動詞「なり」

会話 → 助動詞「だ」

==

【例】

本文種別「会話」

(14) |また|よそ|の|事|を|聞|た|と|て|。|手がら|そう|に|ふれあるく|
やう|な|。|八木|四郎|じや|ない|。|

(色深狹睡夢／「だ」, 助動詞-ダ, 連体形-一般)

(15) |ほん|に|めつた|に|錠|の|ない|所|に|かくす|物|は|おか|れ|ぬ
|ぞ|へ| (陽台遺編・姁閣秘言／「だ」, 助動詞-ダ, 連用形-ニ)

本文種別「割書き」

(16) |いま|来|た|やう|な|かほ|し|て|ふたり|の|座しき|を|のぞき|

(色深狹睡夢／「なり」, 文語助動詞-ナリ, 連体形-一般)

(17) |男|は|あたつ|たら|めつた|に|くだけ|ぬ|北浜|の|意気|張|。|

(北華通情／「なり」, 文語助動詞-ナリ, 連用形-ニ)

●本コーパスが対象とする洒落本は、宝暦年間（1751-1764）から文政年間（1818-1830）にかけての、江戸・京都・大坂を舞台とする作品群である。時代差・地域差を考慮して、江戸を舞台とする作品のみに「だ」を使用することも検討していたが（市村・村山 2017）、その枠組みに当てはまらない例があること、どの段階で上方作品に「だ」を認めるか、といった線引きが困難であることなどから、最終的には全ての「会話」に対して「だ」の使用に倒す方針に決定した。会話以外には、助動詞「なり」（活用型：文語助動詞-ナリ）をあてており、これは先述の「文語」ベース、「口語」ベースの方針にも対応するものである。

●ただし、本文種別：会話の場合でも、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。

【例】

(18) |わたし|が|様|な|数|なら|ぬ|もの|の| (異本郭中奇譚) …未然形

(19) |あきらめ|られ|し|身|なり|せ|ば| (花街鑑) …連用形

(20) |娼妓|など|ほど|きれい|なる|もの|は|ある|まい| (原柳巷花語)
…連体形

(21) |それ|は|そふ|なれ|ど| (誰か面影) …已然形

● 仮定条件を表すもの（「ナラ」、「ナラバ」）については、助動詞「なり」未然形と見たほうがよいものも含まれる可能性があるが、便宜上すべて助動詞「だ」に倒す。否定を表すもののみ、助動詞「なり」未然形として扱う。

【例】

(22) | なんぼ | ふ孝 | に | なつ | て | も | おまへ | ゆへ | なら | 大事 | ない |
(聖遊廓 / 「だ」, 助動詞-ダ, 仮定形-一般)

(23) | 一 | べん | なら | ず | 二 | へん | 迄 | 。 |
(色深狭睡夢 / 「なり」, 文語助動詞-ナリ-断定, 未然形-一般)

● 出現書字形「に」「で」については、助詞／助動詞の区別が不明瞭なものも含まれる。原則的に小椋他編（2011）の規程に従って判断した。

● 助動詞「じゃ」について、動詞「有る」に連なる「じゃあつた」「じゃある」の形式は、室町時代編の方針を引き継ぎ、助動詞「じゃ」連用形-一般として処理した。

【例】

(24) | 此 | 中 | は | きつい | さけ | じゃ | あつ | た |
(異本郭中奇譚 / 「じゃ」助動詞-ジャ, 連用形-一般)

(25) | 何 | 時 | の | お | 輿 | 入 | じゃ | ある |
(当世嘘之川 / 「じゃ」助動詞-ジャ, 連用形-一般)

3.2 活用形

[1] 終止形・連体形の別

● 3.1〔1〕でも述べたとおり、連体形に相当する形態で文末終止を行う場合がある。このような場合は、文末であっても終止形ではなく連体形とした。

【例】

(26) | むかふ | から | も | ちよき | 一 | そう | こぎくる |
(仕懸文庫 / 「漕ぎ来る」動詞-一般, 文語カ行変格, 連体形-一般)

● 終助詞や助動詞に前接する場合、**終止・連体形の区別が困難なケースが多い**。そこで、形態的に明らかなものはその活用形とし、終止・連体同形の物は、極力平安時代編や小椋他(2011)に合わせ、終止形・連体形いずれかに統一した。

● 助動詞「なり」の連体形「なる」が転じた「な」について、終止法で用いられる場合もあるが、これらは連体形として処理を行った。

【例】

- (27) | 今 | は | 来 | ら | れ | ぬ | そう | な |
(色深狹睡夢／「だ」助動詞・助動詞-ダ・連体形-一般)
- (28) | それ | は | まあ | お | 笑 | 止 | な |
(南遊記／「だ」助動詞・助動詞-ダ・連体形-一般)

[2] 意志推量形

●『室町時代編』では助動詞「う」・助動詞「むず」の語形「うず」を立て、未然形+助動詞「う」・「むず」のように、用言と助動詞を分割していたが、本コーパスでは、現代語のコーパスと同じく、ひとまとまりの「意志推量形」として扱った。活用語尾がア段の仮名で表記されている場合も、実際の発音形はオ段であるとみなし、同様に処理を行った。また、本文種別が会話以外の場合でも、意志推量形については口語活用として対処した。

【例】

- (29) | みんな | 釣 | て | 仕 | 廻 | た | の | だらう |
(深川新話／「だ」助動詞，助動詞-ダ，意志推量形)
- (30) | もし | お | 茶 | づ | け | それ | へ | 上 | ませう |
(郭中奇譚／「ます」助動詞，助動詞-マス，意志推量形)
- (31) | とも | に | 出よう | と | する |
(竊潜妻／本文種別「地の文」／「出る」動詞-一般，下一段-ダ行，意志推量形)

[3] 連用形の用法

●次のような連用形の命令法については、連用形として処理した。

【用例】

- (32) | 一寸 | 耳 | かし | と | さ | さ | や | き |
(当世嘘之川／「貸す」，動詞-一般，五段-サ行，連用形-一般)
- (33) | お | ゆ | き | さ | ん | は | や | ふ | お | いで |
(風流裸人形／「出でる」，動詞-非自立可能，下一段-ダ行，連用形-一般)
- (34) | ど | れ | こ | こ | へ | 来 | な |
(青楼阿蘭陀鏡／「来る」，動詞-非自立可能，カ行変格，連用形-一般)
- (35) | ど | ふ | ぞ | 能 | く | し | て | 。 | や | つ | て | 。 | くん | な | よ |
(南閩雑話／「呉れる」，動詞-非自立可能，下一段-ラ行，連用形-撥音便)

3.3 短単位認定

[1] 助詞融合形

●近世後期口語資料の会話文では、助詞とその前後の語との融合がしばしば見られる。こ

の種の融合形について、BCCWJ では、「元の形に戻さずに、融合している複数の最小単位全体で一最小単位とする」と規定している（小椋他編 2011）。本コーパスでも、この方針に従い、融合形を先行語の異語形とみなして処理を行った。

【例】

(36) |それ|で|も| こはけりや (<恐ければ) |どく味|に|さんじ|ませう|か|
(箱まくら / 「怖い」, 形容詞, 形容詞-一般, **仮定形-融合**)

(37) |是|さ|用|が| あらあ (<有るわ) |な|
(甲駅新話 / 「有る」, 動詞-非自立可能, 五段-ラ行, **終止形-融合**)

●活用語の場合は、例のように活用形に「-融合」と表示されるため、融合形だけを抽出することも可能である。ただし、洒落本のテキストは非活用語の融合形も数多く含んでいる。先行語にとっては、異語形も含めて一括検索できるメリットがある一方、後接する助詞は検索対象に含まれない点に注意が必要である。

(38) |そりやあ (<其れは) |ほんとう|の|事|か| (花街鑑)

(39) |でへつ|で|も|相手|だ|。|とんじやかあ (<頓着は) |ねへ| (俠者方言)

(40) |けてう (<けちを) |つける|こたあ (<事は) |ねへ| (同上)

(41) |らちやあ (<埒が) |あか|ねへ| (甲駅新話)

(42) |てめえ|の|とけえ (<所へ) |源四郎|と|いふ|者|が|来|たろふ|が|
(当世左様候)

[2] 終助詞の単位認定

●終助詞について、辞書の立項等に拠らず、基本的に分割して処理している。

【例】

(43) |今宵|は|。|いかふ|さむい|わ|い|な|
(月花余情 / 「わ」「い」「な」 全て助詞-終助詞)

(44) |すま|ぬ|とて|どふ|せう|ぞ|い|の|
(風流裸人形 / 「ぞ」「い」「の」 全て助詞-終助詞)

(45) |又|ね|なんす|か|い|な| (新月花余情 / 「か」「い」「な」 全て助詞-終助詞)

[3] 助動詞「やる」

●動詞連用形に後続する「やる」（「や」）は、「室町時代編 I 狂言」の方針を引き継ぎ、原則助動詞「やる」とする。

【例】


(46) |情|じひ|を|専|に|し|やれ| (原柳巷花語)

(47) |是|みんな|だまり|や| (俠者方言)

3.6 未知語の種別

- 解釈が不明な箇所や単語情報付与が困難な箇所等は「未知語」として扱い、その種別を「品詞」欄に表示する。なお文字列検索は可能である。

表 2 未知語の種別 (品詞)

種別 (品詞)	内容	例
解釈不明	語意、また語境界が不明の場合など形態論情報の付与が困難なものについて、この品詞を付与している。ただし、意味は不詳でも、読みと品詞が明確な場合は、短単位認定した。そのほか、誤字と思われる箇所について『洒落本大成』の翻刻を尊重し、「解釈不明」としたところがある。	ちくてつぼう びりちく はやああい つちやしない 天重 はしり出じ (「走り出し」か) せんだい (「全体」か)
漢文	漢文風の箇所など、読み下しが困難な場合や形態論情報を付与しがたい場合は、まとめて「漢文」とした。また、置き字についても「漢文」とした。	百行衆芸備於身百行之者也 古之言通者 於 而
洒落断片	洒落や掛詞等による臨時的な形態。特に、短単位認定が困難な場合や、語形変化が著しい場合をこれにあてて。	びくまれ くち (「魚籠」と「憎まれ口」を掛ける) くがい くがい で 十 八 がい (苦界を洒落たもの) にいる の つら へ 水 (「～に居る」と「蛙」を掛ける)
歌・呪文 ほか	かけ声や、囃しことばなどで、短単位認定等が困難なものにつき、まとめて「歌・呪文ほか」とした。	つんてれれつとんろん かんのうきうのれんす とんちりちんてんちやんほらほ
絵文字・ 記号等	中納言の検索画面では「☐」で表示される。必要に応じて、『洒落本大成』の該当ページまた外部リンクの参照画像等を確認されたい。	 (国立国語研究所蔵『仕懸文庫』21 オ)
欠損	中納言の検索画面では「↔」で表示される。	省略

参考文献

- 市村太郎(2014)「近世口語資料のコーパス化—狂言・洒落本のコーパス化の過程と課題—」『日本語学 11 月臨時増刊号 日本語史研究と歴史コーパス』33-14 明治書院
- 市村太郎 (2015)「ひまわり版「洒落本コーパス 0.5」利用案内」
http://pj.ninjal.ac.jp/corpus_center/chj/doc/sharebon0.5-doc.pdf (2018年3月28日確認)
- 市村太郎 (2016)「『江戸時代編』の構築と課題」『日本語学会 2016 年度春季大会予稿集』日本語学会
- 市村太郎・小木曾智信 (2016)「文書構造を利用した近世期洒落本の形態素解析」『言語処理学会第 22 回 年次大会発表論文集』言語処理学会
- 市村太郎・河瀬彰宏・小木曾智信(2012)「近世口語テキストの構造化とその課題」『情報処理学会研究報告 人文科学とコンピュータ研究会報告』2012-1
- 市村太郎・村山実和子 (2017)「洒落本コーパス構築の試行」『国立国語研究所論集』12
- 小木曾智信 (2016)『『日本語歴史コーパス』の現状と展望』『国語と国文学』93-5
- 小木曾智信 (2017)「多重の読みを持つテキストのコーパス化」『言語資源活用ワークショップ 2017 発表 論文集』国立国語研究所
- 小木曾智信・市村太郎・鴻野知曉 (2013)「近世口語資料の形態素解析の試み」『第 4 回コーパス日本語学 ワークショップ予稿集』国立国語研究所
- 小椋秀樹・小磯花絵・富士池優美・宮内佐夜香・小西光・原裕 (2011)『『現代日本語書き言葉均衡コーパス』形態論情報規程集 第 4 版 (上) (下)』国立国語研究所
- 国立国語研究所コーパス開発センター(市村太郎ほか)編(2015)『ひまわり版「洒落本コーパス」(日本語歴史コーパス江戸時代編)』http://pj.ninjal.ac.jp/corpus_center/chj/edo.html#share (Ver. 0.5)(2018年3月23日確認)
- 村山実和子・小木曾智信・中村壮範 (2017)「形態論情報の多重化による洒落本コーパスの質的拡張」『第 114 回 CH 研究発表会 論文集』言語処理学会
- 国立国語研究所コーパス開発センター (市村太郎ほか) 編 (2015)『ひまわり版「洒落本コーパス」(日本語歴史コーパス江戸時代編)』
http://pj.ninjal.ac.jp/corpus_center/chj/edo.html#share (Ver. 0.5) (2018年3月20日確認)