

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2024年2月8日(08.02.2024)



(10) 国際公開番号
WO 2024/029152 A1

(51) 国際特許分類:
G10L 15/04 (2013.01) *G06F 40/166* (2020.01)

(21) 国際出願番号: PCT/JP2023/017568

(22) 国際出願日: 2023年5月10日(10.05.2023)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2022-125544 2022年8月5日(05.08.2022) JP

(71) 出願人: 株式会社 N T T ドコモ (NTT DOCOMO, INC.) [JP/JP]; 〒1006150 東京都千代田区永田町二丁目 1 1 番 1 号 Tokyo (JP).

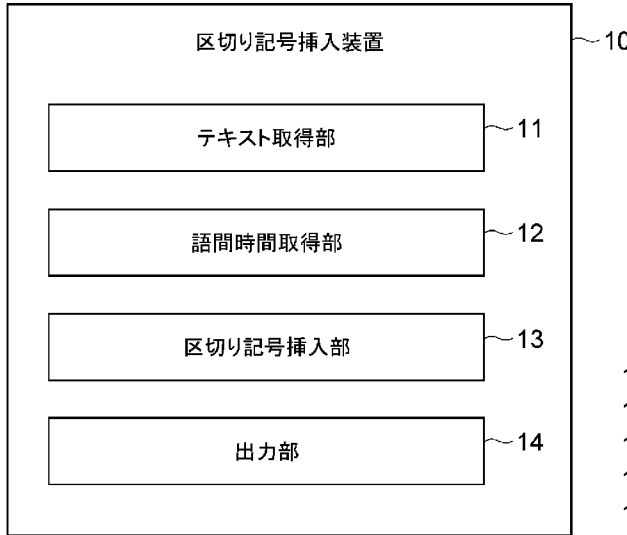
(72) 発明者: 竹 谷 謙 吾 (TAKETANI Kengo); 〒1006150 東京都千代田区永田町二丁目 1 1 番 1 号 山王パークタワー 株式会社 N T T ドコモ 知的財産部内 Tokyo (JP).

(74) 代理人: 長谷川 芳樹, 外 (HASEGAWA Yoshiki et al.); 〒1000005 東京都千代田区丸の内二丁目 1 番 1 号 丸の内 M Y P L A Z A (明治安田生命ビル) 9 階 創英国際特許法律事務所 Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR,

(54) Title: DELIMITER INSERTION DEVICE AND SPEECH RECOGNITION SYSTEM

(54) 発明の名称: 区切り記号挿入装置及び音声認識システム



10 Delimiter insertion device
11 Text acquisition unit
12 Word-spacing time acquisition unit
13 Delimiter insertion unit
14 Output unit

(57) Abstract: This delimiter insertion device comprises: a word-spacing time acquisition unit that acquires a word-spacing time, which is the length of time until the next word is uttered, for each word included in uttered speech; and a delimiter insertion unit that, on the basis of a delimiter insertion model and the word-spacing time, inserts a delimiter into target text, which is text obtained through speech recognition of the uttered speech. The delimiter insertion model outputs delimiter inference information indicating a delimiter in response to input of a delimiter removal sentence. The delimiter insertion unit inserts a delimiter into the target text on the basis of the delimiter inference information obtained by inputting the target text into the delimiter insertion model.

HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告 (条約第21条(3))

(57) 要約: 区切り記号挿入装置は、発話音声に含まれる各単語における次の単語が発話されるまでの時間の長さである語間時間を取得する語間時間取得部と、区切り記号挿入モデル及び語間時間に基づいて、発話音声の音声認識により得られたテキストである対象テキストに区切り記号を挿入する区切り記号挿入部とを備える。区切り記号挿入モデルは、区切り記号除去文の入力に応じて区切り記号を示す区切り記号推測情報を出力する。区切り記号挿入部は、対象テキストを区切り記号挿入モデルに入力することにより得られた区切り記号推測情報に基づいて、対象テキストに区切り記号を挿入する。

明 細 書

発明の名称：区切り記号挿入装置及び音声認識システム

技術分野

[0001] 本発明は、区切り記号挿入装置及び音声認識システムに関する。

背景技術

[0002] 音声認識により得られたテキストに句読点を挿入する技術が知られている。例えば、特許文献1には、統計に基づいて句読点がつけられたテキスト形態の訓練データにより訓練されたエンジンにより、テキストに句読点が挿入される技術が開示されている。

先行技術文献

特許文献

[0003] 特許文献1：特開2012-508903号公報

発明の概要

発明が解決しようとする課題

[0004] 一般的な音声認識エンジンでは、単語列からなるテキストを発話音声から取得した後に、文として妥当な位置に句読点等の区切り記号（d e l i m i t e r）が挿入される。テキスト情報のみの参照により区切り記号が挿入された場合に、その区切り記号の挿入位置が、文としては誤りではないものの、発話者の意図とは異なる場合があった。

[0005] そこで、本発明は、上記問題点に鑑みてなされたものであり、発話音声の音声認識処理により得られたテキストに対して、発話者の意図した通りの位置に区切り記号を挿入することを目的とする。

課題を解決するための手段

[0006] 上記課題を解決するために、本開示の一側面に係る区切り記号挿入装置は、発話音声の音声認識により得られたテキストに含まれる単語の後に、文を区切る区切り記号を挿入する区切り記号挿入装置であって、発話音声に含まれる各単語における次の単語が発話されるまでの時間の長さである語間時間

を取得する語間時間取得部と、区切り記号挿入モデル及び語間時間に基づいて、発話音声の音声認識により得られたテキストである対象テキストに区切り記号を挿入する区切り記号挿入部であって、区切り記号挿入モデルは、区切り記号を含まない文である区切り記号除去文を少なくとも入力とし、区切り記号除去文に含まれる各単語の後に挿入される区切り記号を示す区切り記号推測情報を出し、区切り記号除去文と区切り記号を含む文である区切り記号入り文とのペアを含む学習データを用いた機械学習により生成されるモデルであり、対象テキストを区切り記号除去文として区切り記号挿入モデルに入力することにより得られ、語間時間に応じて調整された区切り記号推測情報に基づいて、対象テキストに区切り記号を挿入する、区切り記号挿入部と、を備える。

[0007] 上記の側面によれば、発話音声における語間時間が取得される。語間時間には、発話時における発話者の意図が反映されている。そして、対象テキストを区切り記号挿入モデルに入力することにより得られ且つ語間情報に応じて調整された区切り記号推測情報が取得される。ここで取得された区切り記号推測情報は、発話音声における各単語の語間時間に応じて調整された情報であるので、発話者の意図が反映された区切り記号を示す。そして、区切り記号推測情報に基づいて対象テキストに区切り記号が挿入されることにより、発話者の意図に沿う位置に区切り記号が挿入されたテキストを得ることが可能となる。

発明の効果

[0008] 発話音声の音声認識処理により得られたテキストに対して、発話者の意図した通りの位置に区切り記号を挿入することが可能となる。

図面の簡単な説明

[0009] [図1]本実施形態の区切り記号挿入装置の機能的構成を示すブロック図である。

[図2]区切り記号挿入装置のハードブロック図である。

[図3]本実施形態の区切り記号挿入装置により解決される課題を説明する図で

ある。

[図4]対象テキストの取得処理を示す図である。

[図5]区切り記号挿入モデルの機会学習に用いられる学習データの第1の例を示す図である。

[図6]区切り記号挿入モデルの構成の第1の例を示す図である。

[図7]区切り記号推測情報を語間時間に基づいて調整するために参照される調整ルール情報の例を示す図である。

[図8]区切り記号推測情報の調整処理の例を示す図である。

[図9]語間時間の補正処理の例を示す図である。

[図10]区切り記号挿入モデルの機会学習に用いられる学習データの第2の例を示す図である。

[図11]区切り記号挿入モデルの構成の第2の例を示す図である。

[図12]区切り記号挿入モデルの機会学習に用いられる学習データの第3の例を示す図である。

[図13]本実施形態の音声認識システムの構成の例を示す機能ブロック図である。

[図14]区切り記号挿入装置における区切り記号挿入方法の処理内容を示すフローチャートである。

[図15]区切り記号挿入プログラムの構成を示す図である。

発明を実施するための形態

[0010] 本発明に係る区切り記号挿入装置及び音声認識システムの実施形態について図面を参照して説明する。なお、可能な場合には、同一の部分には同一の符号を付して、重複する説明を省略する。

[0011] 図1は、本実施形態に係る区切り記号挿入装置の機能的構成を示す図である。本実施形態の区切り記号挿入装置は、発話音声の音声認識により得られたテキストに含まれる単語の後に、文を区切る区切り記号（delimiter）を挿入する装置である。

[0012] 本実施形態では、区切り記号挿入装置10が、英文に挿入される区切り記

号であるピリオド (period)、カンマ (comma)、疑問符 (question mark) 等を、英語の文からなるテキストに挿入する場合の例を説明するが、この例には限定されない。区切り記号挿入装置 10 が、日本語の文に句点及び読点等の区切り記号を挿入する装置であってもよいし、その他の言語の文に、その言語における区切り記号を挿入する装置であってもよい。

[0013] 区切り記号挿入装置 10 は、図 1 に示すように、機能的には、テキスト取得部 11、語間時間取得部 12、区切り記号挿入部 13 及び出力部 14 を備える。これらの各機能部 11～14 は、一つの装置に構成されてもよいし、複数の装置に分散されて構成されてもよい。

[0014] なお、図 1 に示したブロック図は、機能単位のブロックを示している。これらの機能ブロック（構成部）は、ハードウェア及びソフトウェアの少なくとも一方の任意の組み合わせによって実現される。また、各機能ブロックの実現方法は特に限定されない。すなわち、各機能ブロックは、物理的又は論理的に結合した 1 つの装置を用いて実現されてもよいし、物理的又は論理的に分離した 2 つ以上の装置を直接的又は間接的に（例えば、有線、無線などを用いて）接続し、これら複数の装置を用いて実現されてもよい。機能ブロックは、上記 1 つの装置又は上記複数の装置にソフトウェアを組み合わせ実現されてもよい。

[0015] 機能には、判断、決定、判定、計算、算出、処理、導出、調査、探索、確認、受信、送信、出力、アクセス、解決、選択、選定、確立、比較、想定、期待、見做し、報知 (broadcasting)、通知 (notifying)、通信 (communicating)、転送 (forwarding)、構成 (configuring)、再構成 (reconfiguring)、割り当て (allocating、mapping)、割り振り (assigning) などがあるが、これらに限られない。たとえば、送信を機能させる機能ブロック（構成部）は、送信部 (transmitting unit) や送信機 (transmitter) と呼称される。いずれも、上述したとおり、実現方法は特に限定されない。

[0016] 例えば、本発明の一実施の形態における区切り記号挿入装置 10 は、コンピュータとして機能してもよい。図 2 は、本実施形態に係る区切り記号挿入

装置 10 のハードウェア構成の一例を示す図である。区切り記号挿入装置 10 は、物理的には、プロセッサ 1001、メモリ 1002、ストレージ 1003、通信装置 1004、入力装置 1005、出力装置 1006、バス 1007 などを含むコンピュータ装置として構成されてもよい。

[0017] なお、以下の説明では、「装置」という文言は、回路、デバイス、ユニットなどに読み替えることができる。区切り記号挿入装置 10 のハードウェア構成は、図に示した各装置を 1 つ又は複数含むように構成されてもよいし、一部の装置を含まずに構成されてもよい。

[0018] 区切り記号挿入装置 10 における各機能は、プロセッサ 1001、メモリ 1002 などのハードウェア上に所定のソフトウェア（プログラム）を読み込ませることで、プロセッサ 1001 が演算を行い、通信装置 1004 による通信や、メモリ 1002 及びストレージ 1003 におけるデータの読み出し及び／又は書き込みを制御することで実現される。

[0019] プロセッサ 1001 は、例えば、オペレーティングシステムを動作させてコンピュータ全体を制御する。プロセッサ 1001 は、周辺装置とのインターフェース、制御装置、演算装置、レジスタなどを含む中央処理装置（CPU: Central Processing Unit）で構成されてもよい。例えば、図 1 に示した各機能部 11～14 などは、プロセッサ 1001 で実現されてもよい。

[0020] また、プロセッサ 1001 は、プログラム（プログラムコード）、ソフトウェアモジュールやデータを、ストレージ 1003 及び／又は通信装置 1004 からメモリ 1002 に読み出し、これらに従って各種の処理を実行する。プログラムとしては、上述の実施の形態で説明した動作の少なくとも一部をコンピュータに実行させるプログラムが用いられる。例えば、区切り記号挿入装置 10 の各機能部 11～15 は、メモリ 1002 に格納され、プロセッサ 1001 で動作する制御プログラムによって実現されてもよい。上述の各種処理は、1 つのプロセッサ 1001 で実行される旨を説明してきたが、2 以上のプロセッサ 1001 により同時又は逐次に実行されてもよい。プロセッサ 1001 は、1 以上のチップで実装されてもよい。なお、プログラム

は、電気通信回線を介してネットワークから送信されても良い。

[0021] メモリ 1002 は、コンピュータ読み取り可能な記録媒体であり、例えば、ROM (Read Only Memory)、EPROM (Erasable Programmable ROM)、EEPROM (Electrically Erasable Programmable ROM)、RAM (Random Access Memory) などの少なくとも 1 つで構成されてもよい。メモリ 1002 は、レジスタ、キャッシュ、メインメモリ（主記憶装置）などと呼ばれてもよい。メモリ 1002 は、本発明の一実施の形態に係る疑似データ生成方法及び文生成方法を実施するために実行可能なプログラム（プログラムコード）、ソフトウェアモジュールなどを保存することができる。

[0022] ストレージ 1003 は、コンピュータ読み取り可能な記録媒体であり、例えば、CD-ROM (Compact Disc ROM) などの光ディスク、ハードディスクドライブ、フレキシブルディスク、光磁気ディスク（例えば、コンパクトディスク、デジタル多用途ディスク、Blu-ray（登録商標）ディスク）、スマートカード、フラッシュメモリ（例えば、カード、スティック、キードライブ）、フロッピー（登録商標）ディスク、磁気ストリップなどの少なくとも 1 つで構成されてもよい。ストレージ 1003 は、補助記憶装置と呼ばれてもよい。上述の記憶媒体は、例えば、メモリ 1002 及び／又はストレージ 1003 を含むデータベース、サーバその他の適切な媒体であってもよい。

[0023] 通信装置 1004 は、有線及び／又は無線ネットワークを介してコンピュータ間の通信を行うためのハードウェア（送受信デバイス）であり、例えばネットワークデバイス、ネットワークコントローラ、ネットワークカード、通信モジュールなどともいう。

[0024] 入力装置 1005 は、外部からの入力を受け付ける入力デバイス（例えば、キーボード、マウス、マイクロフォン、スイッチ、ボタン、センサなど）である。出力装置 1006 は、外部への出力を実施する出力デバイス（例えば、ディスプレイ、スピーカー、LEDランプなど）である。なお、入力装置 1

005 及び出力装置 1006 は、一体となった構成（例えば、タッチパネル）であってもよい。

[0025] また、プロセッサ 1001 やメモリ 1002 などの各装置は、情報を通信するためのバス 1007 で接続される。バス 1007 は、単一のバスで構成されてもよいし、装置間で異なるバスで構成されてもよい。

[0026] また、区切り記号挿入装置 10 は、マイクロプロセッサ、デジタル信号プロセッサ（DSP：Digital Signal Processor）、ASIC（Application Specific Integrated Circuit）、PLD（Programmable Logic Device）、FPGA（Field Programmable Gate Array）などのハードウェアを含んで構成されてもよく、当該ハードウェアにより、各機能ブロックの一部又は全てが実現されてもよい。例えば、プロセッサ 1001 は、これらのハードウェアの少なくとも 1 つで実装されてもよい。

[0027] 図 3 を参照して、本実施形態の区切り記号挿入装置 10 により解決される課題を説明する。図 3 に示される発話音声 sp11, sp21 はそれぞれ、異なる意図により発話者により発話されたものである。発話音声 sp11, sp21 の音声認識により得られた音声認識結果 sr11, sr21 は、同一のテキスト「I know it's been there forever」である。発話音声 sp11 において、単語「know」の次の単語が発話されるまでの語間時間は 1.2 秒である。一方、発話音声 sp21 において、単語「know」の次の単語が発話されるまでの語間時間は 0.1 秒である。

[0028] 従来の区切り記号を挿入する技術では、音声認識結果として得られたテキストに、文として妥当な位置に区切り記号が挿入されるので、音声認識結果 sr11, sr21 のそれぞれに、従来の区切り記号挿入技術により区切り記号を挿入して得られた音声認識結果 sr12, sr22 は、元となる発話音声異なるにも関わらず同一である。

[0029] 音声認識結果 sr22 は、発話音声 sp21 の発話者の意図のとおり、単語「forever」の後にピリオドを有する。これに対して、発話音声

s p 1 1 が単語「k n o w」の後にピリオドが挿入されることを意図されたものであるにも関わらず、音声認識結果 s r 1 2 は、単語「f o r e v e r」の後にピリオドを有する。このピリオドの位置は、発話者の意図する位置とは異なる。

[0030] 本実施形態の区切り記号挿入装置 1 0 は、発話音声における語間時間を用いて区切り記号を挿入するので、発話音声 s p 1 1 に基づいて得られた音声認識結果 s r 1 1 に対して、単語「k n o w」及び単語「f o r e v e r」のそれぞれの後の位置にピリオドを挿入して音声認識結果 s r 1 3 を得る。一方、区切り記号挿入装置 1 0 は、発話音声 s p 2 1 に基づいて得られた音声認識結果 s r 2 1 に対して、単語「f o r e v e r」の後の位置にピリオドを挿入して音声認識結果 s r 2 3 を得る。音声認識結果 s r 1 3, s r 2 3 は、発話音声 s p 1 1, s p 2 1 の発話者の意図した通りの位置に区切り記号を有する。

[0031] 次に、区切り記号挿入装置 1 0 の各機能部について説明する。テキスト取得部 1 1 は、区切り記号を挿入する対象のテキストである対象テキストを取得する。対象テキストは、発話音声の音声認識により得られたテキストである。語間時間取得部 1 2 は、発話音声に含まれる各単語における次の単語が発話されるまでの時間の長さである語間時間を取得する。

[0032] 図 4 は、対象テキスト及び語間時間の取得の例を示す図である。テキスト取得部 1 1 は、発話音声 s p 3 に基づいて対象テキスト t x 1 「I k n o w i t ' s b e e n t h e r e f o r e v e r」を取得する。テキスト取得部 1 1 は、周知の音声認識処理技術及びその他の技術により、発話音声の音声認識を行うことにより、対象テキストを取得してもよい。

[0033] 語間時間取得部 1 2 は、対象テキストに含まれる各単語における、次の単語が発話されるまでの時間の長さを語間時間 i t として取得する。語間時間取得部 1 2 は、発話音声 s p 3 の音声認識時における無音時間を語間時間 i t として取得してもよい。無音時間は、例えば、音量が所定の程度未満である時間である。また、語間時間取得部 1 2 は、発話音声の音声認識に用いら

れる音声認識エンジンから、音声認識の度に逐次的に音声認識結果を取得し、音声認識結果の取得及び更新の時間間隔を単語の発生の間の擬似的な無音時間とみなして、更新時間間隔を語間時間 i_t として取得してもよい。

[0034] 区切り記号挿入部 13 は、区切り記号挿入モデル及び語間時間に基づいて、対象テキストに区切り記号を挿入する。具体的には、区切り記号挿入部 13 は、対象テキストを区切り記号挿入モデルに入力することにより得られる区切り記号推測情報に基づいて、対象テキストに区切り記号を挿入する。区切り記号推測情報は、語間時間に応じて調整された情報を含む。

[0035] 区切り記号挿入モデルは、区切り記号を含まない文である区切り記号除去文を少なくとも入力とし、区切り記号除去文に含まれる各単語の後に挿入される区切り記号を示す区切り記号推測情報を出力する。また、区切り記号挿入モデルは、区切り記号除去文と区切り記号を含む文である区切り記号入り文とのペアを含む学習データを用いた機械学習により生成される。

[0036] 図 5 は、区切り記号挿入モデルの機会学習に用いられる学習データの第 1 の例を示す図である。図 6 は、区切り記号挿入モデルの構成の第 1 の例を示す図である。

[0037] 図 5 に示されるように、区切り記号挿入モデル $md1$ の機会学習に用いられる学習データの一例である学習データ $td1$ は、区切り記号除去文 $id1$ と区切り記号入り文 $od1$ とのペアからなる。

[0038] 区切り記号入り文 $od1$ は、文を構成する単語列及び各単語の後に挿入される区切り記号を示すラベルである区切り記号ラベルを含む。区切り記号を示すラベルは、図 5 等において、模式的に以下のように図示される。

<O>…区切り記号なし

<P>…ピリオド、句点

<C>…カンマ、読点

<Q>…クエスチョンマーク、疑問符

区切り記号除去文 $id1$ は、区切り記号入り文 $od1$ から区切り記号ラベルを除去した文であってもよい。

- [0039] 区切り記号挿入モデルm d 1の機会学習では、学習過程の区切り記号挿入モデルm d 1に区切り記号除去文i d 1を入力し、区切り記号挿入モデルm d 1から得られた出力と教師データである区切り記号入り文o d 1との誤差に基づいて、区切り記号挿入モデルm d 1を構成する重み及びパラメータ等が更新される。
- [0040] 図6に示されるように、学習済みの区切り記号挿入モデルm d 1は、区切り記号除去文s d 1の入力に応じて、区切り記号推測情報d p 1を出力する。
- [0041] 区切り記号挿入モデルm d 1は、ニューラルネットワークを含んで構成されるモデルであってもよい。さらに具体的には、区切り記号挿入モデルm d 1は、入力された文に含まれる各単語の後に挿入される区切り記号を予測する系列ラベリングタスクを解決する系列ラベリングモデルとして構成されてもよい。
- [0042] 学習済みのニューラルネットワークを含むモデルである区切り記号挿入モデルm d 1は、コンピュータにより読み込まれ又は参照され、コンピュータに所定の処理を実行させ及びコンピュータに所定の機能を実現させるプログラムとして捉えることができる。
- [0043] 即ち、本実施形態の学習済みの区切り記号挿入モデルm d 1は、CPU及びメモリを備えるコンピュータにおいて用いられる。具体的には、コンピュータのCPUが、メモリに記憶された学習済みの区切り記号挿入モデルm d 1からの指令に従って、ニューラルネットワークの入力層に入力された入力データに対し、各層に対応する学習済みの重み付け係数（パラメータ）と応答関数等に基づく演算を行い、出力層から結果（確率）を出力するよう動作する。
- [0044] 区切り記号推測情報d p 1は、区切り記号除去文s d 1に含まれる各単語の後に挿入されうる各種の区切り記号の尤度である記号挿入尤度、及び、各単語の後に区切り記号を未挿入とすることに関する尤度である記号なし尤度を含む。そして、記号挿入尤度及び記号なし尤度のうちの最大の尤度に基づ

いて、各単語の後に、複数の種類のうちのいずれかの区切り記号を挿入すること又は区切り記号を未挿入とすることが決定（ラベリング）される。

[0045] 図6に例示される区切り記号推測情報d p 1は、以下のとおり、単語「I」に関する各区切り記号の記号挿入尤度及び記号なし尤度を含む。

<O> : 90%、<C> : 5%、<P> : 2%、<Q> : 3%

従って、区切り記号を未挿入とすること（ラベル<O>）の記号なし尤度が最大であるので、単語「I」に対して区切り記号なしのラベル<O>がラベリングされる。

[0046] 同様に、区切り記号推測情報d p 1において、以下のとおり、単語「know」に関する各区切り記号の記号挿入尤度及び記号なし尤度を含む。

<O> : 60%、<C> : 5%、<P> : 30%、<Q> : 5%

従って、区切り記号を未挿入とすること（ラベル<O>）の記号なし尤度が最大であるので、単語「know」に対して区切り記号なしのラベル<O>がラベリングされる。

[0047] 区切り記号挿入部13は、区切り記号推測情報の語間時間による調整の一例として、区切り記号挿入モデルから出力された区切り記号推測情報を、語間時間に基づいて調整してもよい。具体的には、区切り記号挿入部13は、区切り記号推測情報に含まれる記号挿入尤度及び記号なし尤度を、語間時間に基づいて調整してもよい。

[0048] 区切り記号挿入部13は、対象テキストに含まれる単語のうちの一の単語における語間時間が第1の時間である場合に、当該一の単語の記号なし尤度を上げるように調整し、又は／及び、複数の種類の区切り記号のうち少なくとも一種の当該一の単語の区切り記号の記号挿入尤度を下げるように調整し、当該一の単語における語間時間が第1の時間より長い第2の時間である場合に、複数の種類の区切り記号のうち少なくとも一種の当該一の単語の区切り記号の記号挿入尤度を上げるように調整し、又は／及び、当該一の単語の記号なし尤度を下げるように調整してもよい。

[0049] そして、区切り記号挿入部13は、調整された記号挿入尤度及び記号なし

尤度のうちの最大の尤度に基づいて、当該一の単語の後に、複数の種類のうちのいずれかの区切り記号を挿入し、又は区切り記号を未挿入とする。

[0050] このような記号挿入尤度及び記号なし尤度の調整のために、区切り記号挿入部 13 は、例えば、調整ルール情報を参照して尤度を調整してもよい。図 7 は、区切り記号推測情報を語間時間に基づいて調整するために参照される調整ルール情報の例を示す図である。調整ルール情報は、区切り記号挿入部 13 がアクセス可能な記憶手段に記憶されていてもよいし、区切り記号挿入部 13 の内部にテーブルとして備えられていてもよい。

[0051] 図 7 に示されるように、調整ルール情報は、語間時間の範囲ごとに、語間時間の長さを表す語間時間カテゴリ及び尤度の調整の内容を示す尤度調整情報が関連付けられた情報である。例えば、語間時間 (x) が 0.1 秒以下である場合には、語間時間カテゴリは「無」であり、「記号なし尤度を 50% 上げる」ことが尤度の調整として実施されることが調整ルールとして規定されている。また、例えば、語間時間 (x) が 0.5 秒より長く且つ 1.0 秒以下である場合には、語間時間カテゴリは「中」であり、「記号挿入尤度を 50% 上げる」ことが尤度の調整として実施されることが調整ルールとして規定されている。

[0052] 図 8 は、区切り記号推測情報の調整処理の例を示す図である。図 8 に示される区切り記号推測情報 d p 2 1 は、区切り記号挿入モデル m d 1 から出力された調整処理前の区切り記号推測情報の一部を示す。

[0053] 区切り記号推測情報 d p 2 1 は、単語「k n o w」に関する各区切り記号の記号挿入尤度 l h 2 1 及び記号なし尤度を含む。調整前の区切り記号推測情報 d p 2 1 に従うならば、区切り記号を未挿入とすること（ラベル<0>）の記号なし尤度が最大であるので、単語「k n o w」に対して区切り記号なしのラベル<0>がラベリングされる。

[0054] 区切り記号挿入部 13 は、単語「k n o w」の語間時間 i t 2 に基づいて、区切り記号推測情報 d p 2 1 の各尤度を調整する。具体的には、語間時間取得部 12 により取得された単語「k n o w」の語間時間 i t 2 は 0.9 秒

であるので、区切り記号挿入部 13 は、調整ルール情報（図 7）を参照して、語間時間 0.9 秒に関連付けられた尤度調整情報「ピリオド、カンマ、クエスチョンマークの記号挿入尤度を 50% 上げる」を取得し、カンマ、ピリオド及びクエスチョンマークの記号挿入尤度 1h21 を、取得した尤度調整情報に従って調整する。

[0055] そして、区切り記号挿入部 13 は、記号挿入尤度 1h21 のそれぞれの値を 50% 増加させて、調整後の区切り記号推測情報 dp22 を得る。区切り記号推測情報 dp22 において、ピリオド（ラベル<P>）の記号挿入尤度が最大であるので、区切り記号挿入部 13 は、単語「know」に対して区切り記号「ピリオド」のラベル<P>をラベリングする。そして、区切り記号挿入部 13 は、対象テキストに含まれる単語「know」の後に、ラベリングされたラベル<P>に基づいて、ピリオドを挿入する。

[0056] このように、対象テキストに含まれる一の単語の語間時間が長いほど、記号挿入尤度が上げられ、又は／及び、記号なし尤度が下げられるように相対的に調整され、一の単語の語間時間が短いほど、記号挿入尤度が下げられ、又は／及び、記号なし尤度が上げられるように相対的に調整されるので、発話者の意図が、記号挿入尤度及び記号なし尤度に反映される。そして、調整された記号挿入尤度及び記号なし尤度に基づいて、区切り記号が挿入され又は区切り記号が未挿入とされるので、発話者の意図に沿う適切な位置に区切り記号が挿入されたテキストを得ることが可能となる。

[0057] 区切り記号挿入部 13 は、区切り記号推測情報の調整に用いるための語間時間を所定の条件に従って補正してもよい。図 9 は、語間時間の補正処理の例を示す図である。区切り記号挿入部 13 は、対象テキストに含まれる全単語のうちの一以上の単語の語間時間の長さの程度が所与の程度より長い場合に、全単語のそれぞれの前記語間時間を所与の程度で短くなるように補正してもよい。そして、区切り記号挿入部 13 は、補正された語間時間である補正語間時間に基づいて、記号挿入尤度及び／又は記号なし尤度を調整してもよい。

[0058] 図9に示される対象テキスト $t \times 31$ は、補正前の語間時間 $i t 31$ を含む。ここでは、一例として、対象テキストに含まれる全ての単語のうち、半数以上の単語の語間時間が、語間時間カテゴリ「小」に対応する語間時間以上である場合に、全ての単語の語間時間を0.5秒減ずることが、補正処理の内容として予め設定されていることとする。

[0059] この場合において、区切り記号挿入部13は、対象テキスト $t \times 31$ に含まれる単語の語間時間 $i t 31$ のうちの全ての単語の語間時間が、語間時間カテゴリ「小」に対応する語間時間以上であることを判定する。そして、区切り記号挿入部13は、対象テキスト $t \times 32$ に示されるように、語間時間 $i t 31$ をそれぞれ0.5秒減じて、補正後の語間時間である補正語間時間 $i t 32$ を得る。区切り記号挿入部13は、補正語間時間 $i t 32$ に基づいて、対象テキスト $t \times 32$ に含まれる各単語の記号挿入尤度及び／又は記号なし尤度を調整する。

[0060] 発話者の発話が、全体的に語間時間が長くなるような傾向を有する場合には、区切り記号の挿入後のテキストにおいて、発話者が意図しない位置に過剰に区切り記号が挿入される可能性があるところ、上記の語間時間の補正処理によれば、語間時間の長さが所定の程度より長い場合に、語間時間が短くなるように補正された補正語間時間に基づいて記号挿入尤度及び／又は記号なし尤度が調整されるので、発話者の意図が好適に反映された適切な位置に区切り記号が挿入されたテキストを得ることが可能となる。

[0061] 次に、区切り記号推測情報の語間時間に応じた調整処理の第2の例を説明する。図10は、区切り記号挿入モデルの機会学習に用いられる学習データの第2の例を示す図である。図11は、区切り記号挿入モデルの構成の第2の例を示す図である。

[0062] 図10に示されるように、区切り記号挿入モデル $md2$ の機会学習に用いられる学習データの一例である学習データ $td2$ は、区切り記号除去文 $id2$ と区切り記号入り文 $od2$ とのペアからなる。

[0063] 区切り記号入り文 $od2$ は、図5を参照して説明した区切り記号入り文 o

d 1 と同様に、文を構成する単語列及び各単語の後に挿入される区切り記号を示す区切り記号ラベルを含む。区切り記号除去文 i d 2 は、区切り記号入り文 o d 2 から区切り記号ラベルが除去された文、及び、当該文を構成する各単語にそれぞれ関連付けられた語間時間 i t 4 を含む。

[0064] 区切り記号挿入モデル m d 2 の機会学習では、学習過程の区切り記号挿入モデル m d 2 に区切り記号除去文 i d 2 を入力し、区切り記号挿入モデル m d 2 から得られた出力と教師データである区切り記号入り文 o d 2 との誤差に基づいて、区切り記号挿入モデル m d 2 を構成する重み及びパラメータ等が更新される。

[0065] 区切り記号挿入モデル m d 2 は、ニューラルネットワークを含んで構成されるモデルであってもよい。さらに具体的には、区切り記号挿入モデル m d 2 は、入力された文に含まれる各単語の後に挿入される区切り記号を予測する系列ラベリングタスクを解決する系列ラベリングモデルとして構成されてもよい。

[0066] 図 1 1 に示されるように、学習済みの区切り記号挿入モデル m d 2 は、区切り記号除去文 s d 2 の入力に応じて、区切り記号推測情報 d p 2 を出力する。区切り記号除去文 s d 2 は、区切り記号除去文 s d 2 を構成する各単語に関連付けられた語間時間 i t 5 を含む。区切り記号推測情報 d p 2 は、区切り記号除去文 s d 2 に含まれる各単語に関する記号挿入尤度及び記号なし尤度を含む。

[0067] 学習済みのニューラルネットワークを含むモデルである区切り記号挿入モデル m d 2 は、コンピュータにより読み込まれ又は参照され、コンピュータに所定の処理を実行させ及びコンピュータに所定の機能を実現させるプログラムとして捉えることができる。

[0068] 即ち、本実施形態の学習済みの区切り記号挿入モデル m d 2 は、C P U 及びメモリを備えるコンピュータにおいて用いられる。具体的には、コンピュータの C P U が、メモリに記憶された学習済みの区切り記号挿入モデル m d 2 からの指令に従って、ニューラルネットワークの入力層に入力された入力

データに対し、各層に対応する学習済みの重み付け係数（パラメタ）と応答関数等に基づく演算を行い、出力層から結果（確率）を出力するよう動作する。

[0069] 図6を参照して説明した例では、区切り記号挿入モデルm d 1から出力された区切り記号推測情報d p 1に含まれる記号挿入尤度及び記号なし尤度が、語間時間に基づいて調整される。これに対して、区切り記号挿入モデルm d 2では、語間時間i t 4を含む区切り記号除去文i d 2が機会学習における学習データとして入力に用いられ、語間時間i t 5を含む区切り記号除去文s d 2の入力に応じて区切り記号推測情報d p 2を出力するので、区切り記号推測情報d p 2に含まれる記号挿入尤度及び記号なし尤度は、区切り記号挿入モデルm d 2における演算により、語間時間に応じた調整がされた値である。

[0070] 区切り記号挿入モデルm d 2において、以下のような、単語「I」に関する各区切り記号の記号挿入尤度及び記号なし尤度が算出される。

<O> : 90%、<C> : 5%、<P> : 2%、<Q> : 3%

従って、区切り記号推測情報d p 2は、単語「I」に関する各区切り記号の記号挿入尤度及び記号なし尤度のうちの最大の尤度を有する区切り記号なしのラベル<O>が、単語「I」に対してラベリングされたことを示す情報（I=<O>）を含む。そして、区切り記号挿入部13は、対象テキストに含まれる単語「I」の後に、ラベリングされたラベル<O>に基づいて、区切り記号を未挿入とすることを決定する。

[0071] また、区切り記号挿入モデルm d 2において、以下のような、単語「know」に関する各区切り記号の記号挿入尤度及び記号なし尤度が算出される。

<O> : 10%、<C> : 20%、<P> : 70%、<Q> : 0%

図6に例示した区切り記号推測情報d p 1では、区切り記号を未挿入とすること（ラベル<O>）の記号なし尤度が最大であったのに対して、区切り記号挿入モデルm d 2により算出された各尤度においては、ピリオドを挿入す

ること（ラベル＜P＞）の記号挿入尤度が最大である。従って、区切り記号推測情報 d p 2 は、ピリオドを挿入することが単語「k n o w」に対してラベリングされたことを示す情報（k n o w＝＜P＞）を含む。そして、区切り記号挿入部 1 3 は、対象テキストに含まれる単語「k n o w」の後に、ラベリングされたラベル＜P＞に基づいて、ピリオドを挿入する。

[0072] このように、語間時間が各単語に関連付けられた区切り記号除去文が学習データとして用いられた機会学習により区切り記号挿入モデル m d 2 が生成され、各単語の語間時間を含む区切り記号除去文が区切り記号挿入モデル m d 2 の入力とされるので、各単語に語間時間が関連付けられた対象テキストを区切り記号挿入モデルに入力することにより、区切り記号挿入モデルからの出力に対する語間時間に基づく別途の調整処理を行うことなく、語間時間により調整された区切り記号推測情報を得ることが可能となる。そして、区切り記号推測情報に基づいて対象テキストに区切り記号が挿入されることにより、発話者の意図に沿う位置に区切り記号が挿入されたテキストを容易に得ることが可能となる。

[0073] 次に、区切り記号挿入モデルの他の例について説明する。図 1 2 は、区切り記号挿入モデルの機会学習に用いられる学習データの第 3 の例を示す図である。図 1 2 に示される例では、学習データ t d 3 は、区切り記号除去文 i d 3 と区切り記号入り文 o d 3 とのペアからなる。

[0074] 区切り記号入り文 o d 3 は、図 5 及び図 1 0 を参照して説明した区切り記号入り文 o d 1 , o d 2 と同様に、文を構成する単語列及び各単語の後に挿入される区切り記号を示す区切り記号ラベルを含む。区切り記号除去文 i d 3 は、区切り記号入り文 o d 3 から区切り記号ラベルが除去された文、及び、当該文を構成する各単語にそれぞれ関連付けられた語間時間に加えて、シチュエーション情報 s t 3 を含む。シチュエーション情報 s t 3 は、発話音声が発話されたときの状況を示す情報であって、タグとして区切り記号除去文 i d 3 に付与されてもよい。

[0075] シチュエーション情報 s t 3 は、以下のような、発話音声が発話されたと

きの状況に応じたバリエーションを有してもよい。

会議：＜M e e t i n g＞

講演：＜L e c t u r e＞

雑談：＜C h a t t i n g＞

[0076] 学習データ $t d 3$ を用いた区切り記号挿入モデルの機会学習では、学習過程の区切り記号挿入モデルに区切り記号除去文 $i d 3$ を入力し、区切り記号挿入モデルから得られた出力と教師データである区切り記号入り文 $o d 3$ との誤差に基づいて、区切り記号挿入モデルを構成する重み及びパラメータ等が更新される。

[0077] 学習データ $t d 3$ を用いた機械学習による学習済みの区切り記号挿入モデルは、シチュエーション情報及び各単語に関連付けられた語間時間を含む区切り記号除去文（対象テキスト）の入力に応じて、区切り記号推測情報を出力する。出力された区切り記号推測情報は、区切り記号除去文に含まれる各単語に関する記号挿入尤度及び記号なし尤度、並びに、最大の尤度に基づく区切り記号又は区切り記号未挿入のラベルを含む。

[0078] 区切り記号挿入部 1 3 は、シチュエーション情報が関連付けられた対象テキストを区切り記号挿入モデルに入力することに応じて得られた区切り記号推測情報に基づいて、各単語の後にラベルの応じた区切り記号を挿入し、又は、区切り記号を未挿入とすることを決定する。

[0079] このように、シチュエーション情報が関連付けられた区切り記号除去文が学習データとして用いられた機会学習により区切り記号挿入モデルが生成され、シチュエーション情報を含む区切り記号除去文が区切り記号挿入モデルの入力とされるので、シチュエーション情報が関連付けられた対象テキストを区切り記号挿入モデルに入力することにより、発話音声が発話されたときの状況に応じた発話の傾向が考慮された区切り記号推測情報を得ることが可能となる。

[0080] 再び図 1 を参照して、出力部 1 4 は、区切り記号挿入部 1 3 により区切り記号が挿入された対象テキストである区切り記号挿入テキストを出力する。

出力の態様は限定されず、出力部 14 は、区切り記号挿入テキストを所定の表示手段に表示させてもよいし、所定の記憶手段に区切り記号挿入テキストを記憶させてもよいし、所定の装置に区切り記号挿入テキストを送信してもよい。

[0081] 図 13 は、本実施形態の音声認識システムの構成の例を示す機能ブロック図である。図 13 に示されるように、音声認識システム 20 は、区切り記号挿入装置 10 を含んで構成され、音声認識結果取得部 21 及び音声認識結果出力部 22 を備える。

[0082] 音声認識結果取得部 21 は、音声認識エンジンにより音声認識された結果のテキスト及び当該テキストにおける各単語の語間時間を第 1 の音声認識結果として取得する。

[0083] 区切り記号挿入装置 10 のテキスト取得部 11 は、第 1 の音声認識結果に含まれるテキストを対象テキストとして取得する。区切り記号挿入装置 10 の語間時間取得部 12 は、第 1 の音声認識結果に含まれる語間時間を取得する。

[0084] 区切り記号挿入装置 10 の区切り記号挿入部 13 は、第 1 の音声認識結果に含まれるテキストを対象テキストとして、区切り記号を挿入する処理を実施する。

[0085] 音声認識結果出力部 22 は、区切り記号挿入部 13 により区切り記号が挿入された対象テキストを、第 2 の音声認識結果として出力する。

[0086] 音声認識システム 20 によれば、音声認識処理において語間時間が考慮されない音声認識エンジンから得られた第 1 の音声認識結果に基づいて、第 1 の音声認識結果に含まれるテキストを対象テキストとして、第 1 の音声認識結果に含まれる語間時間に応じて調整された区切り記号推測情報を取得できるので、発話者の意図に沿う位置に区切り記号が挿入された対象テキストを含む第 2 の音声認識結果を得ることが可能となる。

[0087] 図 14 は、区切り記号挿入装置 10 における区切り記号挿入方法の処理内容を示すフローチャートである。

- [0088] ステップS 1において、テキスト取得部 1 1 は、発話音声の音声認識により得られたテキストであって、区切り記号を挿入する対象のテキストである対象テキストを取得する。
- [0089] ステップS 2において、語間時間取得部 1 2 は、発話音声における各単語の語間時間を取得する。
- [0090] ステップS 3において、区切り記号挿入部 1 3 は、区切り記号挿入モデルに対象テキストを入力する。
- [0091] ステップS 4において、区切り記号挿入部 1 3 は、区切り記号推測情報を取得する。区切り記号推測情報は、語間時間に応じた調整された各単語の記号挿入尤度及び記号なし尤度に基づく、各単語に対するいずれかの区切り記号を挿入すること又は区切り記号を未挿入とすることを示すラベルを含む。
- [0092] ステップS 5において、区切り記号挿入部 1 3 は、区切り記号推測情報に基づいて、対象テキストに区切り記号を挿入する。
- [0093] ステップS 6において、出力部 1 4 は、区切り記号が挿入された対象テキストである区切り記号挿入テキストを出力する。
- [0094] 次に、図 1 5 を参照して、コンピュータを、本実施形態の区切り記号挿入装置 1 0 として機能させるための区切り記号挿入プログラムについて説明する。図 1 5 は、区切り記号挿入プログラムの構成を示す図である。区切り記号挿入プログラム P 1 は、区切り記号挿入装置 1 0 における区切り記号挿入処理を統括的に制御するメインモジュール m 1 0、テキスト取得モジュール m 1 1、語間時間取得モジュール m 1 2、区切り記号挿入モジュール m 1 3 及び出力モジュール m 1 4 を備えて構成される。そして、各モジュール m 1 1 ~ m 1 4 により、テキスト取得部 1 1、語間時間取得部 1 2、区切り記号挿入部 1 3 及び出力部 1 4 のための各機能が実現される。
- [0095] なお、区切り記号挿入プログラム P 1 は、通信回線等の伝送媒体を介して伝送される態様であってもよいし、図 1 5 に示されるように、記録媒体 M 1 に記憶される態様であってもよい。
- [0096] 以上説明した本実施形態の区切り記号挿入装置 1 0、区切り記号挿入方法

、区切り記号挿入プログラム P 1 によれば、発話音声における語間時間が取得される。語間時間には、発話時における発話者の意図が反映されている。そして、対象テキストを区切り記号挿入モデルに入力することにより得られ且つ語間情報に応じて調整された区切り記号推測情報が取得される。ここで取得された区切り記号推測情報は、発話音声における各単語の語間時間に応じて調整された情報であるので、発話者の意図が反映された区切り記号を示す。そして、区切り記号推測情報に基づいて対象テキストに区切り記号が挿入されることにより、発話者の意図に沿う位置に区切り記号が挿入されたテキストを得ることが可能となる。

[0097] 本開示に係る発明は、例えば、以下のように把握される。

[0098] 本開示の第 1 の一側面に係る区切り記号挿入装置は、発話音声の音声認識により得られたテキストに含まれる単語の後に、文を区切る区切り記号を挿入する区切り記号挿入装置であって、前記発話音声に含まれる各単語における次の単語が発話されるまでの時間の長さである語間時間を取得する語間時間取得部と、区切り記号挿入モデル及び前記語間時間に基づいて、前記発話音声の音声認識により得られたテキストである対象テキストに区切り記号を挿入する区切り記号挿入部であって、前記区切り記号挿入モデルは、前記区切り記号を含まない文である区切り記号除去文を少なくとも入力とし、前記区切り記号除去文に含まれる各単語の後に挿入される区切り記号を示す区切り記号推測情報を出力し、前記区切り記号除去文と区切り記号を含む文である区切り記号入り文とのペアを含む学習データを用いた機械学習により生成されるモデルであり、前記対象テキストを前記区切り記号除去文として前記区切り記号挿入モデルに入力することにより得られ、前記語間時間に応じて調整された前記区切り記号推測情報に基づいて、前記対象テキストに区切り記号を挿入する、区切り記号挿入部と、を備える。

[0099] 上記の側面によれば、発話音声における語間時間が取得される。語間時間には、発話時における発話者の意図が反映されている。そして、対象テキストを区切り記号挿入モデルに入力することにより得られ且つ語間情報に応じ

て調整された区切り記号推測情報が取得される。ここで取得された区切り記号推測情報は、発話音声における各単語の語間時間に応じて調整された情報であるので、発話者の意図が反映された区切り記号を示す。そして、区切り記号推測情報に基づいて対象テキストに区切り記号が挿入されることにより、発話者の意図に沿う位置に区切り記号が挿入されたテキストを得ることが可能となる。

[0100] 第2の側面に係る区切り記号挿入装置では、第1の側面に係る区切り記号挿入装置において、前記区切り記号挿入部は、前記区切り記号挿入モデルから出力された前記区切り記号推測情報を、前記語間時間に基づいて調整することとしてもよい。

[0101] 上記の側面によれば、区切り記号推測情報に、語間時間に表された発話者の意図を確実に反映させることが可能となる。

[0102] 第3の側面に係る区切り記号挿入装置では、第2の側面に係る区切り記号挿入装置において、前記区切り記号推測情報は、前記区切り記号除去文に含まれる各単語の後に複数の種類の区切り記号のそれぞれを挿入することに関する尤度である記号挿入尤度、及び、各単語の後に区切り記号を未挿入とすることに関する尤度である記号なし尤度を含み、前記区切り記号挿入部は、前記対象テキストに含まれる単語のうちの一の単語における前記語間時間が第1の時間である場合に、前記一の単語の前記記号なし尤度を上げるように調整し、又は／及び、複数の種類の区切り記号のうち少なくとも一種の前記一の単語の区切り記号の前記記号挿入尤度を下げるように調整し、前記一の単語における前記語間時間が前記第1の時間より長い第2の時間である場合に、複数の種類の区切り記号のうち少なくとも一種の前記一の単語の区切り記号の前記記号挿入尤度を上げるように調整し、又は／及び、前記一の単語の前記記号なし尤度を下げるように調整し、前記記号挿入尤度及び前記記号なし尤度のうちの最大の尤度に基づいて、前記一の単語の後に、複数の種類のうちのいずれかの区切り記号を挿入し、又は区切り記号を未挿入とすることとしてもよい。

- [0103] 上記の側面によれば、対象テキストに含まれる一の単語の語間時間が長いほど、記号挿入尤度が上げられ、又は／及び、記号なし尤度が下げられるように相対的に調整され、一の単語の語間時間が短いほど、記号挿入尤度が下げられ、又は／及び、記号なし尤度が上げられるように相対的に調整されるので、発話者の意図が、記号挿入尤度及び記号なし尤度に反映される。そして、調整された記号挿入尤度及び記号なし尤度に基づいて、区切り記号が挿入され又は区切り記号が未挿入とされるので、発話者の意図に沿う適切な位置に区切り記号が挿入されたテキストを得ることが可能となる。
- [0104] 第４の側面に係る区切り記号挿入装置では、第３の側面に係る区切り記号挿入装置において、前記区切り記号挿入部は、前記対象テキストに含まれる全単語のうちの一以上の単語の前記語間時間の長さの程度が、所与の程度より長い場合に、前記全単語のそれぞれの前記語間時間を所与の程度で短くなるように補正した補正語間時間に基づいて、前記記号挿入尤度及び／又は前記記号なし尤度を調整することとしてもよい。
- [0105] 発話者の発話が、全体的に語間時間が長くなるような傾向を有する場合には、区切り記号の挿入後のテキストにおいて、発話者が意図しない位置に過剰に区切り記号が挿入される可能性があるところ、上記の側面によれば、語間時間の長さが所定の程度より長い場合に、語間時間が短くなるように補正された補正語間時間に基づいて記号挿入尤度及び／又は記号なし尤度が調整されるので、発話者の意図が好適に反映された適切な位置に区切り記号が挿入されたテキストを得ることが可能となる。
- [0106] 第５の側面に係る区切り記号挿入装置では、第１の側面に係る区切り記号挿入装置において、前記区切り記号挿入モデルは、前記区切り記号除去文に含まれる各単語の語間時間を入力として更に含み、前記語間時間が各単語に関連付けられた前記区切り記号除去文と前記区切り記号入り文とのペアからなる学習データを用いた機械学習により生成され、前記語間時間により調整された前記区切り記号推測情報を出し、前記区切り記号挿入部は、各単語に前記語間時間が関連付けられた前記対象テキストを前記区切り記号挿入モ

デルに入力することとしてもよい。

[0107] 上記の側面によれば、語間時間が各単語に関連付けられた区切り記号除去文が学習データとして用いられた機会学習により区切り記号挿入モデルが生成され、各単語の語間時間を含む区切り記号除去文が区切り記号挿入モデルの入力とされるので、各単語に語間時間が関連付けられた対象テキストを区切り記号挿入モデルに入力することにより、区切り記号挿入モデルからの出力に対する語間時間に基づく別途の調整処理を行うことなく、語間時間により調整された区切り記号推測情報を得ることが可能となる。そして、区切り記号推測情報に基づいて対象テキストに区切り記号が挿入されることにより、発話者の意図に沿う位置に区切り記号が挿入されたテキストを容易に得ることが可能となる。

[0108] 第6の側面に係る区切り記号挿入装置では、第5の側面に係る区切り記号挿入装置において、前記区切り記号推測情報は、前記区切り記号除去文に含まれる各単語の後に複数の種類の区切り記号のそれぞれを挿入することに関する尤度である記号挿入尤度、及び、各単語の後に区切り記号を未挿入とすることに関する尤度である記号無し尤度を含むこととしてもよい。

[0109] 上記の側面によれば、各単語に関して、語間時間により調整された記号挿入尤度及び記号なし尤度を含む区切り記号推測情報を得ることが可能となる。調整された記号挿入尤度及び記号なし尤度に基づいて対象テキストに区切り記号が挿入されることにより、発話者の意図に沿う位置に区切り記号が挿入されたテキストを容易に得ることが可能となる。

[0110] 第7の側面に係る区切り記号挿入装置では、第5または6の側面に係る区切り記号挿入装置において、前記区切り記号挿入モデルは、前記区切り記号除去文に対応する発話音声が発話されたときの状況を示すシチュエーション情報を入力として更に含み、前記語間時間が各単語に関連付けられると共に前記シチュエーション情報が関連付けられた前記区切り記号除去文と前記区切り記号入り文とのペアからなる学習データを用いた機械学習により生成され、前記区切り記号挿入部は、前記シチュエーション情報が関連付けられた

前記対象テキストを前記区切り記号挿入モデルに入力することとしてもよい。

[0111] 上記の側面によれば、シチュエーション情報が関連付けられた区切り記号除去文が学習データとして用いられた機会学習により区切り記号挿入モデルが生成され、シチュエーション情報を含む区切り記号除去文が区切り記号挿入モデルの入力とされるので、シチュエーション情報が関連付けられた対象テキストを区切り記号挿入モデルに入力することにより、発話音声が発話されたときの状況に応じた発話の傾向が考慮された区切り記号推測情報を得ることが可能となる。

[0112] 第1の側面に係る音声認識システムでは、第1～7の側面のいずれか一つの側面に係る区切り記号挿入装置を含み、音声認識エンジンにより音声認識された結果のテキスト及び該テキストにおける各単語の前記語間時間を第1の音声認識結果として取得する音声認識結果取得部と、音声認識結果出力部と、を備え、前記区切り記号挿入装置の前記語間時間取得部は、前記第1の音声認識結果に含まれる前記語間時間を取得し、前記区切り記号挿入装置の前記区切り記号挿入部は、前記第1の音声認識結果に含まれるテキストを前記対象テキストとして、区切り記号を挿入し、前記音声認識結果出力部は、前記区切り記号挿入部により区切り記号が挿入された前記対象テキストを、第2の音声認識結果として出力することとしてもよい。

[0113] 上記の側面によれば、音声認識処理において語間時間が考慮されない音声認識エンジンから得られた第1の音声認識結果に基づいて、第1の音声認識結果に含まれるテキストを対象テキストとして、第1の音声認識結果に含まれる語間時間に応じて調整された区切り記号推測情報を取得できるので、発話者の意図に沿う位置に区切り記号が挿入された対象テキストを含む第2の音声認識結果を得ることが可能となる。

[0114] 以上、本実施形態について詳細に説明したが、当業者にとっては、本実施形態が本明細書中に説明した実施形態に限定されるものではないということは明らかである。本実施形態は、特許請求の範囲の記載により定まる本発明

の趣旨及び範囲を逸脱することなく修正及び変更態様として実施することができる。したがって、本明細書の記載は、例示説明を目的とするものであり、本実施形態に対して何ら制限的な意味を有するものではない。

[0115] 情報の通知は、本開示において説明した態様／実施形態に限られず、他の方法を用いて行われてもよい。例えば、情報の通知は、物理レイヤシグナリング（例えば、D C I (Downlink Control Information)、U C I (Uplink Control Information)）、上位レイヤシグナリング（例えば、R R C (Radio Resource Control) シグナリング、M A C (Medium Access Control) シグナリング、報知情報（M I B (Master Information Block)、S I B (System Information Block)））、その他の信号又はこれらの組み合わせによって実施されてもよい。また、R R Cシグナリングは、R R Cメッセージと呼ばれてもよく、例えば、R R C接続セットアップ（R R C Connection Setup）メッセージ、R R C接続再構成（R R C Connection Reconfiguration）メッセージなどであってもよい。

[0116] 本明細書で説明した各態様／実施形態は、L T E (Long Term Evolution)、L T E - A (LTE-Advanced)、S U P E R 3 G、I M T - A d v a n c e d、4 G、5 G、F R A (Future Radio Access)、W - C D M A（登録商標）、G S M（登録商標）、C D M A 2 0 0 0、U M B (Ultra Mobile Broadband)、I E E E 8 0 2 . 1 1 (W i - F i)、I E E E 8 0 2 . 1 6 (W i M A X)、I E E E 8 0 2 . 2 0、U W B (Ultra-WideBand)、B l u e t o o t h（登録商標）、その他の適切なシステムを利用するシステム及び／又はこれらに基づいて拡張された次世代システムに適用されてもよい。また、複数のシステムが組み合わされて（例えば、L T E及びL T E - Aの少なくとも一方と5 Gとの組み合わせ等）適用されてもよい。

[0117] 本明細書で説明した各態様／実施形態の処理手順、シーケンス、フローチャートなどは、矛盾の無い限り、順序を入れ替えてもよい。例えば、本明細書で説明した方法については、例示的な順序で様々なステップの要素を提示しており、提示した特定の順序に限定されない。

- [0118] 本開示において基地局によって行われるとした特定動作は、場合によってはその上位ノード (upper node) によって行われることもある。基地局を有する1つ又は複数のネットワークノード (network nodes) からなるネットワークにおいて、端末との通信のために行われる様々な動作は、基地局及び基地局以外の他のネットワークノード (例えば、MME又はS-GWなどが考えられるが、これらに限られない) の少なくとも1つによって行われ得ることは明らかである。上記において基地局以外の他のネットワークノードが1つである場合を例示したが、複数の他のネットワークノードの組み合わせ (例えば、MME及びS-GW) であってもよい。
- [0119] 情報等 (※「情報、信号」の項目参照) は、上位レイヤ (又は下位レイヤ) から下位レイヤ (又は上位レイヤ) へ出力され得る。複数のネットワークノードを介して入出力されてもよい。
- [0120] 入出力された情報等は特定の場所 (例えば、メモリ) に保存されてもよいし、管理テーブルで管理してもよい。入出力される情報等は、上書き、更新、または追記され得る。出力された情報等は削除されてもよい。入力された情報等は他の装置へ送信されてもよい。
- [0121] 判定は、1ビットで表される値 (0か1か) によって行われてもよいし、真偽値 (Boolean: trueまたはfalse) によって行われてもよいし、数値の比較 (例えば、所定の値との比較) によって行われてもよい。
- [0122] 本開示において説明した各態様／実施形態は単独で用いてもよいし、組み合わせで用いてもよいし、実行に伴って切り替えて用いてもよい。また、所定の情報の通知 (例えば、「Xであること」の通知) は、明示的に行うものに限られず、暗黙的 (例えば、当該所定の情報の通知を行わない) ことによって行われてもよい。
- [0123] 以上、本開示について詳細に説明したが、当業者にとっては、本開示が本開示中に説明した実施形態に限定されるものではないということは明らかである。本開示は、請求の範囲の記載により定まる本開示の趣旨及び範囲を逸脱することなく修正及び変更態様として実施することができる。したがって

、本開示の記載は、例示説明を目的とするものであり、本開示に対して何ら制限的な意味を有するものではない。

[0124] ソフトウェアは、ソフトウェア、ファームウェア、ミドルウェア、マイクロコード、ハードウェア記述言語と呼ばれるか、他の名称で呼ばれるかを問わず、命令、命令セット、コード、コードセグメント、プログラムコード、プログラム、サブプログラム、ソフトウェアモジュール、アプリケーション、ソフトウェアアプリケーション、ソフトウェアパッケージ、ルーチン、サブルーチン、オブジェクト、実行可能ファイル、実行スレッド、手順、機能などを意味するよう広く解釈されるべきである。

[0125] また、ソフトウェア、命令などは、伝送媒体を介して送受信されてもよい。例えば、ソフトウェアが、同軸ケーブル、光ファイバケーブル、ツイストペア及びデジタル加入者回線（DSL）などの有線技術及び／又は赤外線、無線及びマイクロ波などの無線技術を使用してウェブサイト、サーバ、又は他のリモートソースから送信される場合、これらの有線技術及び／又は無線技術は、伝送媒体の定義内に含まれる。

[0126] 本開示において説明した情報、信号などは、様々な異なる技術のいずれかを使用して表されてもよい。例えば、上記の説明全体に渡って言及され得るデータ、命令、コマンド、情報、信号、ビット、シンボル、チップなどは、電圧、電流、電磁波、磁界若しくは磁性粒子、光場若しくは光子、又はこれらの任意の組み合わせによって表されてもよい。

[0127] なお、本開示において説明した用語及び／又は本明細書の理解に必要な用語については、同一の又は類似する意味を有する用語と置き換えてもよい。

[0128] 本明細書で使用する「システム」および「ネットワーク」という用語は、互換的に使用される。

[0129] また、本明細書で説明した情報、パラメータなどは、絶対値で表されてもよいし、所定の値からの相対値で表されてもよいし、対応する別の情報で表されてもよい。例えば、無線リソースはインデックスによって指示されるものであってもよい。

[0130] 上述したパラメータに使用する名称はいかなる点においても限定的な名称ではない。さらに、これらのパラメータを使用する数式等は、本開示で明示的に開示したものと異なる場合もある。様々なチャネル（例えば、P U C C H、P D C C Hなど）及び情報要素は、あらゆる好適な名称によって識別できるので、これらの様々なチャネル及び情報要素に割り当てている様々な名称は、いかなる点においても限定的な名称ではない。

[0131] 本開示で使用する「判断(determining)」、「決定(determining)」という用語は、多種多様な動作を包含する場合がある。「判断」、「決定」は、例えば、判定(judging)、計算(calculating)、算出(computing)、処理(processing)、導出(deriving)、調査(investigating)、探索(looking up、search、inquiry)（例えば、テーブル、データベース又は別のデータ構造での探索）、確認(ascertaining)した事を「判断」「決定」したとみなす事などを含み得る。また、「判断」、「決定」は、受信(receiving)（例えば、情報を受信すること）、送信(transmitting)（例えば、情報を送信すること）、入力(input)、出力(output)、アクセス(accessing)（例えば、メモリ中のデータにアクセスすること）した事を「判断」「決定」したとみなす事などを含み得る。また、「判断」、「決定」は、解決(resolving)、選択(selecting)、選定(choosing)、確立(establishing)、比較(comparing)などした事を「判断」「決定」したとみなす事を含み得る。つまり、「判断」「決定」は、何らかの動作を「判断」「決定」したとみなす事を含み得る。また、「判断（決定）」は、「想定する(assuming)」、「期待する(expecting)」、「みなす(considering)」などで読み替えられてもよい。

[0132] 本開示で使用する「に基づいて」という記載は、別段に明記されていない限り、「のみに基づいて」を意味しない。言い換えれば、「に基づいて」という記載は、「のみに基づいて」と「に少なくとも基づいて」の両方を意味する。

[0133] 本明細書で「第1の」、「第2の」などの呼称を使用した場合には、その要素へのいかなる参照も、それらの要素の量または順序を全般的に限

定するものではない。これらの呼称は、2つ以上の要素間を区別する便利な方法として本明細書で使用され得る。したがって、第1および第2の要素への参照は、2つの要素のみがそこで採用され得ること、または何らかの形で第1の要素が第2の要素に先行しなければならないことを意味しない。

[0134] 「含む(include)」、「含んでいる(including)」、およびそれらの変形が、本明細書あるいは特許請求の範囲で使用されている限り、これら用語は、用語「備える(comprising)」と同様に、包括的であることが意図される。さらに、本明細書あるいは特許請求の範囲において使用されている用語「または(or)」は、排他的論理和ではないことが意図される。

[0135] 本開示において、例えば、英語でのa, an及びtheのように、翻訳により冠詞が追加された場合、本開示は、これらの冠詞の後に続く名詞が複数形であることを含んでもよい。

[0136] 本開示において、「AとBが異なる」という用語は、「AとBが互いに異なる」ことを意味してもよい。なお、当該用語は、「AとBがそれぞれCと異なる」ことを意味してもよい。「離れる」、「結合される」などの用語も、「異なる」と同様に解釈されてもよい。

符号の説明

[0137] 1 0…区切り記号挿入装置、1 1…対象テキスト取得部、1 2…語間時間取得部、1 3…区切り記号挿入部、1 4…出力部、2 0…音声認識システム、2 1…音声認識結果取得部、2 2…音声認識結果出力部、M 1…記録媒体、m 1 0…メインモジュール、m 1 1…対象テキスト取得モジュール、m 1 2…語間時間取得モジュール、m 1 3…区切り記号挿入モジュール、m 1 4…出力モジュール、m d 1, m d 2…記号挿入モデル。

請求の範囲

- [請求項1] 発話音声の音声認識により得られたテキストに含まれる単語の後に、文を区切る区切り記号を挿入する区切り記号挿入装置であって、
- 前記発話音声に含まれる各単語における次の単語が発話されるまでの時間の長さである語間時間を取得する語間時間取得部と、
- 区切り記号挿入モデル及び前記語間時間に基づいて、前記発話音声の音声認識により得られたテキストである対象テキストに区切り記号を挿入する区切り記号挿入部であって、
- 前記区切り記号挿入モデルは、
- 前記区切り記号を含まない文である区切り記号除去文を少なくとも入力とし、
- 前記区切り記号除去文に含まれる各単語の後に挿入される区切り記号を示す区切り記号推測情報を出力し、
- 前記区切り記号除去文と区切り記号を含む文である区切り記号入り文とのペアを含む学習データを用いた機械学習により生成されるモデルであり、
- 前記対象テキストを前記区切り記号除去文として前記区切り記号挿入モデルに入力することにより得られ、前記語間時間に応じて調整された前記区切り記号推測情報に基づいて、前記対象テキストに区切り記号を挿入する、区切り記号挿入部と、
- を備える区切り記号挿入装置。
- [請求項2] 前記区切り記号挿入部は、前記区切り記号挿入モデルから出力された前記区切り記号推測情報を、前記語間時間に基づいて調整する、
- 請求項1に記載の区切り記号挿入装置。
- [請求項3] 前記区切り記号推測情報は、前記区切り記号除去文に含まれる各単語の後に複数の種類の区切り記号のそれぞれを挿入することに関する尤度である記号挿入尤度、及び、各単語の後に区切り記号を未挿入とすることに関する尤度である記号なし尤度を含み、

前記区切り記号挿入部は、

前記対象テキストに含まれる単語のうちの一の単語における前記語間時間が第1の時間である場合に、前記一の単語の前記記号なし尤度を上げるように調整し、又は／及び、複数の種類の区切り記号のうちの少なくとも一種の前記一の単語の区切り記号の前記記号挿入尤度を下げるように調整し、

前記一の単語における前記語間時間が前記第1の時間より長い第2の時間である場合に、複数の種類の区切り記号のうちの少なくとも一種の前記一の単語の区切り記号の前記記号挿入尤度を上げるように調整し、又は／及び、前記一の単語の前記記号なし尤度を下げるように調整し、

前記記号挿入尤度及び前記記号なし尤度のうちの最大の尤度に基づいて、前記一の単語の後に、複数の種類のうちのいずれかの区切り記号を挿入し、又は区切り記号を未挿入とする、

請求項2に記載の区切り記号挿入装置。

[請求項4]

前記区切り記号挿入部は、

前記対象テキストに含まれる全単語のうちの一以上の単語の前記語間時間の長さの程度が、所与の程度より長い場合に、前記全単語のそれぞれの前記語間時間を所与の程度で短くなるように補正した補正語間時間に基づいて、前記記号挿入尤度及び／又は前記記号なし尤度を調整する、

請求項3に記載の区切り記号挿入装置。

[請求項5]

前記区切り記号挿入モデルは、

前記区切り記号除去文に含まれる各単語の語間時間を入力として更に含み、

前記語間時間が各単語に関連付けられた前記区切り記号除去文と前記区切り記号入り文とのペアからなる学習データを用いた機械学習により生成され、

前記語間時間により調整された前記区切り記号推測情報を出力し

、

前記区切り記号挿入部は、各単語に前記語間時間が関連付けられた前記対象テキストを前記区切り記号挿入モデルに入力する、

請求項 1 に記載の区切り記号挿入装置。

[請求項6]

前記区切り記号推測情報は、前記区切り記号除去文に含まれる各単語の後に複数の種類の区切り記号のそれぞれを挿入することに関する尤度である記号挿入尤度、及び、各単語の後に区切り記号を未挿入とすることに関する尤度である記号無し尤度を含む、

請求項 5 に記載された区切り記号挿入装置。

[請求項7]

前記区切り記号挿入モデルは、

前記区切り記号除去文に対応する発話音声が発話されたときの状況を示すシチュエーション情報を入力として更に含み、

前記語間時間が各単語に関連付けられると共に前記シチュエーション情報が関連付けられた前記区切り記号除去文と前記区切り記号入り文とのペアからなる学習データを用いた機械学習により生成され、

前記区切り記号挿入部は、前記シチュエーション情報が関連付けられた前記対象テキストを前記区切り記号挿入モデルに入力する、

請求項 5 または 6 に記載の区切り記号挿入装置。

[請求項8]

請求項 1 に記載の区切り記号挿入装置を含む音声認識システムであって、

音声認識エンジンにより音声認識された結果のテキスト及び該テキストにおける各単語の前記語間時間を第 1 の音声認識結果として取得する音声認識結果取得部と、

音声認識結果出力部と、を備え、

前記区切り記号挿入装置の前記語間時間取得部は、前記第 1 の音声認識結果に含まれる前記語間時間を取得し、

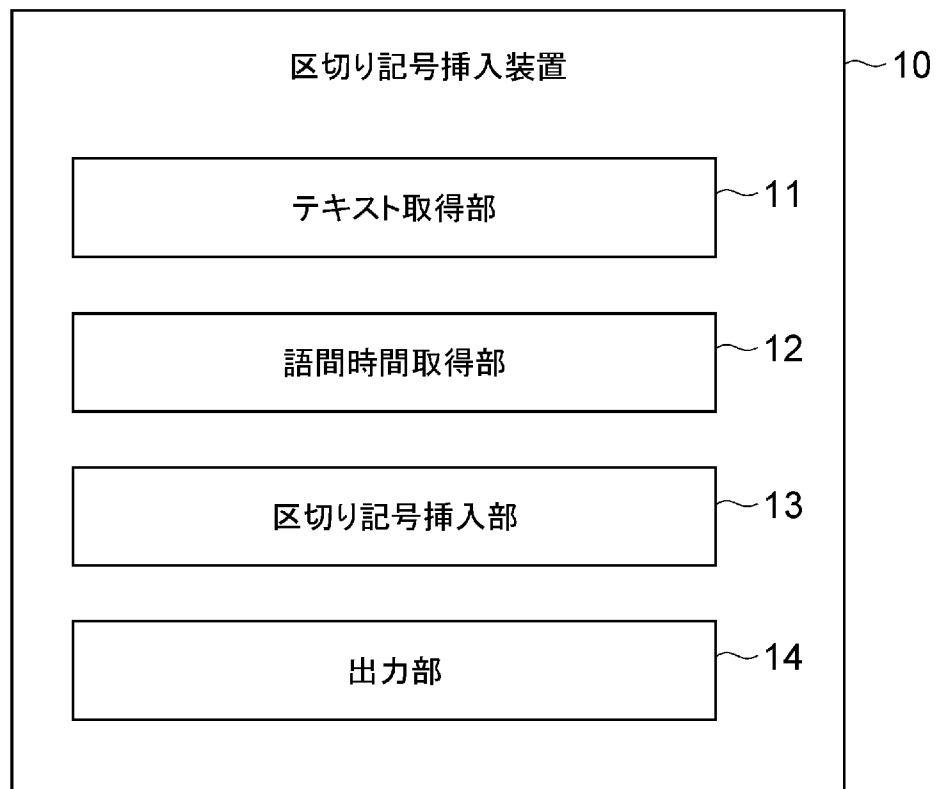
前記区切り記号挿入装置の前記区切り記号挿入部は、前記第 1 の音

声認識結果に含まれるテキストを前記対象テキストとして、区切り記号を挿入し、

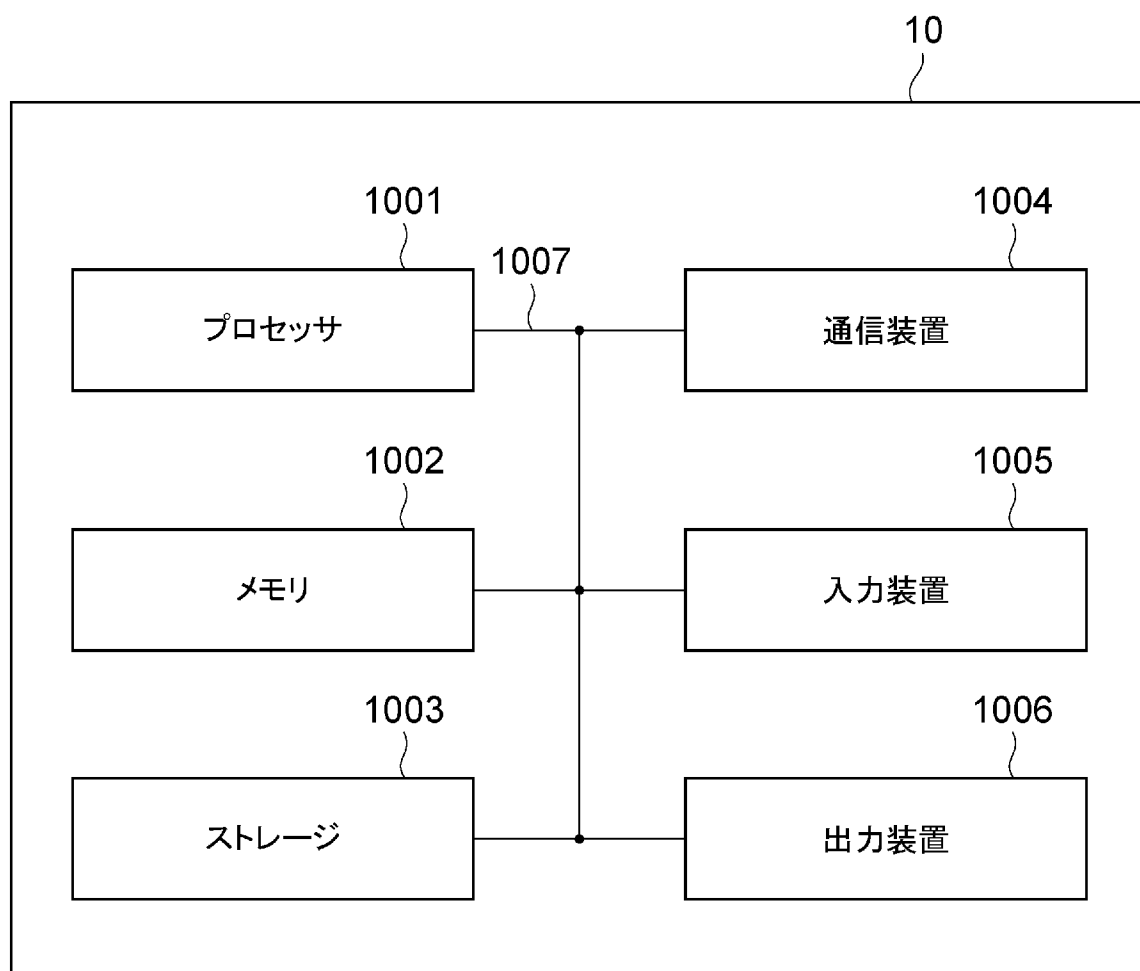
前記音声認識結果出力部は、前記区切り記号挿入部により区切り記号が挿入された前記対象テキストを、第2の音声認識結果として出力する、

を備える音声認識システム。

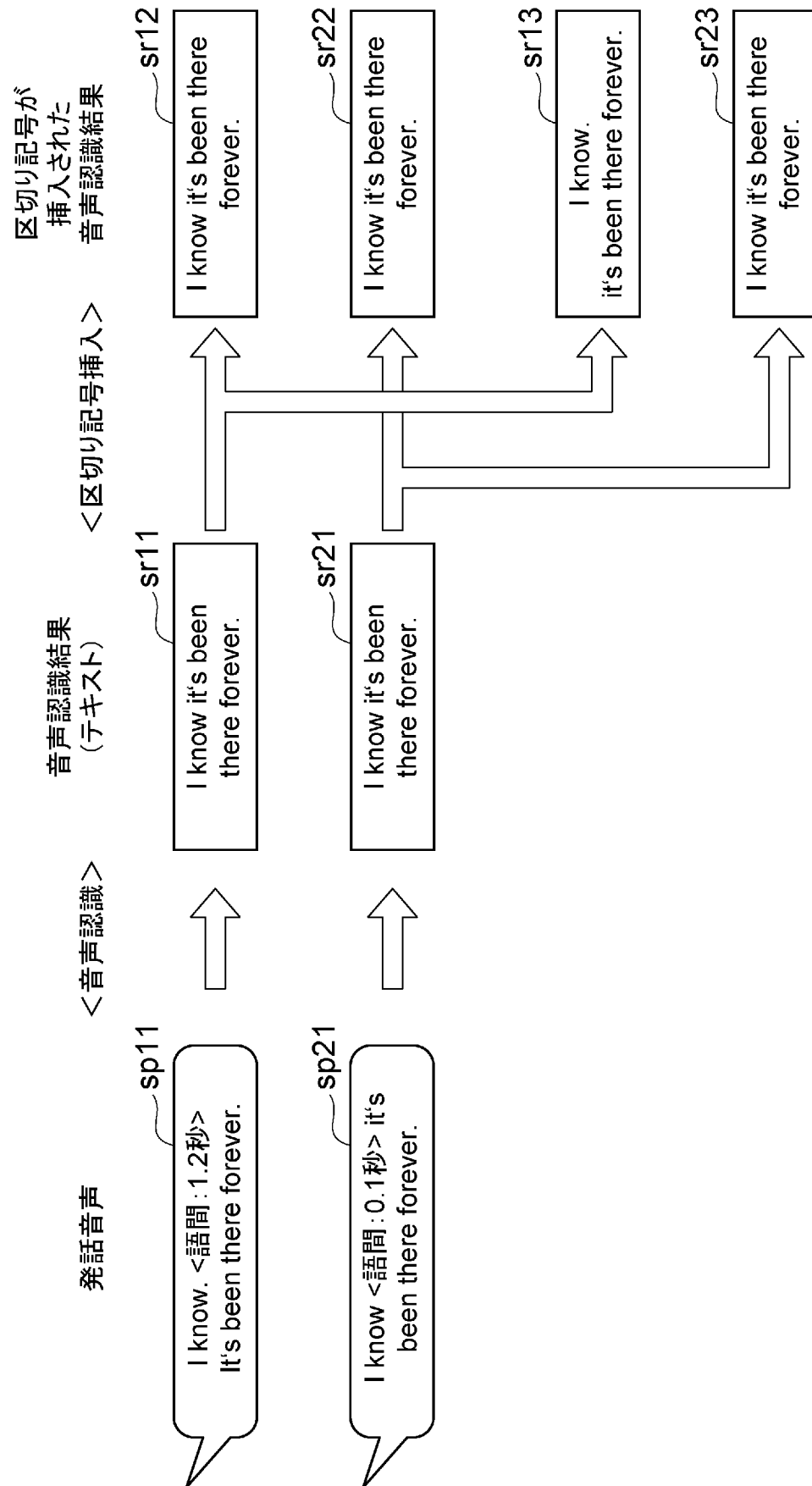
[図1]



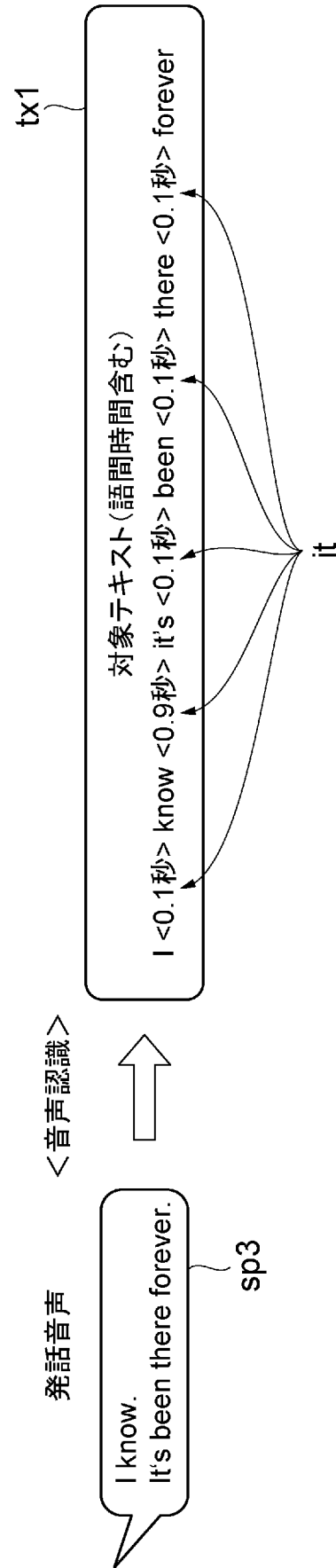
[図2]



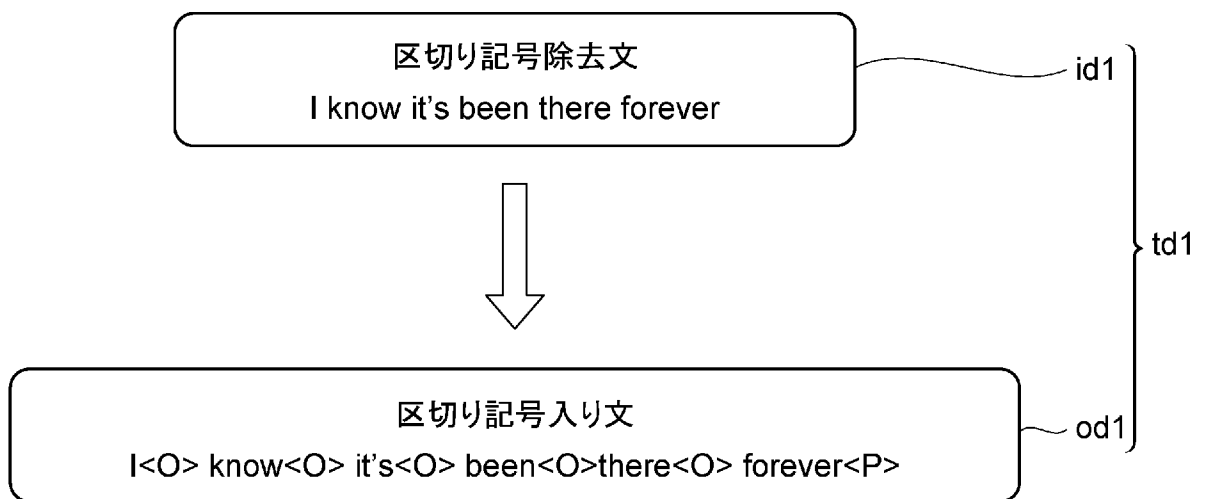
[図3]



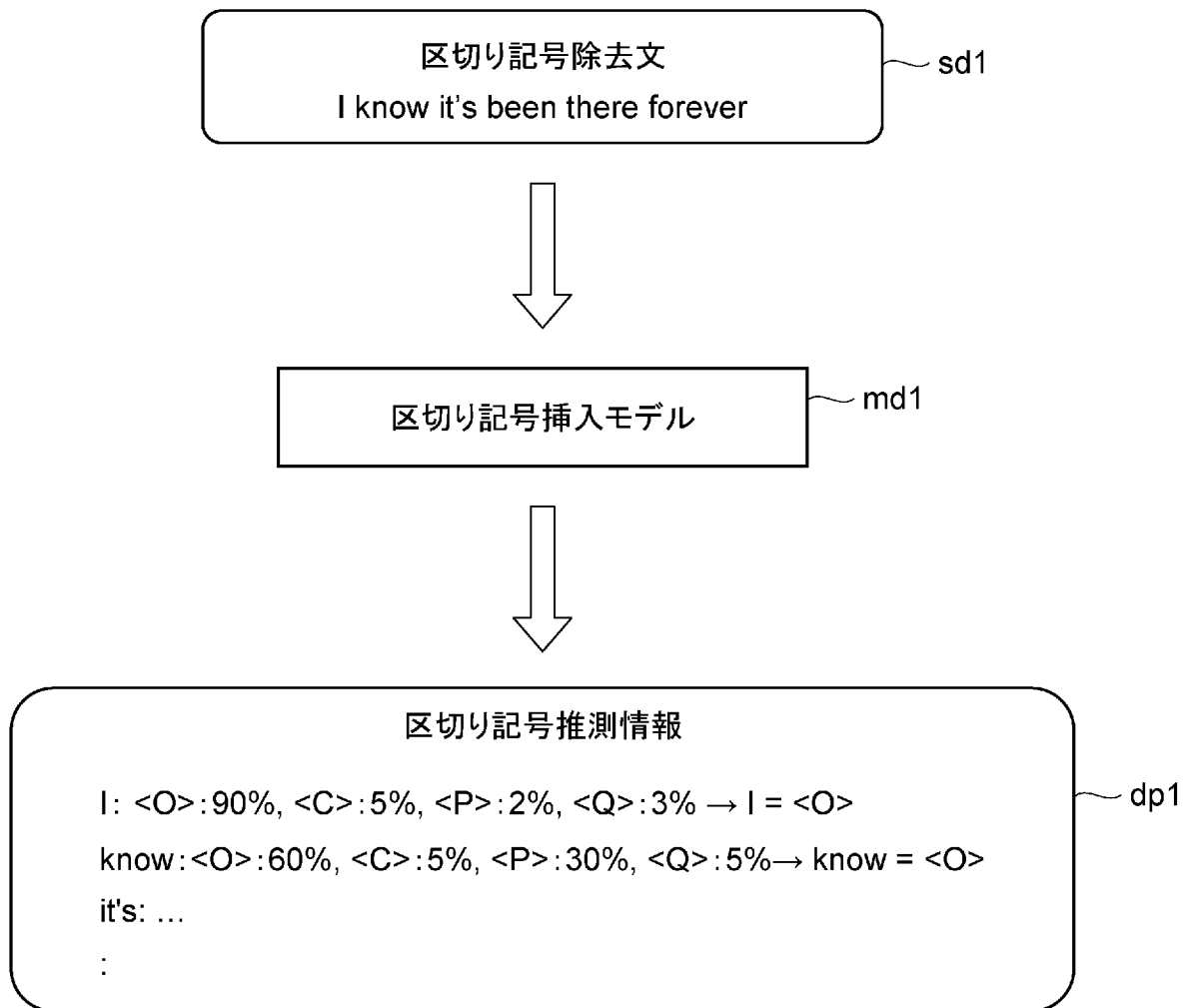
[図4]



[図5]



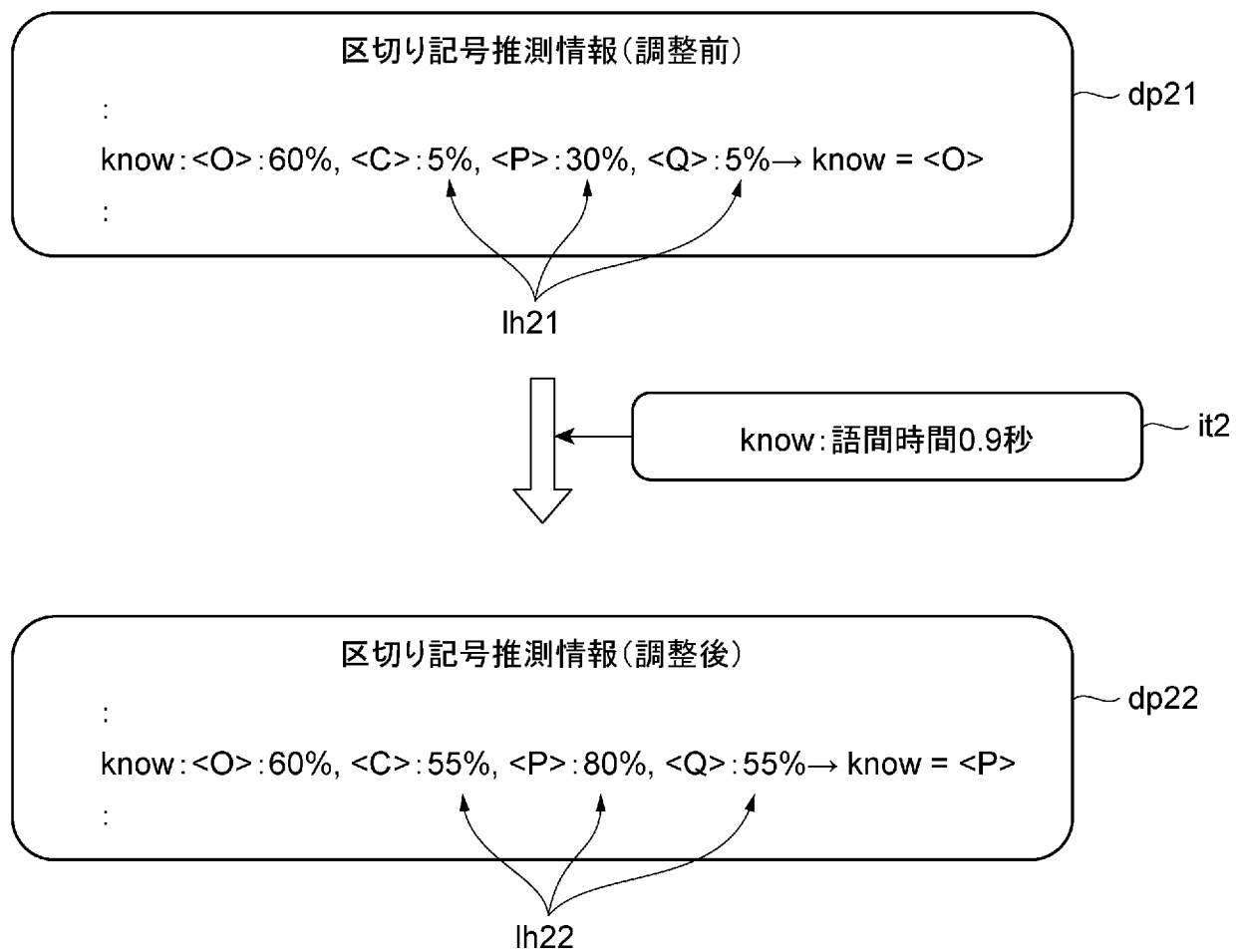
[図6]



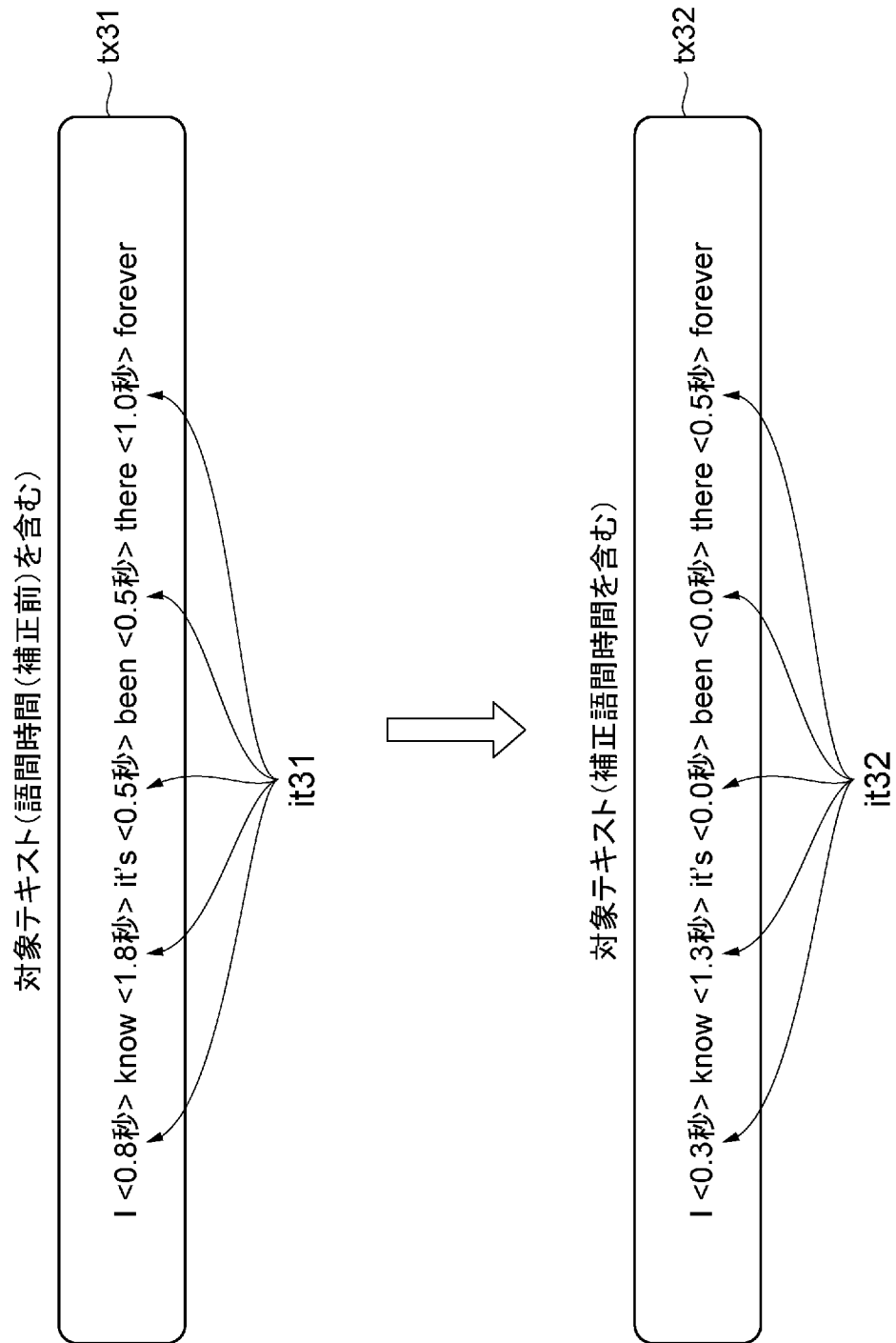
[図7]

語間時間(x秒)	語間時間カテゴリ	尤度調整情報
$x \leq 0.1$	無	区切り記号なしの尤度(記号なし尤度)を50%上げる
$0.1 < x \leq 0.5$	小	ピリオド、カンマ、クエスチョンマークの尤度(記号挿入尤度)を20%上げる
$0.5 < x \leq 1.0$	中	ピリオド、カンマ、クエスチョンマークの尤度(記号挿入尤度)を50%上げる
$1.0 < x$	大	ピリオド、カンマ、クエスチョンマークの尤度(記号挿入尤度)を80%上げる

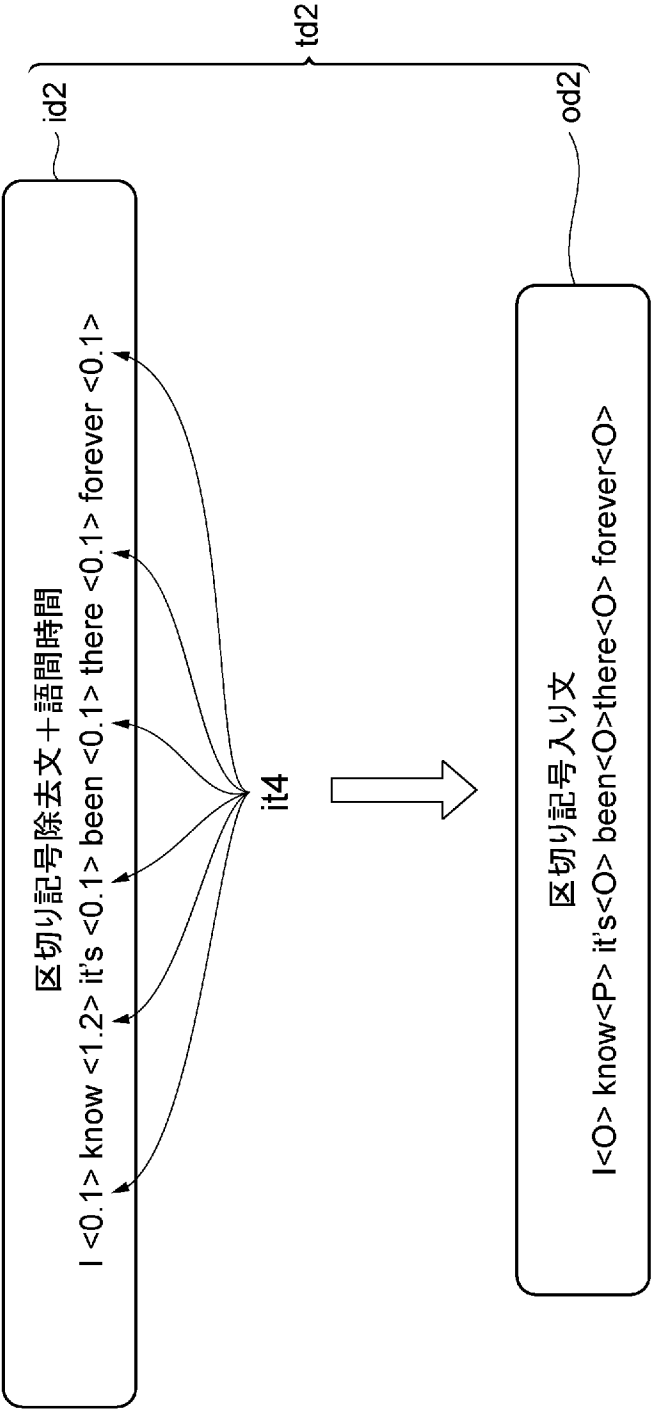
[図8]



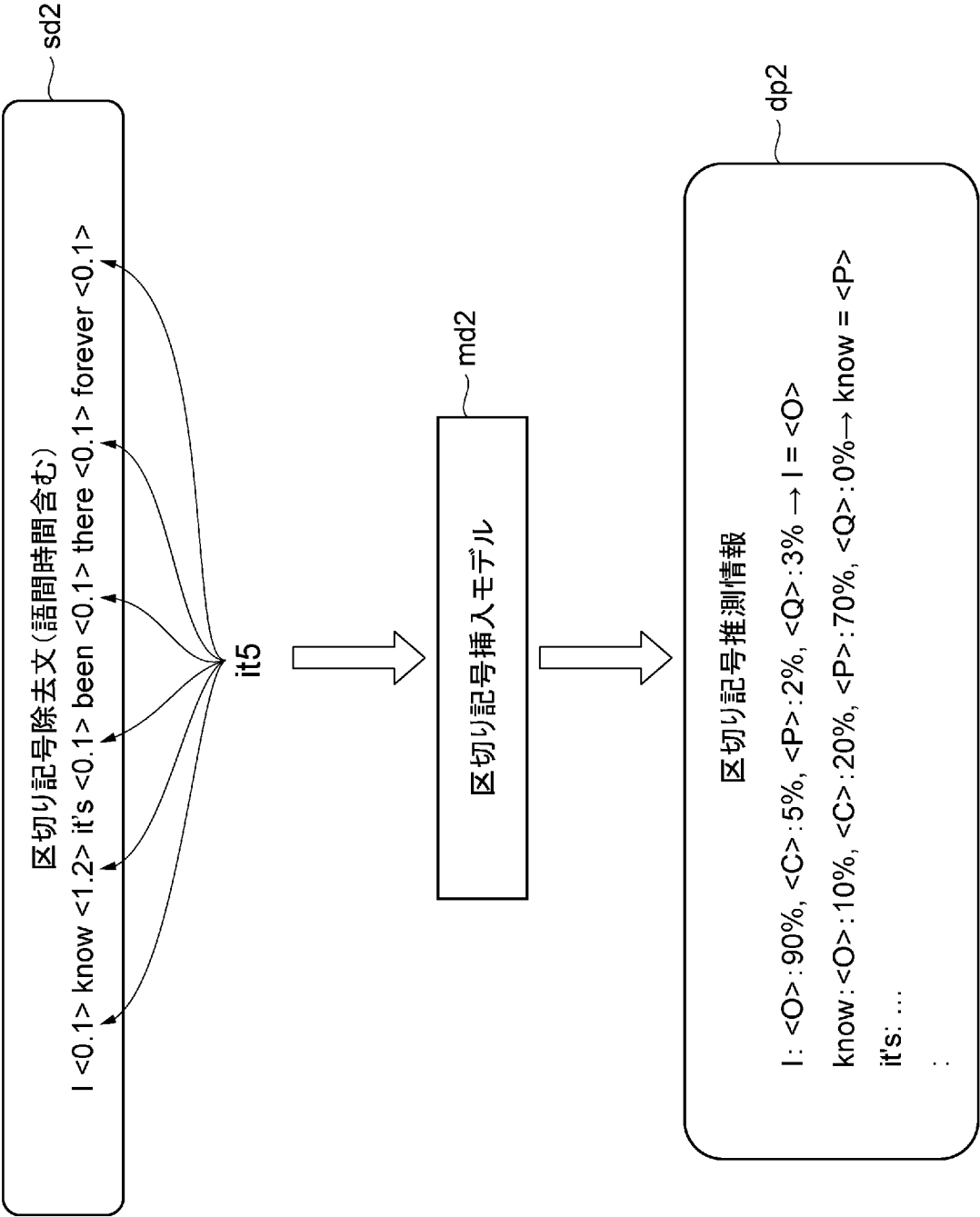
[図9]



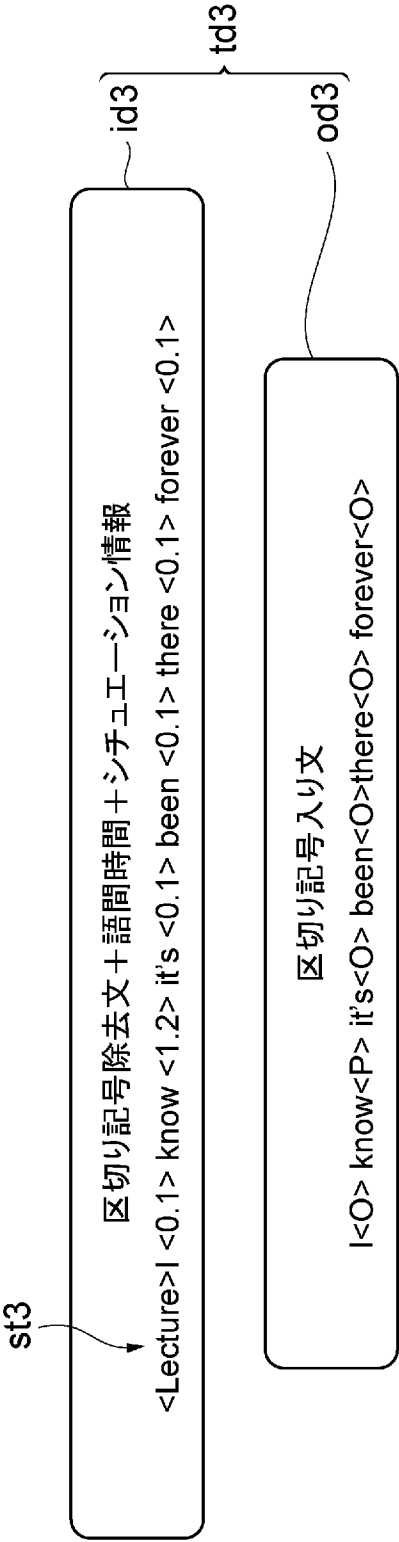
[図10]



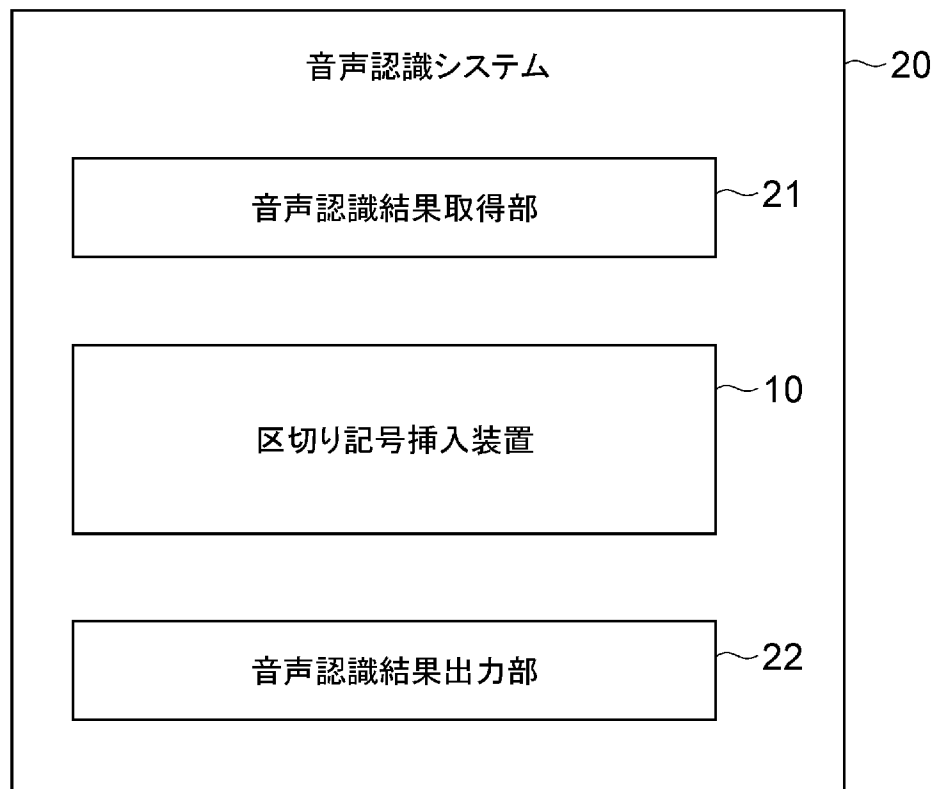
[図11]



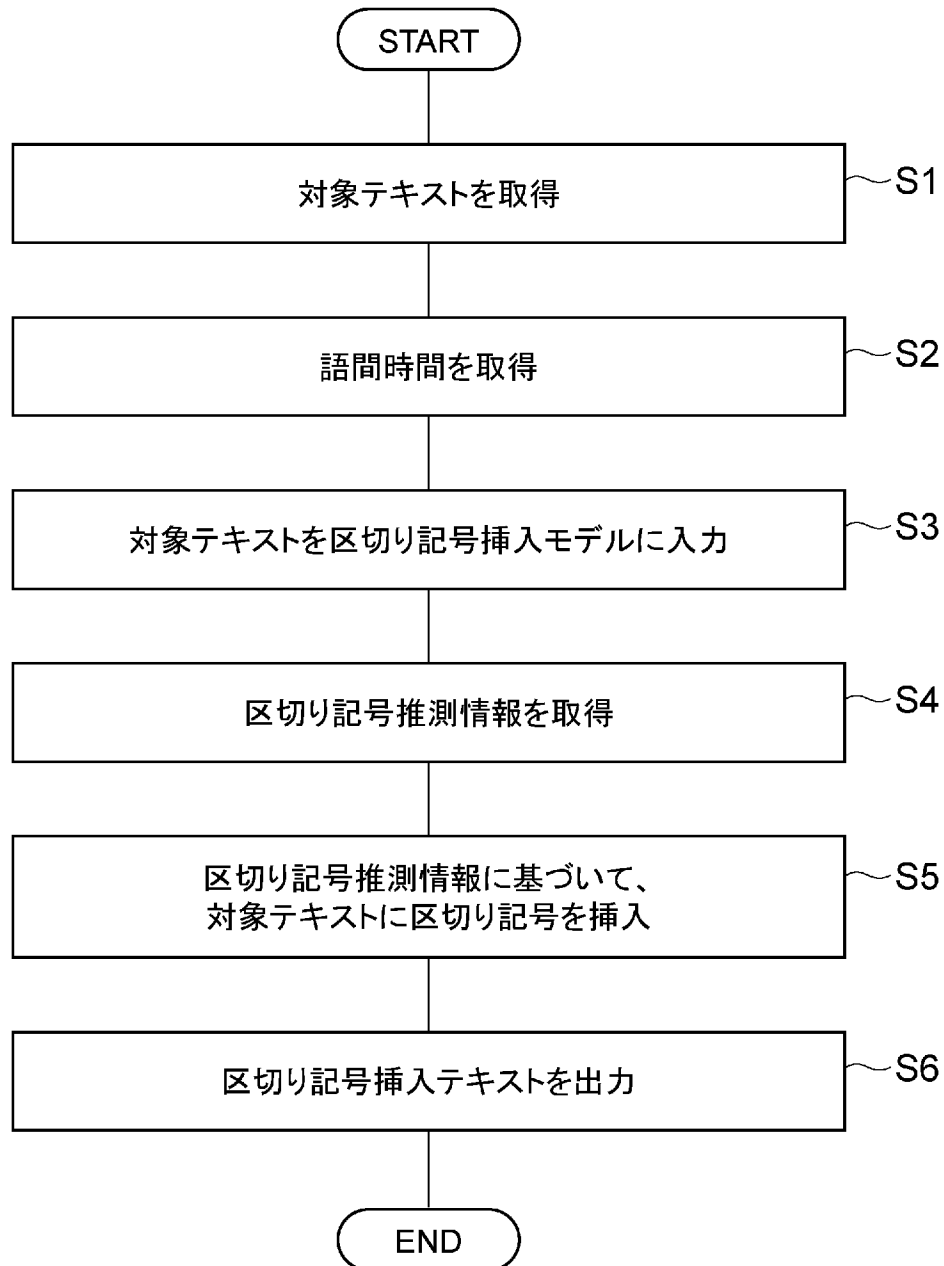
[図12]



[図13]



[図14]



[図15]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2023/017568

A. CLASSIFICATION OF SUBJECT MATTER**G10L 15/04**(2013.01)i; **G06F 40/166**(2020.01)i

FI: G10L15/04 200; G06F40/166

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L15/00-15/34; G06F40/00-40/58

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996
 Published unexamined utility model applications of Japan 1971-2023
 Registered utility model specifications of Japan 1996-2023
 Published registered utility model applications of Japan 1994-2023

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

IEEE Xplore

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	TILK, Ottokar et al. LSTM for Punctuation Restoration in Speech Transcripts. Interspeech 2015. September 2015, pages 683-687 chapter 2-3, fig. 1	1-2, 8
A		3-7
X	JP 2015-219480 A (NIPPON TELEGR. & TELEPH. CORP.) 07 December 2015 (2015-12-07) paragraphs [0026]-[0030], [0054]-[0083], fig. 1, 11	1, 5-6, 8
A		2-4, 7
A	WO 2009/101837 A1 (NEC CORP.) 20 August 2009 (2009-08-20) entire text, all drawings	1-8
A	CN 112927679 A (CHONGDONG INNOVATION PLATFORM (SHENZHEN) CO., LTD.) 08 June 2021 (2021-06-08) entire text, all drawings	1-8

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance
 “E” earlier application or patent but published on or after the international filing date
 “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
 “O” document referring to an oral disclosure, use, exhibition or other means
 “P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
 “X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
 “Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
 “&” document member of the same patent family

Date of the actual completion of the international search

13 July 2023

Date of mailing of the international search report

25 July 2023

Name and mailing address of the ISA/JP

Japan Patent Office (ISA/JP)
 3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915
 Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2023/017568

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP	2015-219480	A	07 December 2015	(Family: none)	
WO	2009/101837	A1	20 August 2009	US 2010/0292989	A1
				entire text, all drawings	
CN	112927679	A	08 June 2021	WO 2022/166218	A1
				entire text, all drawings	

A. 発明の属する分野の分類（国際特許分類（IPC）） G10L 15/04(2013.01)i; G06F 40/166(2020.01)i FI: G10L15/04 200; G06F40/166		
B. 調査を行った分野		
調査を行った最小限資料（国際特許分類（IPC）） G10L15/00-15/34; G06F40/00-40/58		
最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1922-1996年 日本国公開実用新案公報 1971-2023年 日本国実用新案登録公報 1996-2023年 日本国登録実用新案公報 1994-2023年		
国際調査で使用した電子データベース（データベースの名称、調査に使用した用語） IEEE Xplore		
C. 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X	TILK, Ottokar et al., LSTM for punctuation restoration in speech transcripts, Interspeech 2015, 2015.09, p.683-687 2-3章, 図1	1-2, 8
A		3-7
X	JP 2015-219480 A (日本電信電話株式会社) 07.12.2015 (2015 - 12 - 07) [0026]-[0030], [0054]-[0083], 図1, 11	1, 5-6, 8
A		2-4, 7
A	WO 2009/101837 A1 (日本電気株式会社) 20.08.2009 (2009 - 08 - 20) 全文, 全図	1-8
A	CN 112927679 A (CHONGDONG INNOVATION PLATFORM (SHENZHEN) CO., LTD.) 08.06.2021 (2021 - 06 - 08) 全文, 全図	1-8
☐ C欄の続きにも文献が列挙されている。 ☒ パテントファミリーに関する別紙を参照。		
* 引用文献のカテゴリー “A” 特に関連のある文献ではなく、一般的な技術水準を示すもの “E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの “L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す） “O” 口頭による開示、使用、展示等に言及する文献 “P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献 “T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの “X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの “Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの “&” 同一パテントファミリー文献		
国際調査を完了した日 13.07.2023		国際調査報告の発送日 25.07.2023
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号		権限のある職員（特許庁審査官） 中村 天真 5Z 1786 電話番号 03-3581-1101 内線 3591

国際調査報告
パテントファミリーに関する情報

国際出願番号

PCT/JP2023/017568

引用文献			公表日	パテントファミリー文献	公表日
JP	2015-219480	A	07.12.2015	(ファミリーなし)	
WO	2009/101837	A1	20.08.2009	US 2010/0292989 A1	
				全文, 全図	
CN	112927679	A	08.06.2021	WO 2022/166218 A1	
				全文, 全図	