

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第6161656号  
(P6161656)

(45) 発行日 平成29年7月12日(2017.7.12)

(24) 登録日 平成29年6月23日(2017.6.23)

(51) Int.Cl.

F 1

G06F 17/22 (2006.01)  
G06F 17/27 (2006.01)G06F 17/22 664  
G06F 17/27 640

請求項の数 13 (全 21 頁)

(21) 出願番号 特願2015-104613 (P2015-104613)  
 (22) 出願日 平成27年5月22日 (2015.5.22)  
 (65) 公開番号 特開2016-218848 (P2016-218848A)  
 (43) 公開日 平成28年12月22日 (2016.12.22)  
 審査請求日 平成27年12月3日 (2015.12.3)

(73) 特許権者 000004226  
 日本電信電話株式会社  
 東京都千代田区大手町一丁目5番1号  
 (74) 代理人 110001519  
 特許業務法人太陽国際特許事務所  
 (72) 発明者 宮崎 千明  
 東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内  
 (72) 発明者 片山 太一  
 東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内  
 (72) 発明者 平野 徹  
 東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】言語表現書き換え装置、方法、及びプログラム

## (57) 【特許請求の範囲】

## 【請求項 1】

文体、述部機能表現、及び人称代名詞を含み、かつキャラクタに応じた特徴が表れる複数種類の言語的特徴の各々について、該言語的特徴に基づく言語表現の書き換えに関する設定値を設定する設定部と、

前記設定部により設定された設定値に基づいて、入力文の末尾の文節に含まれる機能表現を対象として文体の種類を変換する処理、入力文の末尾の文節に含まれる機能表現を対象として述部機能表現をキャラクタに応じて変換する処理、及び人称代名詞である形態素をキャラクタに応じた人称代名詞に変換する処理を含む複数種類の言語的特徴に基づく言語表現の書き換え処理のうち、少なくとも1種類以上の書き換え処理を前記入力文に適用する書き換え処理部と、

を含む言語表現書き換え装置。

## 【請求項 2】

前記設定部は、文構造、活用形、言いよどみ、方言又は特殊語彙、特定の音素、及びキャラクタを弁別可能だが意味を持たない弁別的無意味表現をさらに含む前記複数種類の言語的特徴の各々について、前記設定値を設定し、

前記書き換え処理部は、前記設定部により設定された設定値に基づいて、複文の入力文を単文に分割すると共に、キャラクタに応じた接続表現を分割箇所に挿入することで文構造を変換する処理、活用形を変換する処理、言いよどみの表現に変換する処理、特定の語彙を方言又は特殊語彙へ変換する処理、特定の音素をキャラクタに応じた音素に変換する

処理、及びキャラクタに応じた弁別的無意味表現を挿入する処理をさらに含む前記複数種類の言語的特徴に基づく言語表現の書き換え処理のうち、少なくとも1種類以上の書き換え処理を入力文に適用する

請求項1記載の言語表現書き換え装置。

**【請求項3】**

前記設定部は、文字種、分かち書き、及び記号類をさらに含む前記複数種類の言語的特徴の各々について、前記設定値を設定し、

前記書き換え処理部は、前記設定部により設定された設定値に基づいて、前記文字種を変換する処理、分かち書きに変換する処理、及び記号類を挿入する処理をさらに含む前記複数種類の言語的特徴に基づく言語表現の書き換え処理のうち、少なくとも1種類以上の書き換え処理を入力文に適用する

10

請求項1又は請求項2記載の言語表現書き換え装置。

**【請求項4】**

文体、述部機能表現、及び人称代名詞を含み、かつキャラクタに応じた特徴が表れる複数種類の言語的特徴の各々について、該言語的特徴に基づく言語表現の書き換えに関する設定値を設定する設定部と、

前記設定部により設定された設定値に基づいて、文体の種類を変換する処理、述部機能表現をキャラクタに応じて変換する処理、及び人称代名詞である形態素をキャラクタに応じた人称代名詞に変換する処理を含む複数種類の言語的特徴に基づく言語表現の書き換え処理を入力文に適用する書き換え処理部と、

20

を含む言語表現書き換え装置。

**【請求項5】**

前記設定部は、文構造、活用形、言いよどみ、方言又は特殊語彙、特定の音素、及びキャラクタを弁別可能だが意味を持たない弁別的無意味表現をさらに含む前記複数種類の言語的特徴の各々について、前記設定値を設定し、

前記書き換え処理部は、前記設定部により設定された設定値に基づいて、文構造を変換する処理、活用形を変換する処理、言いよどみの表現に変換する処理、特定の語彙を方言又は特殊語彙へ変換する処理、特定の音素をキャラクタに応じた音素に変換する処理、及びキャラクタに応じた弁別的無意味表現を挿入する処理のうち、少なくとも1種類以上の処理をさらに含む前記複数種類の言語的特徴に基づく言語表現の書き換え処理を入力文に適用する

30

請求項4記載の言語表現書き換え装置。

**【請求項6】**

前記設定部は、文字種、分かち書き、及び記号類をさらに含む前記複数種類の言語的特徴の各々について、前記設定値を設定し、

前記書き換え処理部は、前記設定部により設定された設定値に基づいて、前記文字種を変換する処理、分かち書きに変換する処理、及び記号類を挿入する処理のうち、少なくとも1種類以上の処理をさらに含む前記複数種類の言語的特徴に基づく言語表現の書き換え処理を入力文に適用する

請求項4又は請求項5記載の言語表現書き換え装置。

40

**【請求項7】**

設定部、及び書き換え処理部を含む言語表現書き換え装置における言語表現書き換え方法であって、

前記設定部が、文体、述部機能表現、及び人称代名詞を含み、かつキャラクタに応じた特徴が表れる複数種類の言語的特徴の各々について、該言語的特徴に基づく言語表現の書き換えに関する設定値を設定し、

前記書き換え処理部が、前記設定部により設定された設定値に基づいて、入力文の末尾の文節に含まれる機能表現を対象として文体の種類を変換する処理、入力文の末尾の文節に含まれる機能表現を対象として述部機能表現をキャラクタに応じて変換する処理、及び人称代名詞である形態素をキャラクタに応じた人称代名詞に変換する処理を含む複数種類

50

の言語的特徴に基づく言語表現の書き換え処理のうち、少なくとも1種類以上の書き換え処理を入力文に適用する

言語表現書き換え方法。

**【請求項8】**

前記設定部は、文構造、活用形、言いよどみ、方言又は特殊語彙、特定の音素、及びキャラクタを弁別可能だが意味を持たない弁別的無意味表現をさらに含む前記複数種類の言語的特徴の各々について、前記設定値を設定し、

前記書き換え処理部は、前記設定部により設定された設定値に基づいて、複文の入力文を単文に分割すると共に、キャラクタに応じた接続表現を分割箇所に挿入することで文構造を変換する処理、活用形を変換する処理、言いよどみの表現に変換する処理、特定の語彙を方言又は特殊語彙へ変換する処理、特定の音素をキャラクタに応じた音素に変換する処理、及びキャラクタに応じた弁別的無意味表現を挿入する処理をさらに含む前記複数種類の言語的特徴に基づく言語表現の書き換え処理のうち、少なくとも1種類以上の書き換え処理を入力文に適用する

請求項7記載の言語表現書き換え方法。

**【請求項9】**

前記設定部は、文字種、分かち書き、及び記号類をさらに含む前記複数種類の言語的特徴の各々について、前記設定値を設定し、

前記書き換え処理部は、前記設定部により設定された設定値に基づいて、前記文字種を変換する処理、分かち書きを変換する処理、及び記号類を挿入する処理をさらに含む前記複数種類の言語的特徴に基づく言語表現の書き換え処理のうち、少なくとも1種類以上の書き換え処理を入力文に適用する

請求項7又は請求項8記載の言語表現書き換え方法。

**【請求項10】**

設定部、及び書き換え処理部を含む言語表現書き換え装置における言語表現書き換え方法であって、

前記設定部が、文体、述部機能表現、及び人称代名詞を含み、かつキャラクタに応じた特徴が表れる複数種類の言語的特徴の各々について、該言語的特徴に基づく言語表現の書き換えに関する設定値を設定し、

前記書き換え処理部が、前記設定部により設定された設定値に基づいて、文体の種類を変換する処理、述部機能表現をキャラクタに応じて変換する処理、及び人称代名詞である形態素をキャラクタに応じた人称代名詞に変換する処理を含む複数種類の言語的特徴に基づく言語表現の書き換え処理を入力文に適用する

言語表現書き換え方法。

**【請求項11】**

前記設定部は、文構造、活用形、言いよどみ、方言又は特殊語彙、特定の音素、及びキャラクタを弁別可能だが意味を持たない弁別的無意味表現をさらに含む前記複数種類の言語的特徴の各々について、前記設定値を設定し、

前記書き換え処理部は、前記設定部により設定された設定値に基づいて、文構造を変換する処理、活用形を変換する処理、言いよどみの表現に変換する処理、特定の語彙を方言又は特殊語彙へ変換する処理、特定の音素をキャラクタに応じた音素に変換する処理、及びキャラクタに応じた弁別的無意味表現を挿入する処理のうち、少なくとも1種類以上の処理をさらに含む前記複数種類の言語的特徴に基づく言語表現の書き換え処理を入力文に適用する

請求項10記載の言語表現書き換え方法。

**【請求項12】**

前記設定部は、文字種、分かち書き、及び記号類をさらに含む前記複数種類の言語的特徴の各々について、前記設定値を設定し、

前記書き換え処理部は、前記設定部により設定された設定値に基づいて、前記文字種を変換する処理、分かち書きを変換する処理、及び記号類を挿入する処理のうち、少なくと

10

20

30

40

50

も1種類以上の書き換え処理をさらに含む前記複数種類の言語的特徴に基づく言語表現の書き換え処理を入力文に適用する

請求項10又は請求項11記載の言語表現書き換え方法。

**【請求項13】**

コンピュータを、請求項1～請求項6の何れか1項記載の言語表現書き換え装置の各部として機能させるための言語表現書き換えプログラム。

**【発明の詳細な説明】**

**【技術分野】**

**【0001】**

本発明は、言語表現書き換え装置、学習装置、方法、及びプログラムに係り、特に、入力されたテキストの言語表現を書き換える言語表現書き換え装置、方法、及びプログラムに関する。 10

**【背景技術】**

**【0002】**

従来、文末の機能語列（文末表現）のみを対象としてテキストを書き換えることにより、言語表現にキャラクタ付けを行う方法が存在する。例えば、著者の属性が付与されたテキストデータを用いて、著者の属性値毎に偏って多く使われる文末表現を抽出し、発話のキャラクタ付けに利用する技術が提案されている（非特許文献1）。

**【先行技術文献】**

**【非特許文献】**

**【0003】**

【非特許文献1】宮崎千明、平野徹、東中竜一郎、牧野俊朗、松尾義博、「発話にキャラクタ性を与えるための文末表現の変換」、人工知能学会研究会資料(SIG-SLUD-68), pp. 4 1-46, 2013. 20

**【発明の概要】**

**【発明が解決しようとする課題】**

**【0004】**

しかし、上記非特許文献1の技術では、文末表現のみを書き換えの対象としているため、例えば、テレビアニメや漫画の登場人物のように個性豊かなキャラクタの言語的特徴を、多様なバリエーションで表現するためには十分ではない、という問題がある。 30

**【0005】**

本発明は、上記問題を解決するために成されたものであり、多様なバリエーションの言語表現への書き換えを実現する言語表現書き換え装置、方法、及びプログラムを提供することを目的とする。

**【課題を解決するための手段】**

**【0006】**

上記目的を達成するために、第1の発明に係る言語表現書き換え装置は、文体、述部機能表現、及び人称代名詞を含み、かつキャラクタに応じた特徴が表れる複数種類の言語的特徴の各々について、該言語的特徴に基づく言語表現の書き換えに関する設定値を設定する設定部と、前記設定部により設定された設定値に基づいて、入力文の末尾の文節に含まれる機能表現を対象として文体の種類を変換する処理、入力文の末尾の文節に含まれる機能表現を対象として述部機能表現をキャラクタに応じて変換する処理、及び人称代名詞である形態素をキャラクタに応じた人称代名詞に変換する処理を含む複数種類の言語的特徴に基づく言語表現の書き換え処理のうち、少なくとも1種類以上の書き換え処理を入力文に適用する書き換え処理部と、を含んで構成することができる。 40

**【0007】**

第1の発明に係る言語表現書き換え装置によれば、設定部が、文体、述部機能表現、及び人称代名詞を含み、かつキャラクタに応じた特徴が表れる複数種類の言語的特徴の各々について、言語的特徴に基づく言語表現の書き換えに関する設定値を設定する。そして、書き換え処理部が、設定部により設定された設定値に基づいて、入力文の末尾の文節に含 50

まれる機能表現を対象として文体の種類を変換する処理、入力文の末尾の文節に含まれる機能表現を対象として述部機能表現をキャラクタに応じて変換する処理、及び人称代名詞である形態素をキャラクタに応じた人称代名詞に変換する処理を含む複数種類の言語的特徴に基づく言語表現の書き換え処理のうち、少なくとも1種類以上の書き換え処理を入力文に適用する。

#### 【0008】

また、第2の発明に係る言語表現書き換え方法は、設定部、及び書き換え処理部を含む言語表現書き換え装置における言語表現書き換え方法であって、前記設定部が、文体、述部機能表現、及び人称代名詞を含み、かつキャラクタに応じた特徴が表れる複数種類の言語的特徴の各々について、該言語的特徴に基づく言語表現の書き換えに関する設定値を設定し、前記書き換え処理部が、前記設定部により設定された設定値に基づいて、入力文の末尾の文節に含まれる機能表現を対象として述部機能表現をキャラクタに応じて変換する処理、及び人称代名詞である形態素をキャラクタに応じた人称代名詞に変換する処理を含む複数種類の言語的特徴に基づく言語表現の書き換え処理のうち、少なくとも1種類以上の書き換え処理を入力文に適用する方法である。10

#### 【0009】

このように、文末表現だけでなく、多様な言語表現の書き換え処理を任意に組み合わせて実施するため、多様なバリエーションの言語表現への書き換えを実現することができる。20

#### 【0010】

また、第1及び第2の発明において、前記設定部は、文構造、活用形、言いよどみ、方言又は特殊語彙、特定の音素、及びキャラクタを弁別可能だが意味を持たない弁別的無意味表現をさらに含む前記複数種類の言語的特徴の各々について、前記設定値を設定し、前記書き換え処理部は、前記設定部により設定された設定値に基づいて、複文の入力文を単文に分割すると共に、キャラクタに応じた接続表現を分割箇所に挿入することで文構造を変換する処理、活用形を変換する処理、言いよどみの表現に変換する処理、特定の語彙を方言又は特殊語彙へ変換する処理、特定の音素をキャラクタに応じた音素に変換する処理、及びキャラクタに応じた弁別的無意味表現を挿入する処理をさらに含む前記複数種類の言語的特徴に基づく言語表現の書き換え処理のうち、少なくとも1種類以上の書き換え処理を入力文に適用することができる。これにより、より多様なバリエーションの言語表現への書き換えを実現することができる。30

#### 【0011】

また、第1及び第2の発明において、前記設定部は、文字種、分かち書き、及び記号類をさらに含む前記複数種類の言語的特徴の各々について、前記設定値を設定し、前記書き換え処理部は、前記設定部により設定された設定値に基づいて、前記文字種を変換する処理、分かち書きに変換する処理、及び記号類を挿入する処理をさらに含む前記複数種類の言語的特徴に基づく言語表現の書き換え処理のうち、少なくとも1種類以上の書き換え処理を入力文に適用することができる。これにより、より多様なバリエーションの言語表現への書き換えを実現することができる。40

#### 【0012】

また、第3の発明に係る言語表現書き換えプログラムは、コンピュータを、上記の言語表現書き換え装置の各部として機能させるためのプログラムである。

#### 【発明の効果】

#### 【0013】

以上説明したように、本発明の言語表現書き換え装置、方法、及びプログラムによれば、文末表現だけでなく、多様な言語表現の書き換え処理を任意に組み合わせて実施するため、多様なバリエーションの言語表現への書き換えを実現することができる。

#### 【図面の簡単な説明】

#### 【0014】

10

20

30

40

50

【図1】本実施形態に係る言語表現書き換え装置の概略構成を示す機能ブロック図である。

【図2】機能表現とその意味との対応表の一例を示す図である。

【図3】文体別機能表現リストの一例を示す図である。

【図4】活用表の一例を示す図である。

【図5】形態素列同士の置換ルールの一例を示す図である。

【図6】キャラクタ別接続表現リストの一例を示す図である。

【図7】キャラクタ別機能表現リストの一例を示す図である。

【図8】崩れ活用ルールの一例を示す図である。

【図9】キャラクタ別人称代名詞リストの一例を示す図である。

10

【図10】キャラクタ別語彙置換ルールの一例を示す図である。

【図11】キャラクタ別音素置換ルールの一例を示す図である。

【図12】本実施形態における言語表現書き換え処理ルーチンの一例を示すフローチャートである。

【発明を実施するための形態】

【0015】

以下、図面を参照して本発明の実施形態を詳細に説明する。

【0016】

<言語表現書き換え装置の構成>

本実施形態に係る言語表現書き換え装置10は、CPUと、RAMと、後述する言語表現書き換え処理ルーチンを実行するための言語表現書き換えプログラムや各種データを記憶したROMとを含むコンピュータで構成することができる。言語表現書き換え装置10は、機能的には、図1に示すように、設定部11と、基本解析部12と、書き換え処理部13とを含む構成で表される。

20

【0017】

言語表現書き換え装置10は、日本語の入力文23(テキストデータ)を入力として受け取り、入力文23の中に含まれる言語表現を、指定された設定に即して書き換えた書き換え文32を出力する。本実施形態では、テレビアニメや漫画の登場人物のような個性の強いキャラクタの発話において頻繁に観察される以下の12種類の言語的特徴に関する書き換え項目を、任意の組み合わせで指定可能にする。

30

【0018】

書き換え項目

(a) 文体、(b) 文字種、(c) 分かち書き、(d) 文構造、(e) 述部機能表現・キャラ語尾、(f) 活用形、(g) 言いよどみ、(h) 人称代名詞、(i) 方言・特殊語彙、(j) 音素置換、(k) 弁別的無意味表現、(l) 記号類

【0019】

設定部11は、後述する書き換え処理部13での言語表現の書き換えに関する設定値が記載された設定ファイル21を読み込む。設定ファイル21には、上記の書き換え項目の各々について、以下のような書き換えに関する設定値(設定を指定するための値、ファイル名、文字列)が記載される。

40

【0020】

(a) 文体変換の設定値(0=無変換、1=常体(だ体)、2=敬体(です・ます体)、3=敬体(でございます体))

(b) 文字種変換の設定値(0=無変換、1=ひらがな化)

(c) 分かち書き変換の設定値(0=無変換、1=読点区切り、2=空白区切り)

(d) 文構造変換用のキャラクタ別接続表現リスト26(詳細は後述)のファイル名(ファイル名を指定しない場合は、無変換とみなす)

(e) 述部機能表現・キャラ語尾変換用のキャラクタ別機能表現リスト27(詳細は後述)のファイル名(ファイル名を指定しない場合は、無変換とみなす)

(f) 活用形変換の設定値(0=無変換、1=崩れた活用形に変換)

50

(g) 言いよどみ変換の設定値 (0 = 無変換、1 = 言いよどみ化)

(h) 人称代名詞置換用のキャラクタ別人称代名詞リスト29(詳細は後述)のファイル名(ファイル名を指定しない場合は、無変換とみなす)

(i) 方言・特殊語彙置換用のキャラクタ別語彙置換ルール30(詳細は後述)のファイル名(ファイル名を指定しない場合は、無変換とみなす)

(j) 音素置換用のキャラクタ別音素置換ルール31(詳細は後述)のファイル名(ルールのファイル名を指定しない場合は、無変換とみなす)

(k) 使用する弁別的無意味表現を示す文字列(弁別的無意味表現を指定しない場合は、無変換とみなす)

(l) 使用する記号類を示す文字列(記号類を指定しない場合は、無変換とみなす)

10

#### 【0021】

設定部11は、読み込んだ設定ファイル21に記載された各項目についての設定値を、設定値データベース(DB)22に記憶する。

#### 【0022】

なお、各項目の詳細については、対応する書き換え処理部13の各処理部の説明で合わせて行うこととし、ここでの説明は省略する。

#### 【0023】

基本解析部12は、入力文23を読み込む。入力文23は、日本語で書かれたテキストデータである。例えば、ブログ、SNS(Social Networking Service)等への書き込み、発話の音声認識結果やテキストチャットなど、文字化された日本語のテキストデータであれば、本実施形態の入力文23として適用可能である。

20

#### 【0024】

基本解析部12は、読み込んだ入力文23を係り受け解析器にかけ、その出力から形態素境界、各形態素の読み、各形態素の品詞、活用語の活用型・活用形、文節境界、及び文節主辞の情報を取得する。なお、基本解析部12では、入力文23の形態素境界、各形態素の読み、各形態素の品詞、活用語の活用型・活用形、文節境界、及び文節主辞の情報が取得できさえすれば、必ずしも係り受け解析器を使用する必要はない。例えば、形態素解析器で形態素境界、読み、品詞、活用語の活用型・活用形を取得し、文節境界や文節主辞は個別のアルゴリズムを使用して求めてよい。

#### 【0025】

30

また、基本解析部12は、図2に示すような機能表現とその意味との対応表を用いて、文字列マッチによって、入力文23に含まれる機能表現の意味ラベルを取得する(参考文献1参照)。なお、機能表現の意味ラベルの取得手法として、例えば参考文献2に記載の、機械学習を用いて適切なラベルを推定する手法を用いてよい。

#### 【0026】

参考文献1：松吉俊、佐藤理史、宇津呂武仁、“日本語機能表現辞書の編纂”、自然言語処理、14.5, 2007

参考文献2：今村賢治、泉朋子、菊井玄一郎、佐藤理史、“述部機能表現の意味ラベルタガー”、言語処理学会第17回年次大会発表論文集、2011

#### 【0027】

40

例えば、「私は寒がりなので、暖かい服装を選んだ。」という入力文23を基本解析部12が読み込んだ場合、以下に示す(1)形態素境界、文節境界、文節主辞、(2)各形態素の読み、(3)各形態素の品詞、活用語の活用型・活用形、(4)機能表現の意味ラベルの情報が取得される。以下、基本解析部12で取得されるこれらの情報をまとめて、基本解析結果という。

#### 【0028】

(1) {私}\_\_は / {寒がり}\_\_な\_\_ので\_\_、 / {暖か}\_\_い / {服装}\_\_を / {選}\_\_ん\_\_だ\_\_。

(2) {ワタシ}\_\_ハ / {サムガリ}\_\_ナ\_\_ノデ\_\_、 / {アタタカ}\_\_イ / {フクソウ}\_\_ヲ / {エラ}\_\_ン\_\_ダ\_\_。

50

(3) {代名詞}\_\_係助詞 / {名詞}\_\_助動詞\_\_接続助詞\_\_読点 / {形容詞語幹}\_\_活用語尾：連体形 / {名詞}\_\_格助詞 / {動詞語幹：バ行五段}\_\_活用語尾：連用形 - 音便\_\_助動詞\_\_句点

(4) { \*}\_\_\* / { \*}\_\_\*\_\_理由\_\_\* / { \*}\_\_\* / { \*}\_\_\* / { \*}\_\_\*\_\_完了\_\_\*

#### 【0029】

上記の基本解析結果では、形態素境界が「\_」、文節境界が「/」、文節主辞が「{ }」で示されている。なお、文節が区切れているところでは、必ず形態素も区切れているため、文節境界は形態素境界でもある。また、活用語（動詞、形容詞、助動詞）の活用語尾は語幹から切り離しておく（例えば、「選\_ん\_だ」）。さらに、(4)において、「\*」は、機能表現ではない、又は機能表現であっても本実施形態で処理対象となる意味ラベルが付与されなかったことを意味する。10

#### 【0030】

書き換え処理部13は、図1に示すように、文体変換部13a、文字種変換部13b、分かち書き変換部13c、文構造変換部13d、述部機能表現・キャラ語尾変換部13e、活用形変換部13f、言いよどみ変換部13g、人称代名詞置換部13h、方言・特殊語彙置換部13i、音素置換部13j、弁別的無意味表現挿入部13k、及び記号類挿入部13lを含む。

#### 【0031】

書き換え処理部13の各処理部には、入力文23について基本解析部12で解析された基本解析結果が各々入力される。ただし、設定値DB22に記憶された各処理部に対応する項目の設定値が、各処理部の書き換え処理を適用しないことを示す場合、その処理部へは基本解析結果は入力されない。例えば、設定値DB22に記憶された項目(a)文体変換の設定値が0=無変換の場合、文体変換部13aには、基本解析結果は入力されない。20

#### 【0032】

以下、基本解析結果が入力された各処理部の書き換え処理について、詳述する。

#### 【0033】

文体変換部13aは、設定値DB22から、項目(a)文体変換の設定値(1=常体(だ体)、2=敬体(です・ます体)、3=敬体(でございます体)のいずれか)を取得する。文体変換部13aは、入力文23の述部の機能表現を、取得した設定値が示す文体に合わせて置換することにより、文体を変換した文を出力する。30

#### 【0034】

具体的には、文体変換部13aは、例えば図3に示すような文体別機能表現リスト24を参照して、入力文23(基本解析結果)の末尾の文節に含まれる機能表現(主辞より後ろの形態素列)の置換後の表記を取得する。図3の例では、文体別機能表現リスト24は、文体(常体(だ体)、敬体(です・ます体)、敬体(でございます体))別に、「文節主辞の品詞」及び「機能表現の意味」毎の置換先の「機能表現の表記」が定められている。

#### 【0035】

なお、文体別機能表現リスト24において、置換の対象となる機能表現が動詞の場合は、置換先の「機能表現の表記」は、置換先の機能表現が要求する動詞の活用語尾の活用形を表すタグ(以下、[活用形]と表記する。例えば、[連用形])を含む表記で定められている。この場合、文体変換部13aは、例えば、図4に示す活用表25を参照し、タグ[活用形]の部分を、置換先の表記に変換する。図4に示す活用表25において、空のセルは、活用語尾を挟まずに語幹と後続の形態素とが接続するものを示す。例えば、一段活用の動詞の未然形(ア段)は、語幹「見」と助動詞「ない」とが直に接続し、「見ない」となる。また、「-」が記載されているセルは、日本語として存在しない組み合わせ(考慮不要の組み合わせ)を示す。

#### 【0036】

上記の「私は寒がりなので、暖かい服装を選んだ。」という入力文23の基本解析結果40

50

に対する文体の変換を例に、文体変換部 13 a の処理をより詳細に説明する。

#### 【0037】

文体変換部 13 a は、基本解析結果から、述部の文節主辞の品詞 = { 動詞語幹 : バ行五段 }、その文節の機能表現の意味ラベル = “完了”の情報を取得する。そして、文体変換部 13 a は、図 3 に示す文体別機能表現リスト 24 を参照して、基本解析結果から取得した情報と一致する「文節主辞の品詞」及び「機能表現の意味」に対応付けられた「機能表現の表記」を取得する。設定値 DB22 から取得された設定値が 2 = 敬体（です・ます体）、又は 3 = 敬体（でございます体）の場合、下記に示すような処理中間結果が得られる。なお、以下では、置換された箇所を <> で表す。

#### 【0038】

敬体（です・ます体）の例：

私は寒がりなので、暖かい服装を選<< [ 連用形 ] ました>>。

敬体（でございます体）の例：

私は寒がりなので、暖かい服装を選<< [ 連用形 - 音便 ] たのでございます>>。

#### 【0039】

上記の処理中間結果のように、機能表現が置換された文節がタグ [ 活用形 ] を含む場合には、文体変換部 13 a は、活用表 25 から、タグ [ 活用形 ] が示す動詞の活用形と、その文節の主辞である動詞の活用型とが一致する活用形の表記を取得する。そして、文体変換部 13 a は、下記に示すように、処理中間結果に含まれるタグ [ 活用形 ] を、活用表 25 から取得した活用形の表記に置換する。

#### 【0040】

敬体（です・ます体）の例：

私は寒がりなので、暖かい服装を選<< [ 連用形 ] ました>>。

私は寒がりなので、暖かい服装を選<< びました>>。

文節「選 [ 連用形 ] ました」における、文節主辞の品詞は「動詞語幹：バ行五段」、タグ「活用形」が示す活用形は「連用形」であるので、この条件にマッチする活用形の表記「び」が取得される。

#### 【0041】

敬体（でございます体）の例：

私は寒がりなので、暖かい服装を選<< [ 連用形 - 音便 ] たのでございます>>。

私は寒がりなので、暖かい服装を選<< んだでございます>>。

文節「選 [ 連用形 - 音便 ] たのでございます」における、文節主辞の品詞は「動詞語幹：バ行五段」、タグ「活用形」が示す活用形は「連用形（音便形）」であるので、この条件にマッチする活用形の表記「ん」が取得される。

#### 【0042】

なお、文節主辞の活用型が「ガ行五段活用」、「バ行五段活用」、「マ行五段活用」、及び「ナ行五段活用」のいずれかである場合は、置換先の機能表現の先頭文字（活用語尾に後続する文字）「て」又は「た」をそれぞれ「で」又は「だ」に置換する。上記の例では、文節主辞の活用型が「バ行五段活用」であるので、「選 [ 連用形 - 音便 ] たのでございます」が、「選んだでございます」に置換されている。

#### 【0043】

なお、上記では、適切な活用語尾を挿入するために、置換先の機能表現が要求する動詞の活用語尾の活用形を表すタグを利用したが、この方法に限定されない。例えば、2つの形態素を結合する際にどのような活用語尾が挿入されるべきかを、何らかの機械学習の手法によって事前に学習しておき、文体変換部 13 a における活用語尾の挿入に利用してもよい。例えば、「選」と「た」との間にどのような活用語尾が入るべきかを推定するモデルを学習しておく。そして、語幹が「選」の動詞を「語幹 + 活用語尾 + 助動詞「た」」の形に置換したい場合には、学習したモデルの出力から置換先の活用語尾を得る、という使い方ができる。

#### 【0044】

10

20

30

40

50

また、別の方法として、置換対象の形態素（列）の前後にどのような形態素が共起しているかを考慮した形態素列同士の置換ルールを用いて文体を変換することも可能である。この場合、例えば、図5に示すような形態素列同士の置換ルールを用いて、文体を変換することができる。

#### 【0045】

なお、本実施形態では、取り得る設定値が0＝無変換、1＝常体（だ体）、2＝敬体（です・ます体）、3＝敬体（でございます体）の4種類の場合について説明するが、他の文体（例えば、「常体（である体）」）へ変換するための設定値を追加してもよい。

#### 【0046】

文字種変換部13bは、設定値DB22に記憶された項目（b）文字種変換の設定値（1＝ひらがな化）にしたがって、漢字をひらがなに変換した文を出力する。具体的には、文字種変換部13bは、下記に示すように、基本解析結果に含まれる各形態素の読み（カタカナで書かれた部分）を全てひらがなに置換する。

#### 【0047】

ひらがな化の例：

わたしはさむがりなので、あたたかいふくそうをえらんだ。

#### 【0048】

なお、本実施形態では、取り得る設定値が0＝無変換、又は1＝ひらがな化の2種類の場合について説明するが、全ての文字をカタカナに変換する「カタカナ化」のオプションを用意してもよい。また、「50%（2回に1回の割合で）ひらがな化する」、「20%（5回に1回の割合で）カタカナ化する」などのように、文字種変換を実施する割合を指定できるようにしてもよい。また、品詞が「名詞」の形態素のみひらがな化する、などのように、文字種変換の対象とする品詞を指定してもよい。

#### 【0049】

分かち書き変換部13cは、設定値DB22から、項目（c）分かち書き変換の設定値（1＝読点区切り、又は2＝空白区切り）を取得する。分かち書き変換部13cは、入力文23の文節境界に、指定された区切り文字を挿入することにより、下記に示すように、分かち書きが変換された文を出力する。なお、区切り文字（読点又は空白）を挿入することにより、読点や空白が連続してしまう場合は、区切り文字を挿入しないこととする。また、「暖かい」と「服装」との間のように、連体修飾関係にある2つの連続する文節の間には区切り文字を挿入しないこととする。

#### 【0050】

分かち書き変換（読点区切り）の例：

私は<<、>>寒がりなので、暖かい服装を<<、>>選んだ。

#### 【0051】

文構造変換部13dは、設定値DB22から、項目（d）文構造変換用のキャラクタ別接続表現リスト26のファイル名を取得する。そして、文構造変換部13dは、入力文23を単文に分割すると共に、取得したファイル名が示すキャラクタ別接続表現リスト26から得られる接続表現を分割箇所に挿入することで、文構造を変換した文を出力する。

#### 【0052】

具体的には、入力文23が複文である場合は、文構造変換部13dは、例えば図6に示すようなキャラクタ別接続表現リスト26のうち、取得したファイル名が示すキャラクタ別接続表現リスト26を参照する。そして、文構造変換部13dは、入力文23から、キャラクタ別接続表現リスト26に記載された意味ラベルを持つ文節を探し、該当する意味ラベルが付与されている形態素を削除し、削除箇所に句点を挿入する。また、文構造変換部13dは、上記意味ラベルが付与されている形態素の直前の形態素が活用語（動詞、形容詞、助動詞）である場合は、その活用語を終止形に変換する。終止形としては、形態素解析結果の一部として出力されることの多い「基本形」を利用してもよいし、活用表25を利用してもよい。そして、文構造変換部13dは、上記意味ラベルに対応する接続表現をキャラクタ別接続表現リスト26から取得し、挿入した句点の後ろに挿入する。なお、

10

20

30

40

50

接続表現と共に読点を挿入してもよい。

#### 【0053】

例えば、設定値DB22に記憶された項目(d)文構造変換用のキャラクタ別接続表現リスト26のファイル名として、キャラAのキャラクタ別接続表現リスト26が指定されているとする。上記の「私は寒がりなので、暖かい服装を選んだ。」という入力文23では、文節「寒がりなので、」の形態素「ので」の意味ラベル「理由」が、キャラクタ別接続表現リスト26に記載された意味ラベルの1つに該当する。そこで、「ので」及びその後の読点「、」を削除すると共に、削除箇所に句点「。」を挿入する。また、「ので」の前の助動詞「な」を終止形「だ」に置換する。さらに、挿入した句点「。」の後に、キャラAのキャラクタ別接続表現リスト26において、意味ラベル「理由」に対応付けられている接続詞「なので」を取得し、読点「、」と共に挿入する。これにより、下記に示すように、入力文23の文構造が変換される。10

#### 【0054】

単文化の例：

私は寒がり<<だ。なので、>>暖かい服装を選んだ。

#### 【0055】

なお、上記では、文の境界を示すために句点を挿入することとしたが、文の境界を示すことができさえすればどのような記号を用いてもよい。

#### 【0056】

また、「は」などの提題を表す助詞又は主格を表す助詞「が」が複数個現れる入力文23は、単文化の対象外、すなわち、文構造変換部13dによる書き換え処理の対象外としてもよい。例えば、「私は、彼が寒がりなので、マフラーを貸してあげました。」という入力文23が、「私は、彼が寒がりだ。なので、マフラーを貸してあげました。」のように書き換えられることを避けるためである。20

#### 【0057】

述部機能表現・キャラ語尾変換部13eは、設定値DB22から、項目(e)述部機能表現・キャラ語尾変換用のキャラクタ別機能表現リスト27のファイル名を取得する。キャラクタ別機能表現リスト27の一例を図7に示す。キャラクタ別機能表現リスト27の構成は、文体別機能表現リスト24と同様である。また、述部機能表現・キャラ語尾変換部13eの処理も、文体変換部13aの処理と同様である。ただし、キャラクタ別機能表現リスト27では、キャラクタの個性を表現可能な述部機能表現及び語尾を任意に定めることができる。図7に示すキャラBのキャラクタ別機能表現リスト27の例のように、キャラクタ付けのために日本語文法の範囲外の表現(キャラ語尾)を用いることもできる。これにより、述部機能表現・キャラ語尾変換部13eでは、文体変換部13aとは異なる言語表現の書き換えを実現することができる。30

#### 【0058】

例えば、設定値DB22に記憶された項目(e)述部機能表現・キャラ語尾変換用のキャラクタ別機能表現リスト27のファイル名として、キャラAのキャラクタ別機能表現リスト27が指定されているとする。この場合、上記の「私は寒がりなので、暖かい服装を選んだ。」という入力文23は、下記のように述部機能表現が変換される。40

#### 【0059】

述部機能表現・キャラ語尾変換の例：

私は寒がりなので、暖かい服装を選ん<<だの>>。

#### 【0060】

活用形変換部13fは、設定値DB22に記憶された項目(f)活用形変換の設定値(1=崩れた活用形に変換)にしたがって、形容詞(語幹及び活用語尾)を崩れた表現に置換することにより、活用形を変換した文を出力する。具体的には、活用形変換部13fは、入力文23に含まれる形容詞の語幹及び活用語尾の表記を、例えば図8に示すような崩れ活用ルール28の「入力(表記)」部分と照合し、対応する「出力」部分の語幹及び活用語尾の表記を取得して置換する。50

**【0061】**

例えば、上記の「私は寒がりなので、暖かい服装を選んだ。」という入力文23の場合、形容詞「暖かい」の「かい」の部分が崩れ活用ルール28の「入力（表記）」の1つに該当する。そこで、「かい」の部分を、対応する「出力（表記）」である「けえ」に置換することで、下記に示すように、入力文23の活用形を崩れた活用形に変換する。

**【0062】**

崩れた活用形に変換の例：

私は寒がりなので、暖<<けえ>>服装を選んだ。

**【0063】**

なお、本実施形態では、例えば図8に示すような崩れ活用ルール28を用いたルールベースの手法で変換処理を行う場合について説明したが、これに限定されない。例えば、何らかの機械学習を用いてコーパス（崩れた活用形が使用されるテキストデータ）から学習しておいたモデルを用いて、崩れていない語幹及び活用語尾の置換先となる崩れた語幹及び活用語尾を推定するなどしてもよい。

10

**【0064】**

言いよどみ変換部13gは、設定値DB22に記憶された項目（g）言いよどみ変換の設定値（1=言いよどみ化）にしたがって、入力文23を、言いよどみが表れた文に変換する。具体的には、言いよどみ変換部13gは、入力文23の文頭の形態素の読み（本実施形態の基本解析結果の例ではカタカナで記載）の1文字目をひらがな化し、文頭の形態素の前に挿入する。なお、挿入したひらがなの後に、読点を挿入してもよい。

20

**【0065】**

例えば、上記の「私は寒がりなので、暖かい服装を選んだ。」という入力文23の場合、下記に示すように、入力文23が言いよどみ化された文に変換される。

**【0066】**

言いよどみ化の例：

<<わ、>>私は寒がりなので、暖かい服装を選んだ。

**【0067】**

なお、上記では、文頭の形態素のみを言いよどみ化の対象とする場合について説明したが、節の先頭の形態素を対象として、「わ、私は寒がりなので、あ、暖かい服装を選んだ。」のように変換してもよいし、各文節の先頭の形態素を対象として、「わ、私は、さ、寒がりなので、あ、暖かい服装を、え、選んだ。」のように変換してもよい。また、言いよどみ化対象の形態素の表記がカタカナの場合は、言いよどみ化対象の形態素の前に挿入する文字をカタカナにしてもよい。また、例えば、「わ、わ、私は・・・」のように、挿入するひらがな又はカタカナを2回以上重ねて挿入してもよい。

30

**【0068】**

人称代名詞置換部13hは、設定値DB22から、項目（h）人称代名詞置換用のキャラクタ別人称代名詞リスト29のファイル名を取得する。そして、人称代名詞置換部13hは、入力文23に含まれる人称代名詞の形態素を、取得したファイル名が示すキャラクタ別人称代名詞リスト29から得られる人称代名詞に置換した文を出力する。

40

**【0069】**

具体的には、人称代名詞置換部13hは、例えば図9に示すようなキャラクタ別人称代名詞リスト29のうち、取得したファイル名が示すキャラクタ別人称代名詞リスト29を参照する。そして、人称代名詞置換部13hは、入力文23において、キャラクタ別人称代名詞リスト29の「入力」部分と表記が一致する形態素を、対応する「出力」部分の表記と置換する。

**【0070】**

例えば、設定値DB22に記憶された項目（h）人称代名詞置換用のキャラクタ別人称代名詞リスト29のファイル名として、キャラAのキャラクタ別人称代名詞リスト29が指定されているとする。この場合、上記の「私は寒がりなので、暖かい服装を選んだ。」という入力文23は、下記のように人称代名詞が置換される。

50

**【0071】**

人称代名詞置換の例：

<<あたし>>は寒がりなので、暖かい服装を選んだ。

**【0072】**

方言・特殊語彙置換部13iは、設定値DB22から、項目(i)方言・特殊語彙置換用のキャラクタ別語彙置換ルール30のファイル名を取得する。そして、方言・特殊語彙置換部13iは、入力文23に含まれる特定の形態素を、取得したファイル名が示すキャラクタ別語彙置換ルール30にしたがって、方言又は特殊語彙に置換した文を出力する。

**【0073】**

具体的には、方言・特殊語彙置換部13iは、例えば図10に示すようなキャラクタ別語彙置換ルール30のうち、取得したファイル名が示すキャラクタ別語彙置換ルール30を参照する。そして、方言・特殊語彙置換部13iは、入力文23において、キャラクタ別語彙置換ルール30の「入力」部分と表記が一致する形態素を、対応する「出力」部分の表記と置換する。キャラクタ別語彙置換ルール30では、特定の置換元の語彙を、目的のキャラクタらしい語彙に変換するルールが定められる。10

**【0074】**

また、方言・特殊語彙置換部13iは、置換先の語彙が活用語(動詞、形容詞、助動詞)である場合は、置換元の語彙(「入力」部分が該当)の活用形に合わせて活用語尾を調整してもよい。なお、助動詞の活用については、動詞型の活用をするものは動詞の活用表を参照し、形容詞型の活用をするものは形容詞の活用表を参照して取得することができる20。

**【0075】**

例えば、設定値DB22に記憶された項目(i)方言・特殊語彙置換用のキャラクタ別語彙置換ルール30のファイル名として、キャラAのキャラクタ別語彙置換ルール30が指定されているとする。この場合、上記の「私は寒がりなので、暖かい服装を選んだ。」という入力文23は、下記のように特定の語彙が方言又は特殊語彙に置換される。

**【0076】**

方言・特殊語彙置換の例：

私は寒がりなので、<<ぬくとい>>服装を選んだ。

**【0077】**

なお、方言・特殊語彙置換部13iで置換された表現は、その他の書き換え処理において更なる書き換え処理が加わらないように保護してもよい。30

**【0078】**

音素置換部13jは、設定値DB22から、項目(j)音素置換用のキャラクタ別音素置換ルール31のファイル名を取得する。そして、音素置換部13jは、入力文23に含まれる特定の文字を、取得したファイル名が示すキャラクタ別音素置換ルール31にしたがって、置換先の文字に置換した文を出力する。

**【0079】**

具体的には、音素置換部13jは、例えば図11に示すようなキャラクタ別音素置換ルール31のうち、取得したファイル名が示すキャラクタ別音素置換ルール31を参照する。そして、音素置換部13jは、入力文23において、キャラクタ別音素置換ルール31の「入力」部分と表記が一致する文字を、対応する「出力」部分の表記と置換する。なお、本実施形態では、音素の置換を文字単位で捉えて置換することとする。例えば、「な」から「にや」、「の」から「によ」への文字的な置換は、「na」から「ny a」へ、「no」から「ny o」への音素的な置換(「n」から「ny」への音素置換)を捉えるためのものである。40

**【0080】**

例えば、設定値DB22に記憶された項目(j)音素置換用のキャラクタ別音素置換ルール31のファイル名として、キャラAのキャラクタ別音素置換ルール30が指定されているとする。この場合、上記の「私は寒がりなので、暖かい服装を選んだ。」という入力50

文23は、下記のように特定の文字が置換される。

**【0081】**

音素的な置換を捉えた文字置換の例：

私は寒がり<<にやによ>>で、暖かい服装を選んだ。

**【0082】**

なお、キャラクタ別音素置換ルール31の「入力」部分と一致する全ての文字を置換する場合に限定されず、文字の置換を実行するか否かを、その文字が属する形態素の品詞や、形態素内での出現位置などを条件（制約）にして決定してもよい。例えば、「ある文字が属する形態素の品詞が名詞である場合は置換しない」、「ある文字が形態素の先頭に位置する場合は置換しない」、などの制約を設けることができる。他にも、「同じ形態素内で複数回の置換を行ってはいけない」という制約や、「連續した2つ以上の文字を置換してはいけない」という制約を設けてもよい。10

**【0083】**

ここで問題になるのが、制約の数が多くなると、制約のあらゆる組み合わせを考慮したルールを人手で定義するのが困難になるという点である。そこで、置換元（入力）の文字、置換先（出力）の文字、置換元文字が属する形態素の品詞、置換元文字の出現位置、置換元文字が属する形態素内で既に実施された置換の回数、置換元文字までの連続文字置換回数などを特徴量として、何らかの機械学習によりモデルを学習しておき、このモデルを使用して、文字の置換を実施するようにしてもよい。

**【0084】**

弁別的無意味表現挿入部13kは、設定値DB22に記憶された項目（k）弁別的無意味表現を示す文字列を取得し、入力文23の末尾に挿入することで、弁別的無意味表現が挿入された文を出力する。本実施形態において、弁別的無意味表現とは、日本語としては何の意味も持たないが、キャラクタの弁別を補助する表現のことを指す。20

**【0085】**

例えば、弁別的無意味表現を示す文字列として、「ピヨン！」が指定されている場合、上記の「私は寒がりなので、暖かい服装を選んだ。」という入力文23には、下記のように弁別的無意味表現が挿入される。

**【0086】**

弁別的無意味表現挿入の例：

私は寒がりなので、暖かい服装を選んだ。<<ピヨン！>>

**【0087】**

なお、本実施形態では、文末の句点の後ろに弁別的無意味表現を挿入する。句点がなければ、句点を挿入したうえで、弁別的無意味表現を句点に後続させる。

**【0088】**

例えば、風貌がカエルのようなキャラクタ（カエル）、ボールのように丸いキャラクタ（ボール）、トゲがたくさん生えたキャラクタ（トゲ）という3種のキャラクタが存在する。この3者が似通った言語表現を使うため、言語的な差異が伝わりづらい場合でも、カエルの発話に弁別的無意味表現「ピヨン！」を挿入することで、ボールやトゲではなく、カエルの発話であることを読み手又は聞き手に対して強く印象付けることができる。40

**【0089】**

なお、上記では、入力文23の末尾に弁別的無意味表現を挿入する場合について説明したが、「ピヨン！私は寒がりなので、暖かい服装を選んだ。」のように、文頭に挿入するなど、その他の箇所に弁別的無意味表現を挿入してもよい。

**【0090】**

記号類挿入部13lは、設定値DB22に記憶された項目（l）記号類を示す文字列を取得し、入力文23の末尾に挿入することで、記号類が挿入された文を出力する。記号類挿入部13lの処理は、挿入する文字列が弁別的無意味表現ではなく記号類である点を除いて、弁別的無意味表現挿入部13kと同様である。なお、本実施形態において、記号類50

とは、（星）や（音符）のような記号や、（\*^o^\*）や（>\_<；）のような顔文字を指すこととする。

#### 【0091】

例えば、記号類を示す文字列として、「（\*^o^\*）」が指定されている場合、上記の「私は寒がりなので、暖かい服装を選んだ。」という入力文23には、下記のように記号類が挿入される。

#### 【0092】

記号類挿入の例：

私は寒がりなので、暖かい服装を選んだ。<<（\*^o^\*）>>

#### 【0093】

上記のように、書き換え処理を行うことが設定された書き換え処理部13の各処理部13a～13lにおいて、入力文23（基本解析結果）に対する書き換え処理が行われ、各処理部13a～13lから書き換え文32が出力される。

#### 【0094】

なお、書き換え処理部13の各処理部13a～13lの書き換え処理は任意の組み合わせで実行することができる。例えば、ある処理部で書き換えられた結果を別の処理部でさらに書き換える場合は、ある処理部の出力を再度、基本解析部12に渡し、新規に取得した基本解析結果を次の処理部に渡せばよい。又は、ある処理部での書き換えに基づいて、入力文23の基本解析結果を書き換えた上で、次の処理部に渡すようにしてもよい。例えば、文体変換部13aで書き換えられた結果を、別の処理部に渡す場合、以下のような基本解析結果を次の処理部へ渡すことができる。

#### 【0095】

(1) {私}\_\_は/\_{寒がり}\_\_な\_\_ので\_\_/\_{暖か}\_\_い/\_{服装}\_\_を/\_{選}\_\_  
<<び>>\_<<ました>>\_

(2) {ワタシ}\_\_ハ/\_{サムガリ}\_\_ナ\_\_ノデ\_\_/\_{アタタカ}\_\_イ/\_{フクソウ}\_\_ヲ/\_{エラ}\_\_<<d umm y>>\_<<d umm y>>\_

(3) {代名詞}\_\_係助詞/\_{名詞}\_\_助動詞\_\_接続助詞\_\_読点/\_{形容詞語幹}\_\_活用語尾：連体形/\_{名詞}\_\_格助詞/\_{動詞語幹：バ行五段}\_\_<<活用語尾：連用形>>\_<<d umm y>>\_句点

(4) {\*}\_\* / {\*}\_\*\_\_理由\_\* / {\*}\_\* / {\*}\_\* / {\*}\_\*\_\_完了  
-\*

#### 【0096】

<<>>箇所は書き換えられた形態素に関する情報を示す。なお、上記の例では、他の処理部で使用される機会のない情報は「d umm y」としているが、正しい情報を付与してもよい。なお、(2)形態素の読みは、文字種変換部13bにおけるひらがな化で使用するが、「び\_ました」が既にひらがなであるため、「d umm y」としても問題ない。また、本実施形態では、機能表現の品詞を使う処理がないため、「ました」の品詞は「d umm y」としても問題ない。

#### 【0097】

書き換え処理部13の各処理部13a～13lの書き換え処理を任意に組み合わせることで、「私は寒がりなので、暖かい服装を選んだ。」という入力文23について、例えば、「(I)<<オレ>>は寒がり<<だ。だから>>、暖<<けえ>>服装を選んだ<<ぜ！>>」や、「(II)<<あ、あたしは、さむ>>がりなの。だ、だから、あたたかいふくそう>>を<<、えらんだの>>」のように個性豊かな書き換えが可能となる。複数の処理部13a～13lの書き換え処理を適用した場合には、最終的な処理部の出力を、書き換え文32として出力する。なお、(I)の例は、(h)人称代名詞、(c)分かち書き、(f)活用形、及び(e)述部機能表現・キャラ語尾の項目についての書き換え処理を組み合わせた例である。また、(II)の例は、(b)文字種、(g)言いよどみ、(h)人称代名詞、(c)分かち書き、及び(e)述部機能表現・キャラ語尾の項目についての書き換え処理を組み合わせた例である。

**【0098】**

なお、本実施形態では、複数の書き換え処理を適用する場合、例えば、以下の点を考慮して、適用する書き換え処理の順番を定めておくことができる。

- ・(d)文構造の変換は、その他11種の書き換え処理よりも先に適用するのが良い。
- ・(g)言いよどみへの変換、及び(b)文字種の変換は、(d)文構造、(i)方言・特殊語彙、(a)文体、(e)述部機能表現・キャラ語尾、及び(h)人称代名詞よりも後に適用するのが良い。ここで、(g)言いよどみへの変換と(b)文字種の変換の適用順序はどちらが先でも構わない。
- ・(i)方言・特殊語彙の置換は、(d)文構造より後、かつ、(a)文体、(e)述部機能表現・キャラ語尾の変換よりも先に適用するのが良い。

本実施形態では、上記3点を踏まえ、(d)文構造、(i)方言・特殊語彙置換、(a)文体、(e)述部機能表現・キャラ語尾、(h)人称代名詞、(b)文字種、(f)活用形、(g)言いよどみ、(j)音素置換、(c)分かち書き、(k)弁別的無意味表現、(l)記号類の順で書き換え処理を適用するものとする。

**【0099】**

なお、図3～図11に示した各リスト又は各ルール内の\*は、任意の文字列を表す。また、図3～図11に示した各リスト又は各ルール内の「例」又は「備考」は、各リスト又は各ルールの説明を補助する適用例等であり、各リスト又は各ルールの項目として定めておく必要はない。

**【0100】**

<言語表現書き換え装置の作用>

次に、本実施形態に係る言語表現書き換え装置10の作用について説明する。目的のキャラクタに応じた言語表現の書き換えに関する設定値が記載された設定ファイル21、及びキャラクタに応じた言語表現への書き換えの対象となる入力文23が言語表現書き換え装置10に入力されると、言語表現書き換え処理装置10において、図12に示す言語表現書き換え処理ルーチンが実行される。

**【0101】**

ステップS11で、設定部11が、設定ファイル21を読み込み、読み込んだ設定ファイル21に記載された各項目についての設定値を、設定値DB22に記憶する。

**【0102】**

次に、ステップS12で、基本解析部12が、入力された入力文23を読み込み、読み込んだ入力文23を解析し、形態素境界、各形態素の読み、各形態素の品詞、活用語の活用型・活用形、文節境界、文節主辞、及び機能表現の意味ラベルの情報を取得する。基本解析部12は、取得した情報を、基本解析結果として、書き換え処理部13へ出力する。

**【0103】**

次に、ステップS13で、上記で詳述したように、書き換え処理部13の各処理部13a～13lで、入力文23（基本解析結果）に対する書き換え処理を行う。次に、ステップS14で、書き換え処理部13が、上記ステップS13の書き換え処理の結果である書き換え文32を出力して、言語表現書き換え処理ルーチンは終了する。

**【0104】**

以上説明したように、本実施形態に係る言語表現書き換え処理装置によれば、(a)文体、(b)文字種、(c)分かち書き、(d)文構造、(e)述部機能表現・キャラ語尾、(f)活用形、(g)言いよどみ、(h)人称代名詞、(i)方言・特殊語彙、(j)音素置換、(k)弁別的無意味表現、及び(l)記号類の12種類の書き換え項目を任意に組み合わせた書き換え処理を実施する。このため、文末表現だけでなく、多様な言語表現の書き換えを任意に組み合わせて実施することができ、多様なバリエーションの言語表現への書き換えを実現することができる。

**【0105】**

また、書き換えの対象を文末表現に限定していないため、非特許文献1の手法では実現が難しかった、テレビアニメや漫画の登場人物のような個性の強いキャラクタらしい言語

10

20

30

40

50

表現への変換が可能となる。

#### 【0106】

例えば、入力文が「私は寒がりなので、暖かい服装を選んだ。」という文の場合、文末表現の変換のみを行う従来技術では、例えば「私は寒がりなので、暖かい服装を選んだぜ！」のようになる。一方、本実施形態のように、多様な言語表現の書き換えを任意に組み合わせて実施することで、例えば、上述したような「（I）オレは寒がりだ。だから、暖けえ服装を選んだぜ！」や、「（II）あ、あたしは、さむがりなの。だ、だから、あたたかいふくそうを、えらんだの」のように多様で、個性豊かな書き換えが可能となる。

#### 【0107】

人と対話をするシステム（対話システム）に本実施形態を適用した場合には、システムを運営する者（システムのキャラクタをデザインする者）は、キャラクタ別の設定ファイルと数種の表現リスト及び変換ルールとを用意するだけで、多様なキャラクタらしさを持つ発話を簡単に作成することができ、対話システムのキャラクタを増やす際にかかるコストを大幅に削減することができる。10

#### 【0108】

なお、本発明は、上述した実施形態に限定されるものではなく、この発明の要旨を逸脱しない範囲内で様々な変形や応用が可能である。

#### 【0109】

例えば、書き換え処理部13の各処理部13a～13lでの書き換え処理の方法は、上述した方法に限定されず、他の方法を適用してもよい。また、書き換え項目は、上記の（a）～（l）に限定されず、キャラクタに応じた発話において、そのキャラクタの特徴が表れる言語的特徴に関する書き換え項目であればよい。20

#### 【0110】

また、上記実施形態では、複数の書き換え処理を組み合わせて適用する場合に、予め定めた順番で適用する場合について説明したが、適用される書き換え処理の順番を指定することにより、適用順を変更可能してもよい。この場合、設定ファイルに書き換え処理の順番を指定する情報も記載しておき、この情報に基づいて、各処理部を機能させるようにすればよい。例えば、「野原の花」という入力文23に対して、図11に示すキャラAのキャラクタ別音素置換ルール31を適用する（j）音素置換と、ひらがな化の（b）文字種とを適用する場合を考える。（j）音素置換（b）文字種の順で書き換え処理を適用した場合には、「のはらによはな」と書き換えられる。一方、（b）文字種（j）音素置換の順で書き換え処理を適用した場合には、「によはらによはにゃ」と書き換えられる。このように、書き換え処理の順番を変えることによっても、キャラクタの個性の強さが異なるような、多様なバリエーションを表現することができる。30

#### 【0111】

また、（b）文字種、（c）分かち書き、及び（l）記号類は、出力される書き換え文32がテキストデータの場合に効果を發揮する書き換え項目である。つまり、最終出力が音声合成による読み上げである対話システム等の発話に適用する場合には、キャラクタ性付与効果を発揮しない。したがって、書き換え文32の出力が音声合成による読み上げのみの場合には、文字種変換部13b、分かち書き変換部13c、及び記号類挿入部13lは、書き換え処理部13の構成から省略してもよい。又は、書き換え文32の出力形態に応じて、（b）文字種、（c）分かち書き、及び（l）記号類の書き換え項目の設定値を変更するようにしてもよい。例えば、書き換え文32の出力が音声合成による読み上げのみの場合であって、設定ファイルに、（b）文字種、（c）分かち書き、及び（l）記号類の書き換え処理を適用することを示す設定値が記載されていた場合には、設定値DB22に設定値を記憶する際に、書き換え処理を適用しないことを示す設定値に変更した上で記憶するようにする。40

#### 【0112】

また、本願明細書中において、プログラムが予めインストールされている実施形態として説明したが、当該プログラムを、コンピュータ読み取り可能な記録媒体に格納して提供50

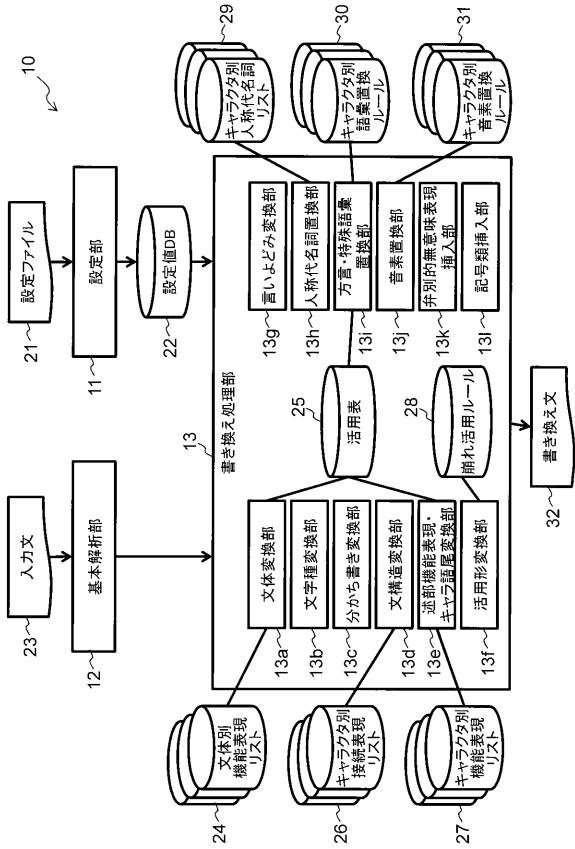
することも可能であるし、ネットワークを介して提供することも可能である。

【符号の説明】

【0 1 1 3】

1 0	言語表現書き換え装置	
1 1	設定部	
1 2	基本解析部	
1 3	書き換え処理部	
1 3 a	文体変換部	
1 3 b	文字種変換部	
1 3 c	分かち書き変換部	10
1 3 d	文構造変換部	
1 3 e	述部機能表現・キャラ語尾変換部	
1 3 f	活用形変換部	
1 3 g	言いよどみ変換部	
1 3 h	人称代名詞置換部	
1 3 i	方言・特殊語彙置換部	
1 3 j	音素置換部	
1 3 k	弁別的無意味表現挿入部	
1 3 l	記号類挿入部	
2 1	設定ファイル	20
2 2	設定値データベース	
2 3	入力文	
2 4	文体別機能表現リスト	
2 5	活用表	
2 6	キャラクタ別接続表現リスト	
2 7	キャラクタ別機能表現リスト	
2 8	崩れ活用ルール	
2 9	キャラクタ別人称代名詞リスト	
3 0	キャラクタ別語彙置換ルール	
3 1	キャラクタ別音素置換ルール	30
3 2	書き換え文	

【図1】



【図2】

意味	機能表現(表記)
否定	ない
疑問	か
完了	た(た)
推量	だろう
	...

【図3】

常体(だ体)		
文節主辞の品詞	機能表現の意味	機能表現の表記 (活用語尾を含む)
形容詞	い	美しい
形容詞	疑問	いか
形容詞	否定	くない
形容詞	完了	かった
	...	
名詞	だ	鳥だ
名詞	疑問	鳥か
名詞	否定	鳥ではない
名詞	完了	鳥だった
	...	
動詞	[終止形]	走る
動詞	[終止形]か	走るか
動詞	[未然形-ア段]ない	走らない
動詞	[連用形-音便]つた	走了た
	...	

敬体(でございます体)		
文節主辞の品詞	機能表現の意味	機能表現の表記 (活用語尾を含む)
形容詞	いです	美しいです
形容詞	疑問	いかですか
形容詞	否定	くあいません
形容詞	完了	かたかったです
	...	
名詞	です	鳥です
名詞	疑問	鳥ですか
名詞	否定	鳥ではありません
名詞	完了	鳥でした
	...	
動詞	[連用形]ます	走ります
動詞	[連用形]ますか	走りますか
動詞	[連用形]ません	走りません
動詞	[連用形]ました	走りました
	...	

敬体(でござります体)		
文節主辞の品詞	機能表現の意味	機能表現の表記 (活用語尾を含む)
形容詞	うござります	美しうござります
形容詞	疑問	うござりますか
形容詞	否定	うございません
形容詞	完了	うございました
	...	
名詞	でござります	鳥でござります
名詞	疑問	鳥でござりますか
名詞	否定	鳥ではございません
名詞	完了	鳥でございました
	...	
動詞	[終止形]でござります	走るのでござります
動詞	[終止形]でござりますか	走るのでござりますか
動詞	[終止形]ではございません	走るのではございません
動詞	[連用形-音便]たのでござります	走ったのでござります
	...	

【図4】

例 見る書く登る選ぶ							
品詞(語幹)	活用型	未然形(ア段)	未然形(カ段)	連用形	連用形(音便)	終止形	連体形
動詞	一段活用	よ	き	い	る	る	れ
	二行五段活用	こ	り	く	く	け	け
	三行五段活用	ろ	り	つ	る	る	れ
	四行五段活用	ほ	び	ん	ぶ	ぶ	べ
形容詞	一	かろ	く	かっ	い	い	かれ

【図5】

入力				出力			
述語種別(文節主辞の品詞)	処理対象の形態素(列)	後続する形態素	置換先の形態素(列)	備考			
名詞	です	*	*	「です」を削除			
名詞	です	の*	な	例えば、「ですので」から「なので」に変換			
名詞	でしょ	*	だろ	例えば、「でしょう」から「だろう」に変換			
...							
動詞	[連用形]まし	た	[連用形]	「まし」を削除			
動詞	[連用形]まし	*	[未然形-オ段]	「ましょ」を削除し、活用語尾を未然形(オ段)に変換			
...							

【図6】

キャラA		キャラB	
意味ラベル	接続詞	意味ラベル	接続詞
理由	なので	だからよお	
逆接確定	けれど	けどよお	
状況	そのとき	そんときよお	
想外	それなのに	なのによお	
	...	...	

【図7】

キャラA		
文節主辞 の品詞	機能表現 の意味	機能表現の表記 (活用語尾を含む)
形容詞	いの	
形容詞	いの?	
形容詞	くないの	
形容詞	かつたの	
名詞	なの	
名詞	なの?	
名詞	じやないの	
名詞	だったの	
動詞	[終止形]	
動詞	[終止形]の?	
動詞	否定 [未然形-ア段]ないの	
動詞	完了 [連用形-音便]たの	
	...	

キャラB		
文節主辞 の品詞	機能表現 の意味	機能表現の表記 (活用語尾を含む)
形容詞	いビヨン	
形容詞	いビヨンか?	
形容詞	くないビヨン	
形容詞	かつたビヨン	
名詞	だビヨン	
名詞	だビヨンか?	
名詞	じやないビヨン	
名詞	だったビヨン	
動詞	[終止形]	
動詞	[終止形]ビヨンか?	
動詞	否定 [未然形-ア段]ないビヨン	
動詞	完了 [連用形-音便]たビヨン	
	...	

【図8】

キャラA	
入力(表記)	出力(表記)
形容詞語幹 活用語尾	形容詞語幹 活用語尾
*か い	*け え
*さ い	*せ え
*た い	*て え
	...

28

キャラB	
入力(表記)	出力(表記)
形容詞語幹 活用語尾	形容詞語幹 活用語尾
*か い	*け え
*さ い	*せ え
*た い	*て え
	...

【図9】

キャラA	
入力	出力
私	あたし
わたし	あたし
あたし	あたし
俺	あたし
オレ	あたし
僕	あたし
...	...
あなた	あんた
...	...

キャラB	
入力	出力
私	ピヨンタ
わたし	ピヨンタ
あたし	ピヨンタ
俺	ピヨンタ
オレ	ピヨンタ
僕	ピヨンタ
...	...
あなた	ソナタ
...	...

29

【図10】

キャラA	
入力	出力
表記	品詞
暖か*	形容詞語幹 活用語尾*
水っぽ*	形容詞語幹 活用語尾*
隅	名詞
	...
表記	品詞
ぬくと*	形容詞語幹 活用語尾*
しゃびしゃび	名詞形容動詞語幹
ぐろ	名詞
	...

キャラB	
入力	出力
表記	品詞
ゲロゲロッ!	感動詞
ゲロゲロッ!	感動詞
ゲロゲロッ!	感動詞
	...

30

【図11】

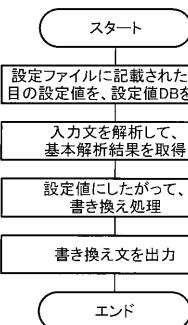
キャラA	
入力	出力
例	
な にや	なんで→にやんで
ぬ にゆ	ぬいぐるみ→にゆいぐるみ
の によ	その時→そによ時

31

キャラB	
入力	出力
例	
け ケロ	お化け→お化ケロ
ヶ ケロ	アップリケ→アップリケロ

キャラC	
入力	出力
例	
し ち	しました→ちました
る ゆ	する→すゆ

【図12】



---

フロントページの続き

(72)発明者 東中 竜一郎  
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

(72)発明者 牧野 俊朗  
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

(72)発明者 松尾 義博  
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

審査官 川 崎 博章

(56)参考文献 特開2014-191484(JP,A)  
特開2012-014311(JP,A)  
特開平10-254877(JP,A)

(58)調査した分野(Int.Cl., DB名)

G 06 F 17 / 20 - 17 / 28