

国立国語研究所学術情報リポジトリ

アニメ・ゲームのスク립ト分析における形態素解析の問題点

メタデータ	言語: Japanese 出版者: 国立国語研究所 公開日: 2024-11-13 キーワード (Ja): アニメの日本語, ゲームの日本語, 形態素解析, 誤解析 キーワード (En): 作成者: 小川, 満梨奈, 山本, 裕子, 岩崎, 理紗 メールアドレス: 所属:
URL	https://doi.org/10.15084/0002000377

アニメ・ゲームのスク립ト分析における形態素解析の問題点

小川 満梨奈 (愛知淑徳大学)

山本 裕子 (愛知淑徳大学)

岩崎 理紗 (愛知淑徳大学 グローバルカルチャーコミュニケーション研究科)

A Study of Issues in Morphological Analysis of Japanese Scripts in Anime and Games

Marina Ogawa (Aichi Shukutoku University)

Hiroko Yamamoto (Aichi Shukutoku University)

Risa Iwasaki (Graduate school of Global Culture and Communication, Aichi Shukutoku University)

要旨

日本語学習者にとってアニメやゲームはリソースの一つであるが、教室で学ぶ日本語とは用いられる語彙が異なる。しかしジャンル別の語彙やその頻度がわかる、学習者にも教師にも活用しやすい語彙リストは公開されていない。そこで、日本語教育で活用できる言語資源として、アニメ・ゲームの語彙リストを作成することにした。アニメやゲームのスク립トはそのまま形態素解析すると誤解析が生じやすい。正確なデータ提供を目指し、まずどこにどの程度誤解析が生じるかを確認するため、アニメ 4 作品、ゲーム 1 作品を対象に形態素解析を行った。その結果、5~10%程度の誤解析が生じることが分かった。そのほとんどがアニメやゲームの語彙の特徴を表すものであった。本発表では、語彙リスト作成に向けて行った形態素解析の手順と誤解析の内容を整理し、アニメやゲームの特徴を可能な限り残した形で整理し、リスト化する方法を検討する。

1. はじめに

日本語教育の世界でも IT 化が進み、形態素解析技術は、日本語で書かれた文章の分析には不可欠な存在になりつつある。そのため、日本語研究だけでなく、日本語教育分野でも便利な技術として活用されている。多くの言語資源がコーパスとして公開され、研究や教育に活用されている。ところが、アニメ・マンガに関しては、日本語教育のみならず母語話者にとっても日本語のインプットの大きなリソースの一つであるにもかかわらず、分析例はいくつかあっても (山本・小川 2022) 公開された言語資源はほとんどないと言ってよい。

スク립トの収集が難しいことに加え、形態素解析においても問題となるものが多く、適切な解析がしにくいことも収集、公開が進まない要因の一つだと考えられる。

発表者らはアニメ・ゲームのスク립トを電子情報化し、MeCab-UniDic を用いてデータ分析を行い、アニメ・ゲームの語彙データベースの作成を目指している。日本語教育に即した形で可能な限り正確なデータを提供するには、誤解析への十分な対応が不可欠である。本稿では、データベースを構築する準備段階として、アニメやゲームのスク립ト解析における問題点を調査・分析した。分析の概要や問題点については後述する。

2. 先行研究の検討と本研究の課題

2.1 形態素解析によって生じる誤解析

形態素解析辞書 UniDic を用いた形態素解析の誤解析については、さまざまな指摘がされてきた。小木曾他 (2010) は、新聞・文学作品・ブログの3ジャンルを対象とし、データ解析を行った。その結果、新聞では表記や語法が統制されているため、誤り率が少なかったのに対し、文学作品やブログでは多様な表現が用いられているため、新聞に比べて誤り率が高い傾向にあったが、UniDic の解析精度はいずれのテキストにおいても 98% 近く以上の精度であったと述べている。廣川 (2014) は、解析精度を中心に扱ったものではないが、電話会話での話し言葉の転記テキストを形態素解析し、その際に生じた誤解析の修正内容を「分節誤り」「品詞誤り」「同音異義語あるいは同字異義語」「人名以外の固有名詞」「その他独特な語彙」「人名・あだ名」の6つのカテゴリーに分けている。

また、誤解析が生じやすい箇所については、中村他 (2008) が、助詞と助動詞の間に誤解析が生じやすいことを指摘している。また、渡邊他 (2022) は、CEJC に収録されている日常会話 (人手修正済みデータ) を対象に、UniDic でどのような語 (品詞) で誤解析が生じるのか分析を行った。その結果、「記号」「代名詞」「接続詞」に誤解析が生じやすいこと、「名詞-固有名詞-人名」と「名詞-固有名詞-一般」は日常会話に頻出するが、UniDic においては未知語として扱われるため、これらの語が本来のカテゴリーではなく、他のカテゴリーに誤解析されやすい傾向が確認されたと述べている。山本他 (2023) は、表記が一般的な場合、それが原因となって単語の切り損ないが生じ、誤解析となること、また、形態素解析は文脈から判断しないため読み情報や品詞情報の誤りも生じることを指摘した。

このように、形態素解析辞書 UniDic は解析精度は高いが、辞書に掲載されていない未知語や、表記が統制されていない場合に、誤解析が生じやすい傾向があることがわかる。

2.2 アニメの語彙的特徴

山本・小川 (2022) では、アニメの言葉の特徴として、名詞や動詞にそのアニメ特有の語が多く出現すること、またキャラクターの特徴を表すために実際の話し言葉に用いられる以上の頻度で発音の変化 (以下、「音変化」) が生じたり、略語のようなくだけた言葉や方言が用いられたりすることも多いと述べられている。

2.3 本研究での課題

以上を踏まえ、本稿では、ジャンルの異なるアニメ・ゲームを取り上げ、形態素解析を行い、アニメ・ゲームの持つ特徴を考慮しながら、誤解析の生じやすい箇所とその原因について整理する。本稿での具体的な研究課題は以下の2点とする。

課題1 アニメ・ゲームのスク립トで誤解析が生じやすいのはどのような部分か。

課題2 課題1について、誤解析が生じやすい原因は何か。

3. 形態素解析の概要

3.1 対象作品

本研究では、ジャンルの異なる¹アニメ 4 作品、『推しの子』『呪術廻戦』『五等分の花嫁』『弱虫ペダル』と、ゲーム 1 作品『東京放課後サモナーズ』を対象に、スクリプトを電子情報化し形態素解析を行った。作品の情報は表 1 にまとめて示す。

表 1 分析対象アニメ・ゲーム

作品名	種類	ジャンル	話数	延べ語数 ²
推しの子	アニメ	ファンタジー、サスペンス	1 話 ³	10,739
呪術廻戦	アニメ	バトル、ファンタジー	3 話	7,047
五等分の花嫁	アニメ	ラブコメ、青春	3 話	6,941
弱虫ペダル	アニメ	スポーツ、青春	3 話	6,481
東京放課後サモナーズ	ゲーム	恋愛、バトル	3 話	11,422

3.2 方法

ここでは、分析の手順と使用ツールについて述べる。

まず、分析対象であるアニメとゲームのスクリプトを電子情報化した。ゲームについては、読みと表記が異なる語彙については、フリガナに従って処理を行った。たとえば、例 1 のように『東京放課後サモナーズ』における役職「召喚主」には「サモナー」とフリガナが振ってあった。それは「サモナー」として解析し、解析結果の表層形には「サモナー（召喚主）」と表記した。ただし、例 2 のように直示表現が関係しているものについては、どちらが主要な意味であるとも判断できないため、読みと表記のどちらも解析を行った。

例 1：その人こそが主様の^{サモナー}召喚主^ナなのです！（東京放課後サモナーズ/サロモン）⁴
 （誤）名詞「サモナー」（未知語）→（正）固有名詞「サモナー（召喚主）」

例 2：偽りの名を以て、閉鎖領域^クに顕現すべし！（東京放課後サモナーズ/サロモン）
 （正）読み：代名詞「此処<閉鎖領域>⁵」
 （正）表記：普通名詞「バトル（閉鎖領域）」+「エリア（閉鎖領域）」

電子情報化したスクリプトは、Web 茶まめを用いて形態素解析を行った。ゲームのスクリプトには句読点を使用されていたが、アニメのスクリプトにはないため、解析精度を比較するため、形態素解析前に統一して句読点は削除した。また、Web 茶まめでは、半角数字は正しく解析されないため、解析前処理として「半角→全角変換」と「数字処理」を行った。

¹ 「d アニメストア」という Web サイトの分類基準と各アニメ作品の公式サイトを参考にジャンルを区別した。

² 「語数」は「短単位数」を表す。ここでは、Web 茶まめによる形態素解析後に誤解析を手作業で修正した総語数を「延べ語数」とする。

³ 『推しの子』が 1 話のみの対象なのは、1 話が約 1 時間 20 分あり、アニメ約 3 話分以上に相当するからである。

⁴ （ ）内は作品名/登場人物名を表す。以下、同様。

⁵ 本研究では、<>で括られた言葉は<>内も解析対象であることを表す。

このようにして Web 茶まめによる形態素解析を行い、得られた結果を手作業で適宜修正した。その後、誤解析が生じた原因に注目して誤解析の分類を行った。

4. 結果

Web 茶まめを用いてアニメ・ゲームの形態素解析した結果、生じた誤解析の割合を以下に示す。

表 2 各作品の誤解析の割合

作品名	延べ語数	誤解析率
推しの子	10,739	5.46%
呪術廻戦	7,047	7.16%
五等分の花嫁	6,941	7.87%
弱虫ペダル	6,481	7.09%
東京放課後サモナーズ	11,422	9.75%

いずれの作品も 5%を超えており、小木曾他 (2010) で言及されている UniDic の解析精度約 98%と比較すると、誤解析率は高い傾向にある。

これらの誤解析は主に、山本・小川 (2022) で指摘されているアニメ・ゲームの語彙の特徴である「各作品特有の用語」、「音変化」、「表記」の 3 つの原因から生じたと考えられる。以下では、この 3 つの原因をもとに誤解析を分類し、実例を示す。

4.1 原因 1：アニメ・ゲームにおける特有の用語

アニメやゲームでは、作品ごとの特有の用語が多く出現する。その特有の語彙は、「固有名詞（特に人名）」、「アニメ・ゲームの特殊語（技名など）」、「感動詞」に現れやすい。これら特有の用語は形態素解析用辞書には存在しない未知語のため、正しく解析されず、誤解析が生じてしまう。また、キャラクターを特徴づけるために、形態素解析用辞書にはない、話し言葉、方言や若者言葉などが使用される例もある。

4.1.1 誤解析 1：固有名詞（特に人名）

固有名詞、特にアニメ・ゲームの登場人物の名前は、接尾辞「さん」や「君」とセットであれば誤解析は少なくなるが、一般的ではない名前がつけられている場合、読みや区切り⁶の誤りが生じやすい。たとえば、『呪術廻戦』の「伏黒恵」は「伏黒」が「伏・黒」と分割して解析される。

また、アニメ・ゲームの登場人物については、作品に関連が深い一般的な名詞が名前として用いられることも多く、そうした場合、例 3 のように固有名詞（人名）ではなく、普通名詞として解析されることがある。

例 3：小野田坂道君と何か勝負するわけ？（弱虫ペダル/寒咲幹）

「坂道」：（誤）普通名詞「坂道」→（正）固有名詞（人名）「サカミチ」

⁶ 「区切り」とは、「適切に区切ることができないという解析ミス」という意味である。以下、同様の意味で「区切り」と記載する。

4.1.2 誤解析2：アニメ・ゲームの特殊語

アニメやゲームでは、作品ごとに特殊語が出てくることが多い。特に、異世界設定やバトルシーンがあるアニメ・ゲームでは、能力名や技名などが特殊な用語である場合が多い。それらは、形態素解析用辞書には存在しない語のため、正しく解析されず、主に区切りによる誤解析が生じる。たとえば、『呪術廻戦』には、「呪術」に関する語として、「呪霊」「呪骸」「呪胎」などが作品固有の語として用いられるが、いずれも一語とは見なされず、「呪・霊」「呪・骸」「呪・胎」と区切られて解析される。

4.1.3 誤解析3：感動詞

アニメ・ゲームの特徴として、感動詞の出現頻度が高く、バリエーションもさまざまであることが挙げられる。正しく解析される部分もある一方で、形態素解析用辞書に存在しない感動詞も多く見られた。これらは、区切りよりも「名詞」として未知語扱いされることが多い。以下は「名詞」として未知語扱いされた例である。

例4：グガ、ガアアアアアアアアッ、ヌハハハ、ワオ、ンギャー など

感動詞は次節で述べる音変化にも影響を受けやすく、それにより区切りの誤りが生じる。たとえば、「えっへ」が「えっへへ」になることにより、「えっ」と「へへ」に区切って解析される誤りである。以下に、「えっへへ」以外に同様の誤解析が生じた例を挙げる。

例5：アッハハ、えっへん、しゃー⁷ など

特にゲームでは、アニメのように字幕にかかる文字数制限がなく、感動詞の音の長さに制限がない。それにより短単位境界の誤解析が多くなる傾向も見られた。これは、感動詞の音変化および表記による区切りの誤りに大きな影響を与えた。

例6：うわああああああああっ。⁸（東京放課後サモナーズ/リョウタ）

また、感動詞が副詞として誤解析されることも多い。特に「ふっ」「ぐっ」の誤解析の頻度が高い。ここでは一語で「感動詞」として使われているため、感動詞として処理した。

例7：フッ 唯一の特技で負けるのが怖いかな？（五等分の花嫁/上杉風太郎）

「フッ」：（誤）副詞「ふっ」→（正）感動詞「ふっ」

4.1.4 誤解析4：キャラクター設定による誤解析

アニメ・ゲームではそれぞれのキャラクターの特徴を明確にするために、方言や略語、若者言葉を使用する登場人物も出てくる。しかし、方言や略語、若者言葉には形態素解析用辞

⁷ 「よっしゃー」の音変化。

⁸ 解析の際はスクリプト通りに句読点は含めなかったが、ここでは読みやすさのためにスクリプト上スペースのある部分に読点を挿入して示す。

書に掲載されていない語も多いため、誤解析が生じる。ここでは、関西出身設定の登場人物が、方言を使用したことによる誤解析の例を紹介する。「地味すぎだ。」という意味の「地味すぎや。」の「や」は、本来助動詞の「だ」の意味になるが、副助詞として解析された。また、「ダメだ」という意味の「あかん」は「あ・かん」と区切られ、それぞれ感動詞、代名詞と解析された。例8に示す。

例8：地味すぎや、あかんで。(弱虫ペダル/鳴子章吉)

地味すぎや、

「や」：(誤) 副助詞「や」→(正) 助動詞「だ」

あかんで。

「あかん」：(誤) フィラー「あー」+代名詞「彼」→(正) 動詞「明く」+助動詞「ず」

4.2 原因2：音変化

アニメ・ゲームではそれぞれのキャラクターの特徴を明確にするために、音変化が生じることも多い。ここでは、音変化が原因による品詞の誤りと区切りの誤りの例を見ていく。

4.2.1 誤解析5：音変化による品詞の誤り

先述したとおり、形態素解析において誤りが生じやすい箇所の指摘はこれまでもされているが、アニメはそれらに加え、音変化が原因により品詞の誤りが生じることがある。たとえば、「ない」が「ねえ」に変化すると、助動詞または形容詞が終助詞として誤解析されることがある。ただし、この助動詞や形容詞の「ない」の音変化「ねえ」には、解析の揺れがある。例9は、「ない」の音変化「ねえ」の正しく解析された例と誤解析が生じた例である。「ない」の音変化「ねえ」が2度出現しているが、「冗談じゃねえ」の「ねえ」は正しく形容詞「ない」として解析されていた。一方で、「並んでねえよ」の「ねえ」は後ろに終助詞がついているにもかかわらず、助動詞「ない」が終助詞「ね」に解析された。例10は終助詞がつかず、助動詞「ない」が終助詞「ね」に解析された例である。

例9：冗談じゃねえ。並んでねえよ。(弱虫ペダル/今泉俊輔)

冗談じゃねえ。

「ねえ」：(正) 形容詞「ない」

並んでねえよ。

「ねえ」：(誤) 終助詞「ね」→(正) 助動詞「ない」

例10：呪いはほっとけねえ。(呪術廻戦/虎杖悠仁)

「ねえ」：(誤) 終助詞「ね」→(正) 助動詞「ない」

このように、音変化には解析の揺れがある。また、音変化により、未知語扱いされたものもあった。例11は、形容詞「痛い」が「イッテ」と音変化したため、一語として扱われ、名詞として未知語に解析された例である。ただし、この場合、カタカナ表記であることも誤解析の原因となっている可能性があることも留意しておきたい。

例 11：うぐっ…。イッテ（五等分の花嫁/上杉風太郎）

「イッテ」：(誤) 名詞「イッテ」(未知語) → (正) 形容詞「痛い」

4.2.2 誤解析6：音変化による区切りの誤り

音変化が原因による誤解析には、品詞誤りだけでなく区切りの誤りもある。「そーいや、あれ」の「いや」は「言えば」の音変化であるが、「嫌」と誤って解析される。形容詞には山本・小川（2022）で述べたように、アニメでは[-ai][-oi]が[-ee]となる音変化を中心に、バリエーションが多い。実際に、形容詞の音変化は誤解析が生じやすい。

例 12：キメえんだよ、ヤベえ、いったーい、よっわ など

また音に対して非標準的な表記がされることもアニメ・ゲームの特徴であると言える。たとえば、感情表現や強調のために文末に「っ」がつく、「うわあああー」などと表記される、などの音のバリエーションである。上記の例に加えて、キャラクターを特徴づけるために、例 13、14 のように特定の人物の発話に頻繁に「っ」がつくこともある。このような非標準的な表記は、正しく区切られず、誤解析の一因となる。

例 13：どこですか、主様。こわいですっ！（東京放課後サモナーズ/サロモン）

(誤) ですっ

(正) ですっ

例 14：起きてくださいいっ！（東京放課後サモナーズ/サロモン）

(誤) くださいいっ

(正) くださいいっ

4.2.3 誤解析7：言い淀み

アニメやゲームにおいて、「な…何で」のように登場人物が驚いたり、戸惑ったりしてことばがすぐに出ない時や、「それはちょっ…」のように途中で口ごもるような時、スクリプトには「語」が分断されたり不完全な形で現れる。このような場合を本研究では「言い淀み」とする。言い淀みは適切に解析されず誤解析が生じる。

例 15：そ… そうだが それは…。(五等分の花嫁/上杉風太郎)

「そ」：(誤) 副詞「そう」→「語」ではないため、短単位として認定せず削除

4.3 原因3：表記が関わる誤解析

アニメやゲームのスクリプトでは、キャラクターを特徴づけるために表記のバリエーションが豊かであり、これにより誤解析が生じやすくなる。

4.3.1 誤解析8：表記と読みのずれ

まず、スクリプト上の表記と語彙素の読みの違いに問題があった。スクリプト上の表記に基づいて解析されると、アニメ・ゲームで用いられているものとは異なる「読み」になるものがある。たとえば、スクリプト上では「私」と漢字で表記されるため、語彙素「私」に解

析される。語彙素「私」の読みは「ワタクシ」であるが、実際の音声では「ワタシ」と発音される場合がほとんどであった。また、「主様」の「主」は本来であれば、「アルジ」と読むが、形態素解析では「主・様」と区切られる。「主」の語彙素読みは「ヌシ」であり、この場合は誤解析となる。その他、息継ぎの「ハッ」はひらがな表記であれば感動詞として正しく解析されるが、スクリプト通りのカタカナ表記の場合、数詞の「八」と誤って解析されることが多い。

さらに表記上、音の区別がつかない場合もある。たとえば、「あ、空？」が「ソラ」と誤って解析されたが、実際は「カラ」であった。

このように複数の読みを持つ漢字語がスクリプトに用いられる場合に、誤解析が生じることがあった。

4.3.2 誤解析9：通常の表記と異なる表記による誤解析

本来、漢字かな交じりで書かれるものにカタカナ表記を用いたり、元の言葉とは部分的に異なる表記を用いたりするような場合、誤解析の原因になる。

たとえば、ある人物の名前を、特定の人物が呼ぶ場合にだけカタカナ表記が用いられ、特定の登場人物のセリフのみ、全てカタカナ表記されたりするような場合が該当する。例16のように、本研究で分析対象としたゲームのスクリプトでは、ある登場人物のセリフがすべてカタカナで表記されていたため、そのセリフの多くに誤解析が生じた。

例16：探シテイタ、探シテイタゾッ。(東京放課後サモナーズ/オニ)

探シテイタ、

(誤) 人名「探 (モトム)」+名詞「シテイタ」(未知語)

→ (正) 動詞「探す」+接続助詞「て」+動詞「居る」+助動詞「た」

探シテイタゾッ。

(誤) 人名「探 (モトム)」+名詞「仕手」+地名「イタ」+名詞「ゾッ」(未知語)

→ (正) 動詞「探す」+助詞「て」+動詞「居る」+助動詞「た」+終助詞「ぞ」

これらは特定の人間関係や特定のキャラクターの特徴を際立たせるために、カタカナ表記を用いていると考えられる。

また、アニメやゲームの特殊な語彙と表記の問題が重なって生じる誤解析もある。たとえば、『東京放課後サモナーズ』では、異世界からの「転校生」を「転光生」と呼ぶ。「転校生」に準じて考えると、正しくは「転光・生」の区切りであるが、「転・光・生」と区切られ解析された。同様に、このゲームに登場する学園は新宿にあり、ゲームの内容に基づき「神宿学園」と名付けられているが、「神宿・学園」とは解析されず、「神・宿学・園」と区切られて解析されてしまう。

4.4 その他の問題点：読みが確定しないもの

3つの原因から誤解析の内容を分類して例を見てきたが、上記で挙げた3つの原因以外にも問題点が確認された。読みが確定しないことによる問題である。

ゲームは音声情報がないため、読みが判別できないことがある。たとえば、以下の例17のように、表記は「道標」で、解析では「ドウヒョウ」と読んでいたが、「ミチシルベ」と

も読める。このような例は誤解析だとは言えないため、本研究では Web 茶まめの解析に従ってそのまま「どうひょう」として残した。これは音声情報がないことによって生じる問題である。

例 17：元の世界への道標もなくなって…。(東京放課後サモナーズ/サロモン)

今回は読みが確定しないという問題は、ゲームにのみ見られたが、アニメにも字幕のみが用いられる場合があり、同様の問題が生じる可能性はある。

5. おわりに

本稿では、Web 茶まめを用いてアニメ・ゲームのスク립トの形態素解析をする際に生じた誤解析について、3つの原因を挙げ、それぞれに該当する例を示した。ジャンルを問わず、どのアニメやゲームにおいても、誤解析が生じやすい箇所は同じであった。いずれも、アニメ・ゲームの世界を形作るための表現が影響して起こるものと言えるだろう。

今回得られた問題点を踏まえつつ、アニメやゲームの特徴を可能な限り残すために、現時点では、コーパス分析システム Co-Chu⁹におけるタグ機能を活用した解析を検討している。正確なデータに基づいた語彙リスト作成を進めていきたい。

文 献

小木曾智信・小椋秀樹・小磯花絵・宮内佐夜香・渡部涼子・伝康晴 (2010) 「形態素解析辞書のベンチマークテスト—IPAdic・NAIST-jdic・UniDic のジャンル別精度比較—」『言語処理学会年次大会発表論文集』16, pp.326-329.

(https://www.anlp.jp/proceedings/annual_meeting/2010/pdf_dir/PA1-14.pdf よりダウンロード可能)

堤智昭・小木曾智信 (2023) 「複数の UniDic 辞書による形態素解析支援ツール『Web 茶まめ』の実装と運用」『情報処理学会論文誌』64:3, pp.749-757.

(<http://doi.org/10.20729/00225271> よりダウンロード可能)

伝康晴・小木曾智信・小椋秀樹・山田篤・峯松信明・内元清貴・小磯花絵 (2007) 「コーパス日本語学のための言語資源：形態素解析用電子化辞書の開発とその応用」『日本語科学』22, pp.101-123.

(<http://doi.org/10.15084/00002185> よりダウンロード可能)

中村純平・伝康晴 (2008) 「形態素解析誤りの多い助詞・助動詞の再解析」『言語処理学会年次大会発表論文集』14, pp.326-329.

(https://www.anlp.jp/proceedings/annual_meeting/2008/pdf_dir/C1-2.pdf よりダウンロード可能)

廣川純子 (2014) 「方言に対応した形態素解析辞書の拡張—『広島大学電話会話コーパス』構築に際して—」『ニダバ』43, pp.31-39.

(<https://ir.lib.hiroshima-u.ac.jp/00051005> よりダウンロード可能)

山本裕子・小川満梨奈 (2022) 「コーパス分析システムを用いた「アニメの日本語」の語彙的・文法的特徴の抽出の試み」『愛知淑徳大学論集-グローバルカルチャー・コミュニ

⁹ Co-Chu については山本他(2022)を参照のこと。

ケーション研究科篇-』14, pp.43-56.

(<https://aska-r.repo.nii.ac.jp/records/8519> よりダウンロード可能)

山本裕子・本間妙・Matthew Lanigan・川村よし子・小森早江子 (2022) 「コーパス分析システムの公開と日本語教育への活用」『ヨーロッパ日本語教育』25, pp.238-248.

(<https://cochu.org/wp/wp-content/uploads/2022/07/AJE2021.pdf> よりダウンロード可能)

山本裕子・川村よし子 (2023) 「形態素解析技術と日本語教育—形態素解析の問題点をどう克服するか—」『日本語教育支援システム研究会第10回国際研究集会予稿集』pp.111-115.

渡邊友香・西川賢哉 (2022) 「『日本語日常会話コーパス』での形態素解析：誤解析箇所分析」『言語資源活用ワークショップ発表論文集』6, pp.226-239.

(<https://repository.ninjal.ac.jp/records/3514> よりダウンロード可能)

関連 URL

形態素解析ツール『Web 茶まめ』	https://chamame.ninjal.ac.jp/
d アニメストア	https://animestore.docomo.ne.jp/animestore/tp_pc
推しの子	https://ichigoproduction.com/Season1/
呪術廻戦	https://jujutsukaisen.jp/
五分の花嫁	https://www.tbs.co.jp/anime/5hanayome/1st/
弱虫ペダル	https://yowapeda.com/series/1st/
東京放課後サモナーズ	https://housamo.jp/