

(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
2017 年 10 月 12 日(12.10.2017)



(10) 国際公開番号

WO 2017/175351 A1

- (51) 国際特許分類:
G10L 15/22 (2006.01) G10L 15/10 (2006.01)
G10L 13/00 (2006.01) G10L 15/28 (2013.01)
- (21) 国際出願番号: PCT/JP2016/061385
- (22) 国際出願日: 2016 年 4 月 7 日(07.04.2016)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人: 株式会社ソニー・インタラクティブエンタテインメント (SONY INTERACTIVE ENTERTAINMENT INC.) [JP/JP]; 〒1080075 東京都港区港南 1 丁目 7 番 1 号 Tokyo (JP).
- (72) 発明者: 本多 真一 (HONDA, Shinichi); 〒1080075 東京都港区港南 1 丁目 7 番 1 号 株式会社ソニー・インタラクティブエンタテインメント内 Tokyo (JP). 菊池 めぐみ (KIKUCHI, Megumi); 〒1080075 東京都港区港南 1 丁目 7 番 1 号 株式会社ソニー・インタラクティブエンタテインメント内 Tokyo (JP). 佐竹 隆史 (SATAKE, Takashi); 〒1080075 東京都港区港南 1 丁目 7 番 1 号 株

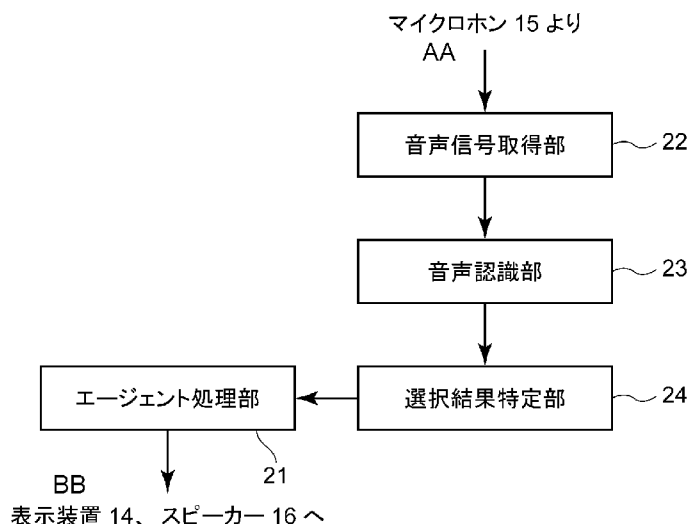
式会社ソニー・インタラクティブエンタテインメント内 Tokyo (JP).

- (74) 代理人: 竹居 信利 (TAKEI Nobutoshi); 〒1600022 東京都新宿区新宿 6 丁目 7-1 エルプリメント新宿 308 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, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーロシェン (AM, AZ, BY, KG, KZ, RU, TJ, TM), ユーロパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC,

[続葉有]

(54) Title: INFORMATION PROCESSING DEVICE

(54) 発明の名称: 情報処理装置



表示装置 14、スピーカー 16 へ

- 21 Agent processing unit
22 Voice signal acquisition unit
23 Voice recognition unit
24 Selection result specification unit
AA From microphone 15
BB To display device 14, speaker 16

(57) Abstract: Provided is an information processing device with which it is possible to enable smooth voice communication with a user. An information processing device for presenting a plurality of options to a user, recognizing the content of an utterance by the user selecting one of the plurality of options, and specifying the option selected by the user on the basis of whether or not a phrase included in the recognized content of the utterance by the user corresponds to an utterance included in a dictionary corresponding to each of the plurality of options readied in advance.

(57) 要約: ユーザーとの間で音声による円滑なコミュニケーションを実現することのできる情報処理装置を提供する。複数の選択肢をユーザーに提示し、当該複数の選択肢のいずれかを選択するユーザーの発話内容を認識し、認識されるユーザーの発話内容に含まれる語句が、予め用意された複数の選択肢のそれぞれに対応する辞書に含まれる語句と対応するか否かに基づいて、ユーザーが選択した選択肢を特定する情報処理装置である。



MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, 添付公開書類:
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, — 国際調査報告 (条約第 21 条(3))
KM, ML, MR, NE, SN, TD, TG).

明 細 書

発明の名称： 情報処理装置

技術分野

[0001] 本発明は、ユーザーから音声による入力を受け付ける情報処理装置、情報処理方法、プログラム、及び情報記憶媒体に関する。

背景技術

[0002] ユーザーが発話する音声を受け付けて、その音声の内容に応じて情報処理を実行する情報処理装置が知られている。このような技術によれば、ユーザーは手を使ってキーボードやタッチパネルで文字を入力する代わりに、声で各種の指示や情報を入力することができる。また、ユーザーの発話から、感情や態度など発話した内容以外のものを読み取る技術も研究されている（例えば非特許文献1 参照）。

先行技術文献

非特許文献

[0003] 非特許文献1：Kikuo Maekawa and Hiroki Mori, "VOICE-QUALITY ANALYSIS OF JAPANESE FILLED PAUSES: A PRELIMINARY REPORT", [online]、[平成27年11月16日検索]、インターネット<URL:<http://www2.ninjal.ac.jp/kikuo/Maekawa-DISS2015.pdf>>

発明の概要

発明が解決しようとする課題

[0004] 上記従来例の技術においては、ユーザーと情報処理装置との間で円滑なコミュニケーションを実現するうえで様々な課題がある。例えば、ユーザーの音声を受け付ける際には、精度よくユーザーの発話内容を認識する必要がある。しかしながら、ユーザーに完全に自由な発話を許すこととすると、ユーザーの発話内容を正確に認識することが難しくなる。そこで、ユーザーから受け付ける情報の種類が限られている場合には、ユーザーに選択肢を提示することによって、ユーザーが発話する内容をこの選択肢に限定することが考

えられる。ところが、ユーザーの発話内容を選択肢だけに制限すると、ユーザーの発話が機械的になり、ユーザーの感情や態度が発話内容に表れにくくなる。そのため、ユーザーに自然な発話をさせたい場合には、ユーザーの発話内容を選択肢だけに制限することは望ましくない。

[0005] また、ユーザーの発話に先立って情報処理装置がユーザーに対する発言を音声で再生する場合、その発言の再生が終了するまでは、ユーザーが発話を開始することができない。そのため、情報処理装置がユーザーにとって興味のない発言などを行うと、円滑なコミュニケーションが損なわれることもあり得る。

[0006] 本発明は上記実情を考慮してなされたものであって、その目的の一つは、ユーザーとの間で音声による円滑なコミュニケーションを実現することのできる情報処理装置、情報処理方法、プログラム、及び情報記憶媒体を提供することにある。

課題を解決するための手段

[0007] 本発明に係る情報処理装置は、複数の選択肢をユーザーに提示する選択肢提示部と、前記複数の選択肢のいずれかを選択する前記ユーザーの発話内容を認識する音声認識部と、前記認識されるユーザーの発話内容に含まれる語句が、予め用意された前記複数の選択肢のそれぞれに対応する辞書に含まれる語句と対応するか否かに基づいて、前記ユーザーが選択した選択肢を特定する選択結果特定部と、を含むことを特徴とする。

[0008] 本発明に係る別の情報処理装置は、ユーザーに提示すべき発言を音声で再生する再生部と、前記ユーザーの発話内容を認識する音声認識部と、を含み、前記再生部は、前記発言の再生中に前記音声認識部が前記ユーザーの発話を検出した場合に、前記発言の再生を中止することを特徴とする。

[0009] 本発明に係る情報処理方法は、複数の選択肢をユーザーに提示するステップと、前記複数の選択肢のいずれかを選択する前記ユーザーの発話内容を認識するステップと、前記認識されるユーザーの発話内容に含まれる語句が、予め用意された前記複数の選択肢のそれぞれに対応する辞書に含まれる語句

と対応するか否かに基づいて、前記ユーザーが選択した選択肢を特定するステップと、を含むことを特徴とする。

[0010] 本発明に係るプログラムは、複数の選択肢をユーザーに提示する選択肢提示部、前記複数の選択肢のいずれかを選択する前記ユーザーの発話内容を認識する音声認識部、及び、前記認識されるユーザーの発話内容に含まれる語句が、予め用意された前記複数の選択肢のそれぞれに対応する辞書に含まれる語句と対応するか否かに基づいて、前記ユーザーが選択した選択肢を特定する選択結果特定部、としてコンピュータを機能させるためのプログラムである。このプログラムは、コンピュータ読み取り可能で非一時的な情報記憶媒体に格納されて提供されてよい。

図面の簡単な説明

[0011] [図1]本発明の実施の形態に係る情報処理装置の構成を示す構成ブロック図である。

[図2]本発明の実施の形態に係る情報処理装置の機能を示す機能ブロック図である。

[図3]選択肢を表示する画面の一例を示す図である。

[図4]ユーザーが行う選択発話の一例を示す図である。

[図5]本発明の実施の形態に係る情報処理装置が実行する処理の流れの一例を示すフロー図である。

発明を実施するための形態

[0012] 以下、本発明の実施形態について、図面に基づき詳細に説明する。

[0013] 図1は、本発明の一実施形態に係る情報処理装置1の構成を示す構成ブロック図である。情報処理装置1は、例えば家庭用ゲーム機や携帯型ゲーム機、パーソナルコンピュータ、スマートホン等であって、図1に示すように、制御部11と、記憶部12と、インタフェース部13と、を含んで構成されている。また、情報処理装置1は、表示装置14、マイクロホン15、スピーカー16、及びカメラ17と接続されている。

[0014] 制御部11はCPU等を含んで構成され、記憶部12に記憶されているプ

プログラムを実行して各種の情報処理を実行する。本実施形態において制御部 11 が実行する処理の具体例については、後述する。記憶部 12 は、RAM 等のメモリデバイスを含み、制御部 11 が実行するプログラム、及び当該プログラムによって処理されるデータを格納する。インタフェース部 13 は、情報処理装置 1 が表示装置 14、マイクロホン 15、スピーカー 16、及びカメラ 17 との間で各種の情報を授受するためのインタフェースである。

[0015] 表示装置 14 は、ヘッドマウントディスプレイや家庭用テレビ受像機、液晶ディスプレイ等であって、情報処理装置 1 が出力する映像信号に応じた画像を画面上に表示する。特に表示装置 14 がヘッドマウントディスプレイのようにユーザーの視界を覆うタイプのデバイスである場合、ユーザーは自分の手元が見えなくなるため、コントローラ等の操作入力デバイスを手で操作しにくくなる。そこで、後述するように情報処理装置 1 がユーザーの声による情報の入力を受け付けることで、ユーザーは自分の手元が確認できない状況でも容易に情報の入力を行うことができる。マイクロホン 15 は、情報処理装置 1 のユーザーが発する声を集音して得られる音声信号を情報処理装置 1 に対して出力する。スピーカー 16 は、情報処理装置 1 が出力する音声信号に従って音声を鳴動させる。カメラ 17 は、ユーザーの様子を示す映像を撮像し、撮像された映像を情報処理装置 1 に入力する。なお、表示装置 14、マイクロホン 15、及びスピーカー 16 は、いずれも情報処理装置 1 の筐体内に内蔵されてもよいし、情報処理装置 1 と有線又は無線により接続される別個の装置であってもよい。

[0016] また、情報処理装置 1 は、ユーザーの頭部の動きを検出するために、3 軸加速度センサーや地磁気センサーなどの姿勢検出センサーを備えたデバイスと接続されてもよい。このようなデバイスをユーザーが装着することで、ユーザーの頭部の動作を検出できる。また、情報処理装置 1 は、ユーザーの操作入力を受け付ける操作入力デバイス（コントローラ等）と接続されてもよい。

[0017] 以下、情報処理装置 1 が実現する機能について、図 2 を用いて説明する。

図2に示すように、情報処理装置1は、機能的に、エージェント処理部21、音声信号取得部22、音声認識部23、及び選択結果特定部24を含んで構成されている。これらの機能は、制御部11が記憶部12に記憶されたプログラムに従って動作することにより実現される。このプログラムは、インターネット等の通信ネットワークを介して情報処理装置1に提供されてもよいし、光ディスク等のコンピュータ読み取り可能な情報記憶媒体に格納されて提供されてもよい。

[0018] エージェント処理部21は、ユーザーとコミュニケーションを行う仮想的なエージェントを実現し、エージェントによるユーザーとの会話処理を実行する。具体的にエージェント処理部21は、会話の一部として、ユーザーに対して選択肢の提示と、その選択肢に対するユーザーの選択の受け入れを行う。本実施形態においてエージェント処理部21は、表示装置14の画面に複数の選択肢を表示することによって、ユーザーに選択肢を提示する。

[0019] 図3は、選択肢を表示する画面の一例を示している。この図の例では、「どこに行く?」というエージェントからユーザーに向けた質問と、その質問に対する回答の候補として、「プール」、「日昇公園」、「梅ランド」、「クール猫パーク」の4個の選択肢が提示されている。さらにこの例では、日本語の各選択肢の語句に対して、英語に翻訳した語句と、選択肢を識別するラベル（ここでは数字）が並んで表示されている。また、この図の例では、4個の選択肢が上から順に並んでいるのではなく、画面中央から見て上下左右の4方向に配置されている。なお、エージェント処理部21は、選択肢を画面に表示させる代わりに、複数の選択肢を順に音声としてスピーカー16から再生してもよい。また、選択肢は画面に表示するが、質問だけは音声として再生してもよい。

[0020] 図3に例示されるような選択肢が提示されると、ユーザーはそこから自分が選択する内容に関する発話を行う。以下では、この選択肢の提示を受けてユーザーが行う発話を、選択発話という。また、ユーザーが選択する選択肢のことを選択結果という。後述するように、本実施形態では、ユーザーに

よる選択発話の内容は必ずしも選択肢の語句そのものと一致しておらずともよい。選択発話の内容は、音声信号取得部 22、音声認識部 23、及び選択結果特定部 24 によって処理され、これにより選択結果特定部 24 がユーザーの選択結果を特定する。エージェント処理部 21 は、特定された選択結果の情報を用いて、続くエージェントの発言内容を決定する。そして、決定したエージェントの発言内容を話す音声信号を生成し、スピーカー 16 に対して出力する。これにより、エージェントの発言がスピーカー 16 から音声として再生される。エージェント処理部 21 は、エージェントの発言を音声として再生する代わりに、表示装置 14 の画面に表示させてもよい。

[0021] エージェント処理部 21 は、ユーザーの選択結果だけでなく、選択発話の特徴を示す各種の特徴量を算出し、その算出結果を用いて続くエージェントの発言内容を決定してもよい。また、カメラ 17 によって撮像されたユーザーの表情や動作などに関する情報を用いてエージェントの発言内容を決定してもよい。ユーザーが選択肢の中から選択を行う場合にも、その選択を積極的行ったか、または消極的行ったかなど、様々な感情が込められていると考えられる。エージェント処理部 21 は、選択発話自体や、選択発話を行った際のユーザーの様子を解析することによって、ユーザーの感情や思考を反映する各種のパラ言語情報、非言語情報などに関する特徴量を求めることができる。このような特徴量を用いて以降のエージェントの振る舞いを決定することによって、単に選択肢の提示と選択結果の受け入れだけを行う場合と比較して、エージェント処理部 21 は人間同士の会話に近い自然なコミュニケーションを実現することができる。

[0022] また、エージェント処理部 21 は、エージェントの発言を音声として再生するだけでなく、エージェントを表すキャラクターなどの画像を表示装置 14 の画面に表示してもよい。さらに、ユーザーの発話内容やエージェントの発言内容などに応じてこのキャラクターが動作する様子を示す映像を表示してもよい。

[0023] 音声信号取得部 22 は、マイクロホン 15 によって集音された音声信号を

取得し、音声認識部 23 に対して出力する。本実施形態において音声信号取得部 22 は、エージェントの質問に対して応答するユーザーの声を含んだ音声信号を取得することになる。

[0024] 音声認識部 23 は、音声信号取得部 22 が取得した音声信号を解析することによって、ユーザーの発話内容を特定する。このような発話内容の特定は、隠れマルコフモデルを用いた統計的手法や動的時間伸縮法など、各種公知の音声認識技術などを用いて実現できる。特定されたユーザーの発話内容は、選択結果特定部 24 による選択結果の特定に用いられる。

[0025] なお、音声認識部 23 は、ユーザーが発話を終えてから音声認識処理を実行するのではなく、ユーザーが発話を行っている間、随時音声認識処理を実行してもよい。この場合、ユーザーが発話を続けている間、音声認識部 23 は、音声信号取得部 22 によって順次取得される所定時間分の音声信号に対して音声認識処理を実行し、その所定時間分の音声信号が何の音素に対応するかを特定する。このような処理を時間の経過に伴って繰り返し実行することによって、リアルタイムでユーザーの発話内容を特定することができる。

[0026] 選択結果特定部 24 は、音声認識部 23 による音声認識の結果特定される選択発話の内容に応じて、ユーザーの選択結果を特定する。ここで、選択肢として提示された語句そのもの（例えば「プール」など）をユーザーが発話し、かつ、その語句が音声認識部 23 によって正しく認識された場合、選択結果特定部 24 は、ユーザーが発話した選択肢をそのまま選択結果として特定すればよい。しかしながら、ユーザーは選択肢の語句そのものを発話するとは限らない。また、ユーザーが選択肢の語句そのものを発話した場合であっても、周囲の雑音やユーザー本人の滑舌、音声認識処理の精度などの問題から、ユーザーの発話内容を常に音声認識部 23 が正確に認識できるとは限らない。そこで選択結果特定部 24 は、音声認識部 23 による音声認識結果に基づいて、各種の判定処理を実行して、ユーザーの選択結果を特定する。

[0027] 以下、選択結果特定部 24 が選択結果を特定するために実行する処理の具体例について、説明する。まず、辞書を利用して選択結果を特定する具体例

について説明する。

[0028] 辞書を使用する第1の例として、複数の選択肢のそれぞれに対応する関連語辞書を用いる例について説明する。この例では、選択肢の語句のそれぞれに対して、選択肢の語句そのものと、これに関連する語句とを含んだ関連語辞書を予め用意しておく。例えば選択肢の語句が固有名詞の場合、関連語辞書にはその略称や愛称などが含まれてよい。選択結果特定部24は、音声認識部23によって認識されたユーザーの発話内容が関連語辞書に登録されたいずれかの語句と一致した場合に、その関連語辞書に対応する選択肢が選択されたと判定する。同じ語句についての様々な表現を予め関連語辞書に登録しておくことで、表示された選択肢をユーザーが別の表現で発話した場合にも、ユーザーの選択結果を特定できる。

[0029] また、関連語辞書は、選択肢の語句やその関連語句を様々な言語に翻訳した翻訳語を含んでもよい。図3の例では日本語と英語の双方で選択肢の語句が表示されているが、それ以外の外国語についても、選択肢の語句の翻訳表現を関連語辞書に登録しておく。こうすれば、ユーザーが外国語で選択肢の内容を発話した場合であっても、選択結果を特定できる。なお、関連語辞書には、外国語の翻訳表現だけに限らず、選択肢の語句を各地方の方言で表現した語句が含まれてもよい。また、語句の文字自体は同じであってもユーザーによっては発話時のイントネーションが異なり、このようなイントネーションの違いにより音声認識処理で違う語句と認識される可能性がある。そこで、関連語辞書は同じ語句を各種のイントネーションで発話した表現を含んでもよい。また、通常の英語表現のほかに、日本人訛りの英語（いわゆるカタカナ英語）などを含んでもよい。

[0030] 次に辞書を使用する第2の例として、選択肢に付加されるラベルを示すラベル辞書を用いる例について説明する。この例では、エージェント処理部21は、各選択肢を表示する際に、その選択肢を識別するためのラベルを付加して表示する。図3の例では、各選択肢に数字のラベルが付加されている。この場合にユーザーは、選択肢の語句の代わりに、選択肢に付加されたラベ

ルを示す語句を発話してもよい。例えば図3の例でプールを選択したい場合、ユーザーは「いち」と発話してもよい。選択肢に付加されるラベルは、数字に限らず、アルファベットやギリシア文字（「 α 」、「 β 」…）、フォネティックコード（「A l f a」、「B r a v o」…）、五十音（「あ」「い」「う」「え」…）、いろは仮名（「い」「ろ」「は」「に」…）など、各種のものあってよい。また、ラベルは「○」「△」「□」などの記号であってもよい。この場合、ユーザーは「まる」「さんかく」などと発話することによって、選択肢を選択できる。

[0031] この第2の例では、ラベルに対応する語句を登録したラベル辞書が予め用意され、エージェント処理部21が各選択肢にラベルを付加して提示する際に、各選択肢に付加されたラベルを示す語句を含んだラベル辞書が対応づけられる。選択結果特定部24は、ユーザーの発話内容がラベル辞書に含まれる語句に一致する場合、そのラベルが付加された選択肢を選択結果として特定する。関連語辞書の例と同様に、ラベル辞書には、「いち」や「まる」などのラベルを直接表す語句だけでなく、その語句を各国語に翻訳した語句や、イントネーション違いの語句などが含まれてよい。

[0032] また、ラベル辞書には、選択肢に付加されるラベルを示す語句に限らず、選択肢の表示態様を示す語句が含まれてもよい。このような表示態様を示す語句の例としては、選択肢の画面内における表示位置を示す語句が挙げられる。具体的に、図3の例では複数の選択肢が上下左右に並んで配置されている。この場合、ラベル辞書には、「うえ」「した」「みぎ」「ひだり」といった位置を表す語句が含まれてよい。こうすれば、ユーザーは、例えば「うえ」と発話することで上方向に表示された「プール」を選択できる。また、表示態様を示す語句は、表示順序を示す数字や序数などの語句を含んでもよい。こうすれば、選択肢に数字のラベルが付加されていない場合であっても、ユーザーは「いちばん」「ひとつめ」などの数字や序数を示す語句を発話することで、その数によって指定される選択肢を選択できる。また、各選択肢が互いに異なる文字色や背景色で表示されている場合、ラベル辞書はその

文字色や背景色（表示色）の名前（例えば「あか」「あお」など）を含んでよい。これにより、ユーザーは表示色で選択肢を指定できる。

[0033] 以上説明した第１の例、及び第２の例では、いずれも選択肢に対応づけられた辞書を用いてユーザーの発話内容が特定の語句に一致するか否かのマッチングを行うこととしているが、いずれの場合にも限られた語句だけを含んだ辞書を用いることになる。そのため、不特定多数の語句を含んだ辞書を用いて自然言語の音声認識処理を行う場合と比較すれば、処理負荷はそれほど高くなく、認識精度も高くなる。なお、選択結果特定部２４は、この第１の例と第２の例を組み合わせ用いてもよい。この場合、選択結果特定部２４は、関連語辞書及びラベル辞書の全てに含まれる語句をユーザーの発話内容と比較して、一致する語句を探せばよい。

[0034] 以上の説明では、音声認識部２３によって認識されたユーザーの発話内容が、いずれかの辞書に含まれる語句と完全に一致することを前提とした。しかしながら、ユーザーの発話内容は、様々な理由から各辞書に含まれる語句と完全に一致しない場合がある。そこで、選択結果特定部２４は、各種の判断基準を用いることによって、完全に一致しない場合であっても、ユーザーの発話内容が辞書に含まれる語句と対応すると判断してもよい。以下、このような判断基準のいくつかの例について、説明する。なお、以下の説明では、関連語辞書やラベル辞書に含まれる語句（すなわち、選択結果特定部２４がユーザーの発話内容と比較すべき語句）を、比較対象語句という。以下に例示する判断基準によってユーザーの発話内容がいずれかの比較対象語句に対応すると判断された場合、選択結果特定部２４は、その比較対象語句に対応する選択肢をユーザーが選択したとみなす。

[0035] 判断基準の第１の例として、装飾表現に関する判断基準について説明する。ユーザーは選択肢の語句やその関連語句、あるいは選択肢に付与されたラベルを示す語句等を単独で発話するとは限らず、様々な語句（装飾表現）を前後に付け加える場合がある。特にユーザーが通常の会話に近い態度で発話をする場合、ユーザーは単独の語句を発話するよりも比較対象語句を含んだ

センテンスを話す可能性が高い。例えば図3の例では、1番のプールを選択しようとする場合に、「プールかな…」「イチかな。」「イチでしょ！」など、選択肢の語句やラベルを示す語句等を一部に含んだ内容を発話すると考えられる。選択結果特定部24は、このような場合であってもユーザーの選択結果を正しく特定することが望まれる。そこで選択結果特定部24は、ユーザーの発話内容に含まれる一部の音素列が比較対象語句と一致する場合、一致する比較対象語句に対応する選択肢を選択結果として特定する。このようにユーザーの発話全体のうちの一部の語句が比較対象語句とマッチするかを選択結果特定部24が判断することで、ユーザーは自然な会話に近い態度で選択肢を選ぶ発話を行うことができる。

[0036] 次に第2の例として、表現の揺らぎに関する判断基準について説明する。ユーザーの発話内容は、音声認識処理における誤認識やユーザーの発話の癖などによって、比較対象語句に近いが完全には一致しない語句として認識される場合がある。このような場合に対応するために、選択結果特定部24は、比較対象語句とユーザーの発話内容との間で表現の揺らぎとして予め定められたパターンのずれが生じている場合に、両者が一致すると判断する。

[0037] 具体例として、ユーザーの滑舌が悪い場合、特に子音が誤認識される場合がある。そこで選択結果特定部24は、ユーザーの発話内容と比較対象語句とで母音が一致した場合に、子音が一致せずとも両者が対応すると判断してもよい。一例として、「イキ」「イシ」などの語句がユーザーの発話内容に含まれる場合に、選択結果特定部24はユーザーが日本語の「イチ」と発話したと推定し、1番目の選択肢を選択結果として特定してもよい。また、英語の「p e a c h」が比較対象語句に含まれており、ユーザーの発話内容に対する音声認識結果に、「b e a c h」のように一部の子音だけが異なる別の語句が含まれる場合、「p e a c h」を選択結果として特定してもよい。

[0038] また、選択結果特定部24は、ユーザーの発話内容から長音を省略した語句が比較対象語句に一致する場合に、両者が対応すると判断してもよい。例えばユーザーが日本語で「イーチ」と発話した場合に、1番目の選択肢を選

択結果として特定する。同様に、英語など他の言語においても、「a l p h a a a」などのように同じ母音が連続する語句をユーザーが発話した場合に、母音を短縮した比較対象語句が選択されたと判断してもよい。また、訛りや方言などによって変化する音がある場合、選択結果特定部24は、そのような音の変化した語句を比較対象語句に対応すると判断する判断基準を用いてもよい。例えば、「シガグ」というユーザーの発話を日本語の「シカク」を意味するものと判断してもよい。また、イギリス英語とアメリカ英語で同じ単語の発音の仕方に違いがある場合に、このような音の変化した語句を比較対象語句に対応すると判断してもよい。具体的には、v i t a m i nにおける最初のiの発音や、h e r bの最初のhの発音などのように、イギリス英語とアメリカ英語で発音の仕方に違いがあることが知られている。このような判断基準を用いて比較対象語句とユーザーの発話内容とを比較することで、方言などにより表現が変化した語句を全て辞書に含めなくとも、このような語句が比較対象語句と対応していると判断することができる。

[0039] 次に第3の例として、部分一致の判断基準について説明する。この例では、選択結果特定部24は、比較対象語句とユーザーの発話内容とが所定の基準の範囲内で部分的に一致する場合に、両者が対応すると判断する。例えば人名は、その前半部分が元の名前と一致するが後半部分が一致しない愛称で呼ばれることがある。具体的には、「マルサス」という名前であれば「マルちゃん」「マルくん」「マルさん」などの呼ばれ方をする可能性があり得る。また、「クリスティーナ」は、「クリス」、「クリスィー」などと呼ばれる可能性がある。そこで、例えば選択結果特定部24は、先頭の所定数の音素、または先頭の所定割合の音素が比較対象語句と一致する語句がユーザーの発話内容に含まれる場合に、ユーザーの発話内容と比較対象語句とが対応すると判断する。こうすれば、ユーザーが選択肢の語句の一部を変化させた表現を用いた場合に、このような変化した表現を網羅的に辞書に登録しておかなくとも、選択結果を正しく特定できる。なお、人名に限らず、選択肢の語句が「c u t e」でユーザーが「c u t y」と発話した場合など、語尾が

活用する語句などの場合にも、同様の判断基準でユーザーの選択結果を特定することができる。

[0040] 以上説明した判断基準は、組み合わせて用いられてよい。これにより、ユーザーが多様な表現で選択肢を選ぶ発話を行ったとしても、選択結果特定部24はユーザーの選択結果を特定することができる。

[0041] 前述したように、エージェント処理部21は、選択結果特定部24によって特定された選択結果に応じて、続くエージェントの発言を決定する。さらにエージェント処理部21は、ユーザーの選択結果だけでなく、選択発話から読み取られるユーザーの感情や態度に応じて、その後の発言を変化させてもよい。発話内容からユーザーの感情や態度を判断する手法としては、公知のものが利用可能である。

[0042] 具体的に、エージェント処理部21は、選択発話を記録した音声信号から算出されるパラ言語情報の特徴量を用いてユーザーの感情を推定する。パラ言語情報は、話速、音量（声圧）、声の抑揚、イントネーション（ユーザーが強調する語句など）、言葉遣い（丁寧口調かくだけた口調かなど）、発言量などを含んでよい。また、マイクロホン15が取得する音声信号に限らず、カメラ17によって撮像されるユーザーの映像から得られる各種の情報を利用してもよい。映像から得られる情報としては、ユーザーの表情、視線、瞬き、また頷きや手振りなどの動作（ジェスチャー）が挙げられる。

[0043] また、エージェント処理部21は、選択肢が提示されてからユーザーが選択発話を開始するまでの経過時間（以下、思考時間という）を特徴量の一つとして算出し、感情の評価処理に用いてもよい。思考時間が短いほど、ユーザーは迷いなく選択肢を選んでいると推定され、思考時間が長い場合、ユーザーは選択の際に迷いがあったと考えられる。このような情報を参照することで、ユーザーの感情を読み取ることができる。思考時間を用いてユーザーの感情を評価する場合、ユーザーが選んだ選択結果が選択肢として提示された際の提示順序に応じて、評価基準を変化させてもよい。複数の選択肢が画面上に表示される場合、ユーザーは上から順に選択肢を見て、どの選択肢を

選ぶか決定すると考えられる。つまり、後に提示される選択肢に目を通すまでには時間がかかると想定される。そのためエージェント処理部21は、後に提示される選択肢が選択結果として特定された場合、先に提示される選択肢が選択結果として特定された場合と比較して、思考時間が実際より短いものとして感情の評価を行う。こうすれば、ユーザーが選択肢を視認してから選択発話を行うまでの時間に応じて評価を行うことができる。

[0044] また、前述したようにユーザーは選択発話において選択肢の語句そのものを発話する代わりに、ラベルを示す語句を発話する場合がある。このような場合にも、その語句が選択肢の語句を指しているものとして、感情の評価を行えばよい。例えばユーザーが「Aだよ」という発話を行った場合、Aというラベルが「かわいい」という選択肢に付加されていたのであれば、エージェント処理部21はユーザーが「かわいい」と発話したときと同様の評価基準で感情の評価を行う。その結果、ユーザーが本心からかわいいと思っているか、あるいは皮肉っぽく発言したのか、などを評価することができる。一方、Aというラベルが「リンゴ」という選択肢に付加されていた場合には、ユーザーがリンゴについて発言したという前提の下に感情の評価を行う。この場合には、その発言が皮肉かどうか、などの評価を行う必要はない。

[0045] エージェント処理部21は、以上説明した感情の評価処理によって得られる評価結果を、発言内容の決定に用いてもよいし、その他の出力内容に反映させてもよい。例えばエージェント処理部21がゲームの処理を実行する場合、評価結果に応じてゲームの進行を変化させてもよい。この例では、ユーザーが選択発話を強い感情で行ったと評価される場合、ゲームで有利な効果が発生させてもよい。また、エージェント処理部21は、評価結果そのものをユーザーに提示してもよい。また、感情の評価結果によっては、エージェント処理部21は、選択結果特定部24によって特定された選択結果を変更してもよい。一例として、ユーザーがある選択結果を選んだ際に、選択発話に対して感情の評価処理を実行した結果、その発言に消極的であったと評価された場合、エージェント処理部21は別の選択肢が選択結果として特定さ

れたものとして処理を続行してもよい。

[0046] 以上説明したように、情報処理装置 1 は、感情の評価結果を利用することで、複数の選択肢のうちの一つをユーザーに選択させる場合であっても、選択肢の数以上にその後の処理内容を様々に変化させることができる。また、感情の評価結果をフィードバックした結果をユーザーに提示することで、情報処理装置 1 はユーザーが自然に感情を込めた発話を行うように誘導することができる。これにより、選択肢をユーザーに選ばせるにもかかわらず、人間同士の会話に近い自然なコミュニケーションを行っているようにユーザーに感じさせることができる。

[0047] 以下、エージェント処理部 2 1 による会話のタイミング制御に関して説明する。通常、エージェント処理部 2 1 は、音声認識部 2 3 によってユーザーの発話の終了が検出されてから、その発話内容を解析した結果に基づいてエージェントの次の発言の内容を決定し、決定した発言を出力する。しかしながら、周囲の騒音などが原因でユーザーの発話が終了したことを検出できない場合がある。そこで音声認識部 2 3 は、選択肢が提示されてから所定時間が経過しても発話の終了が検出できない場合に、無条件で音声認識処理を終了し、それまで取得された音声信号に対する認識処理結果を用いて次の処理を実行してもよい。一般的に、自由会話の場合と異なり、選択肢の提示に対するユーザーの回答にはそれほど時間がかからないと考えられることから、所定時間で音声認識処理を終了しても、その中にユーザーの選択内容が含まれる可能性は高いと想定される。

[0048] また、前述したようにユーザーの発話中にリアルタイムで音声認識処理を実行する場合、選択結果特定部 2 4 もリアルタイムで音声認識処理の結果を用いて選択発話内から比較対象語句と対応する語句の検出を行ってもよい。その結果、対応する語句が検出されれば、その結果に基づいて選択結果特定部 2 4 は選択結果を特定できるので、その時点で音声認識部 2 3 による音声認識処理を中断してもよい。こうすれば、ユーザーが発話を終了した後、間を置かずに適切なタイミングでエージェントの発言を再生させることができ

、音声認識処理による会話の遅延を回避できる。さらに、選択結果の特定ができた後にユーザーの発話の終了が検出できない状態が所定時間続いた場合、エージェント処理部 21 は次の発言の再生を開始してもよい。選択結果の特定ができていれば、ユーザーの発話の終了が検出できない場合であっても、自然にユーザーの選択発話を受けた次の会話を提示することができる。

[0049] 具体的なタイミング制御について、図 4 を例として説明する。図 4 は、ユーザーが行った選択発話の一例を示している。この図の例では、ユーザーは選択肢が提示されたタイミング（時刻 t_0 ）から少し間を空けて選択発話を始め（時刻 t_1 ）、さらに選択発話中においても冒頭にフィラーの発声（すなわち、言語的に意味をなさない発声）をしてから（時刻 t_2 ）、選択した選択肢に対応するラベルを示す語句（「イチ」）を発話している。この語句自体の発話は時刻 t_3 に終了しているが、その後もユーザーは発話を続け、時刻 t_4 に発話を終了している。このような場合において、選択結果特定部 24 は、時刻 t_3 を経過した後、それまでに取得された音声信号に対する音声認識処理の結果を用いて、ユーザーが 1 番の選択肢を選択したことを特定できる。この特定処理が時刻 t_4 よりも前に完了すれば、エージェント処理部 21 はユーザーを待たせずにユーザーの発話に対するエージェントの応答を出力することができる。この場合、音声認識部 23 は、選択結果特定部 24 が選択結果の特定を完了した時点で、それ以降の選択発話に対する音声認識処理を中断してもよい。また、エージェント処理部 21 は、時刻 t_3 までに得られた音声信号を用いて、ユーザーの感情の評価処理を実行してもよい。

[0050] なお、エージェント処理部 21 は、選択結果特定部 24 が選択結果として特定した発話部分（図 4 の例における時刻 t_2 から時刻 t_3 まで）とその前後の部分に対して、それぞれ独立に感情の評価処理を実行してもよい。これにより、ユーザーが選択を迷っている状態、決定した選択内容を発話している状態、及び発話し終えた後の状態のそれぞれについて感情の評価を行うことができる。なお、エージェント処理部 21 は、これらの状態の全てについ

て感情の評価処理を行うのではなく、一部の状態についてのみ感情の評価処理を実行してもよい。

[0051] また、エージェントの発言をエージェント処理部 21 が音声として再生している間に、ユーザーがその発言を遮って自分の発言を始めたい場合がある。特にエージェントの発言時間が長くなると、このような場合が生じ得る。また、エージェントの発言が複数の選択肢を順に提示するものである場合、ユーザーが最初の方に提示された選択肢を選択しようとしているのであれば、ユーザーは後の選択肢を聞かずに選択発話を開始したいこともあると考えられる。そこで音声認識部 23 は、エージェントの発言が再生されている間にも、ユーザーの発話を監視し、発話内容の認識処理を実行してもよい。

[0052] 音声認識部 23 は、特にユーザーがエージェントの発言を遮って発話をする可能性のある状態に限って、上述したような監視処理を実行してもよい。例えばエージェント処理部 21 が再生しようとするエージェントの発言が所定の長さ未満であれば、音声認識部 23 はエージェントの発言が完了するまでユーザーの発話は無視することとし、所定の長さ以上の場合にユーザーの発話を監視してもよい。この場合の所定の長さは、例えば再生時間を単位として決定されてもよいし、語数、モーラ数などによって決定されてもよい。

[0053] さらに音声認識部 23 は、ユーザーの発話の検出処理の内容を時間とともに変化させてもよい。具体的に、エージェントが発言を開始したばかりの状態では、ユーザーがその発言を遮りたい場合は少ないと想定される。一方で、エージェントの発言がある程度続いていると、ユーザーがエージェントの発言を遮って自分の発話を開始したいという欲求は高まると考えられる。そこで、音声認識部 23 は、エージェントの発言開始時を起算点として、所定の長さの発言が再生されるまではユーザーの発話の検出を行わず、所定の長さの発言が再生されたタイミング以降、ユーザーの発話を監視してもよい。また、監視自体はエージェントの発言開始時から開始するが、ユーザーがエージェントの発言を遮る発話を行ったと検出する検出基準を、所定の長さの発言がされる前後で変更してもよい。具体的に、例えばエージェント処理部

21は、エージェントが所定の長さの発言を終えるまでは、ユーザーが「やめて」などの予め定められた語句を発話したことを音声認識部23が検出した場合に限ってエージェントの発言を中止することとし、エージェントが所定の長さを超えて発言している状態では、ユーザーが何らかの発話を行ったと判定された場合にエージェントの発言を中止してもよい。

[0054] 以下、本実施形態に係る情報処理装置1が実行する処理の流れの一例について、図5のフロー図を用いて説明する。

[0055] まずエージェント処理部21は、ユーザーに対する質問、及びその回答候補となる複数の選択肢を表示装置14の画面に表示させてユーザーに提示する(S1)。その後、音声信号取得部22による音声信号の取得、及び音声認識部23による取得された音声信号に対する音声認識処理が開始される(S2)。ユーザーが選択発話を行っている間、所定時間が経過するごとに、音声認識部23は、ユーザーの発話が終了したか否かを判定する(S3)。具体的に音声認識部23は、ユーザーの音声が一定時間検出できなかったり、S1の選択肢の提示から所定時間以上の時間が経過したりした場合に、ユーザーの発話が終了したと判定する。

[0056] ユーザーの発話が終了していなければ、選択結果特定部24がユーザーの選択結果を特定済みか否かを判定する(S4)。特定済みであれば、S3に戻ってユーザーの発話の終了が検出されるまで処理が繰り返される。まだ特定済みでなければ、選択結果特定部24は、それまでに得られる音声認識処理の結果を辞書に含まれる比較対象語句と対比して、選択結果の特定を試みる(S5)。選択結果が特定できれば、選択結果特定部24はその選択結果をエージェント処理部21に対して出力する(S6)。その後、S3に戻ってユーザーの発話が終了したか否かの判定が行われる。

[0057] S3でユーザーの発話が終了した場合、エージェント処理部21はそれまで得られた音声信号に基づいてユーザーの感情を評価する評価処理を実行する(S7)。そして、S6で出力されるユーザーの選択結果と、S7における感情の評価処理の結果に基づいて、エージェントの次の発言内容を決定し

、出力する（S8）。これにより、ユーザーの選択に応じた会話が行われる。

[0058] なお、以上の説明ではエージェント処理部21がユーザーに選択肢を提示する場合の処理について説明したが、これに限らず、エージェント処理部21は選択肢を提示せずにユーザーに任意の回答が可能な質問を提示してもよい。また、エージェント処理部21は、まず任意の回答が可能な質問をユーザーに提示し、これに対するユーザーの回答の認識に失敗した場合に、その質問に対する回答候補の選択肢を提示するようにしてもよい。

[0059] 以上説明した本実施形態に係る情報処理装置1によれば、選択肢を提示することでユーザーの回答内容を絞り込むことができ、誤認識の可能性を低減できる。その一方で、ユーザーが選択肢の語句そのものを回答しない場合であっても、ユーザーの回答がどの選択肢を選択するものなのかを特定できるので、ユーザーに自然な発言での回答を促すことができる。

[0060] なお、本発明の実施の形態は、以上説明したものに限られない。以上の説明では選択肢はエージェントによる質問とともに提示されることとしたが、情報処理装置1は、エージェントによらずに単にユーザーに質問と選択肢を提示してもよい。

[0061] また、以上の説明ではユーザーのすぐそばにあるローカルの情報処理装置1がエージェントの発言内容の決定、音声認識処理、選択結果の特定、感情の評価処理などの処理を全て実行することとしたが、これらの処理の一部、又は全部は通信ネットワークを介して接続されたサーバ装置によって実行されてもよい。特にサーバ装置が選択肢を決定し、ローカルの端末装置から得られる音声信号に対する音声認識処理を行って選択結果を特定する場合、このサーバ装置が本発明における情報処理装置として機能することになる。

符号の説明

[0062] 1 情報処理装置、11 制御部、12 記憶部、13 インタフェース部、14 表示装置、15 マイクロホン、16 スピーカー、17 カメラ、21 エージェント処理部、22 音声信号取得部、23 音声認識部

、 2 4 選択結果特定部。

請求の範囲

- [請求項1] 複数の選択肢をユーザーに提示する選択肢提示部と、
前記複数の選択肢のいずれかを選択する前記ユーザーの発話内容を認識する音声認識部と、
前記認識されるユーザーの発話内容に含まれる語句が、予め用意された前記複数の選択肢のそれぞれに対応する辞書に含まれる語句と対応するか否かに基づいて、前記ユーザーが選択した選択肢を特定する選択結果特定部と、
を含むことを特徴とする情報処理装置。
- [請求項2] 請求項1に記載の情報処理装置において、
前記複数の選択肢のそれぞれに対応する辞書は、当該選択肢の語句に関連する語句、及び当該選択肢の語句を他の言語に翻訳した語句のいずれか少なくとも一つを含む
ことを特徴とする情報処理装置。
- [請求項3] 請求項1に記載の情報処理装置において、
前記選択肢提示部は、前記複数の選択肢のそれぞれに対して、当該選択肢を識別するラベルを付加して前記ユーザーに提示し、
前記複数の選択肢のそれぞれに対応する辞書は、当該選択肢に付加されたラベルを示す語句を含む
ことを特徴とする情報処理装置。
- [請求項4] 請求項1に記載の情報処理装置において、
前記複数の選択肢のそれぞれに対応する辞書は、当該選択肢の表示態様を示す語句を含む
ことを特徴とする情報処理装置。
- [請求項5] 請求項4に記載の情報処理装置において、
前記選択肢の表示態様を示す語句は、前記選択肢の表示位置、表示順序、及び、表示色のいずれか少なくとも一つを含む
ことを特徴とする情報処理装置。

- [請求項6] 請求項 1 に記載の情報処理装置において、
前記選択結果特定部は、前記発話内容に含まれる語句が、前記辞書に含まれる語句の一部と一致する場合に、両者が対応すると判定することを特徴とする情報処理装置。
- [請求項7] 請求項 1 に記載の情報処理装置において、
前記選択結果特定部は、前記発話内容に含まれる語句と前記辞書に含まれる語句との間のずれが予め定められたパターンに該当する場合に、両者が対応すると判定することを特徴とする情報処理装置。
- [請求項8] 請求項 1 に記載の情報処理装置において、
前記ユーザーの発話を含む音声信号の特徴量を算出することによって、前記ユーザーの感情を評価し、当該感情の評価結果と、前記ユーザーが選択した選択肢と、に応じた出力を行う出力部をさらに含むことを特徴とする情報処理装置。
- [請求項9] 請求項 8 に記載の情報処理装置において、
前記特徴量は、前記複数の選択肢が前記ユーザーに提示されてから、前記ユーザーが前記複数の選択肢のいずれかを選択する発話を行うまでの経過時間を含み、
前記出力部は、前記経過時間を用いて前記ユーザーの感情を評価する際に、前記選択結果特定部が特定する選択肢が前記ユーザーに提示された際の提示順序に応じて、評価基準を変化させることを特徴とする情報処理装置。
- [請求項10] 請求項 1 に記載の情報処理装置において、
ユーザーに提示すべき発言を音声で再生する再生部をさらに含み、
前記再生部は、前記発言の再生中に前記音声認識部が前記ユーザーの発話を検出した場合に、前記発言の再生を中止することを特徴とする情報処理装置。
- [請求項11] 請求項 10 に記載の情報処理装置において、

前記音声認識部は、前記発言の長さに応じて、前記発言の再生中に前記ユーザーの発話を検出するか否かを決定する

ことを特徴とする情報処理装置。

[請求項12]

請求項10に記載の情報処理装置において、

前記音声認識部は、前記発言の再生開始からの経過時間に応じて、前記ユーザーの発話の検出基準を変化させる

ことを特徴とする情報処理装置。

[請求項13]

ユーザーに提示すべき発言を音声で再生する再生部と、

前記ユーザーの発話内容を認識する音声認識部と、

を含み、

前記再生部は、前記発言の再生中に前記音声認識部が前記ユーザーの発話を検出した場合に、前記発言の再生を中止する

ことを特徴とする情報処理装置。

[請求項14]

請求項13に記載の情報処理装置において、

前記音声認識部は、前記発言の長さに応じて、前記発言の再生中に前記ユーザーの発話を検出するか否かを決定する

ことを特徴とする情報処理装置。

[請求項15]

請求項13に記載の情報処理装置において、

前記音声認識部は、前記発言の再生開始からの経過時間に応じて、前記ユーザーの発話の検出基準を変化させる

ことを特徴とする情報処理装置。

[請求項16]

複数の選択肢をユーザーに提示するステップと、

前記複数の選択肢のいずれかを選択する前記ユーザーの発話内容を認識するステップと、

前記認識されるユーザーの発話内容に含まれる語句が、予め用意された前記複数の選択肢のそれぞれに対応する辞書に含まれる語句と対応するか否かに基づいて、前記ユーザーが選択した選択肢を特定するステップと、

を含むことを特徴とする情報処理方法。

[請求項17]

複数の選択肢をユーザーに提示する選択肢提示部、

前記複数の選択肢のいずれかを選択する前記ユーザーの発話内容を認識する音声認識部、及び、

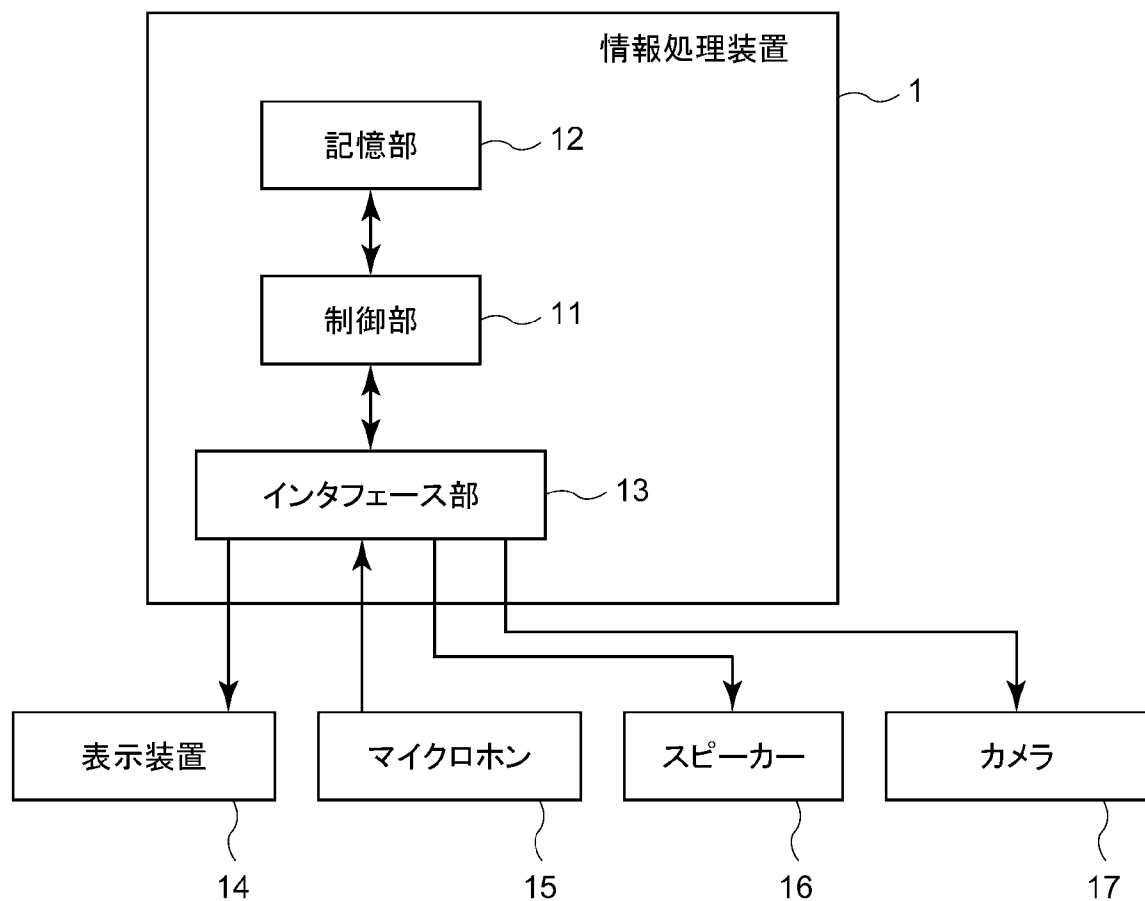
前記認識されるユーザーの発話内容に含まれる語句が、予め用意された前記複数の選択肢のそれぞれに対応する辞書に含まれる語句と対応するか否かに基づいて、前記ユーザーが選択した選択肢を特定する選択結果特定部、

としてコンピュータを機能させるためのプログラム。

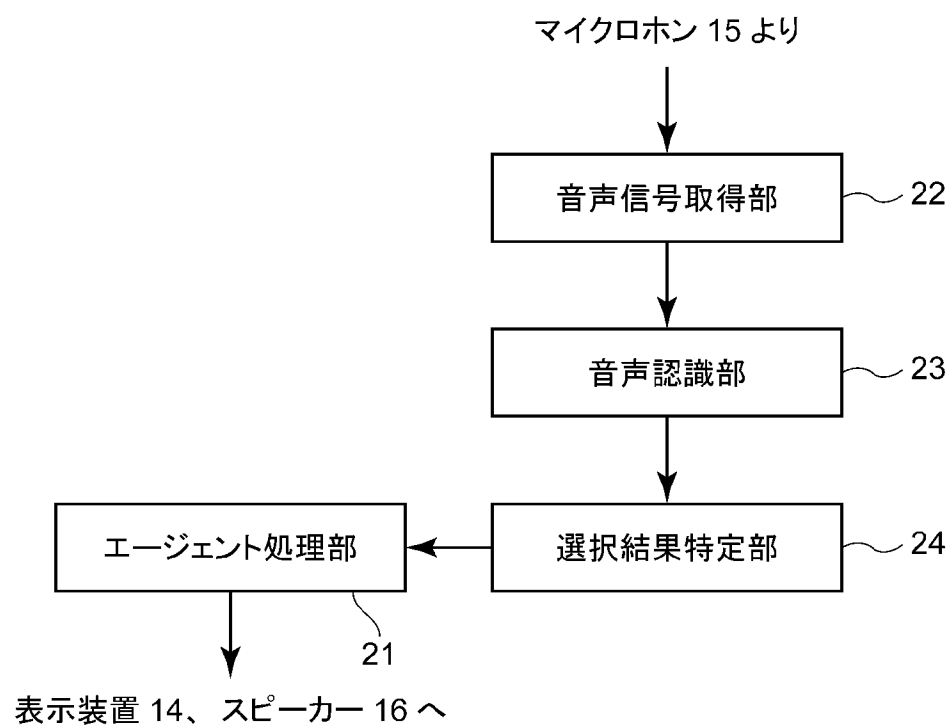
[請求項18]

請求項17に記載のプログラムを記憶した、コンピュータ読み取り可能な情報記憶媒体。

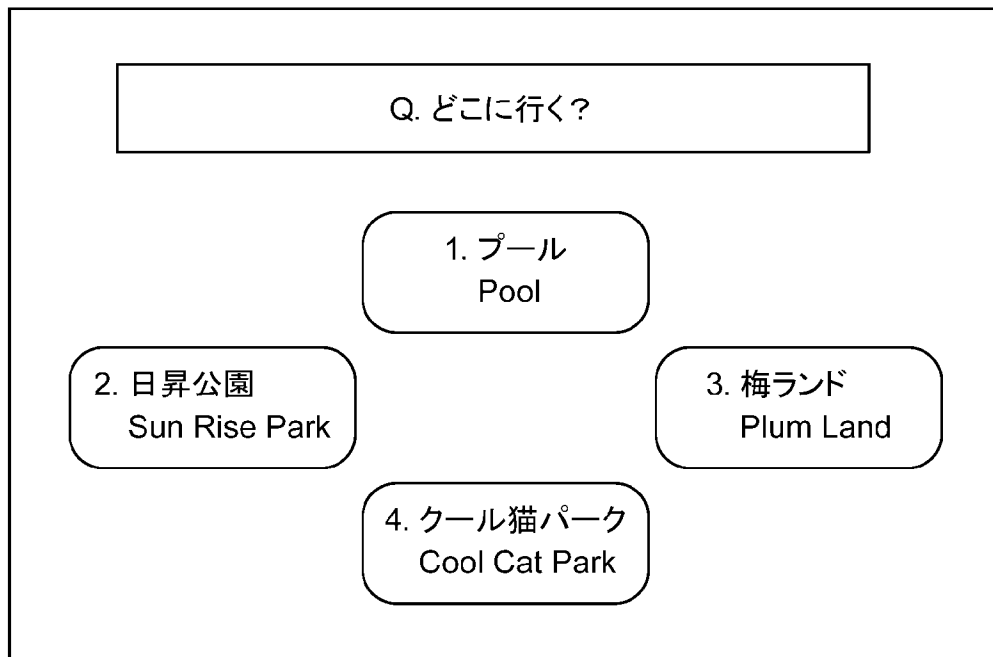
[図1]



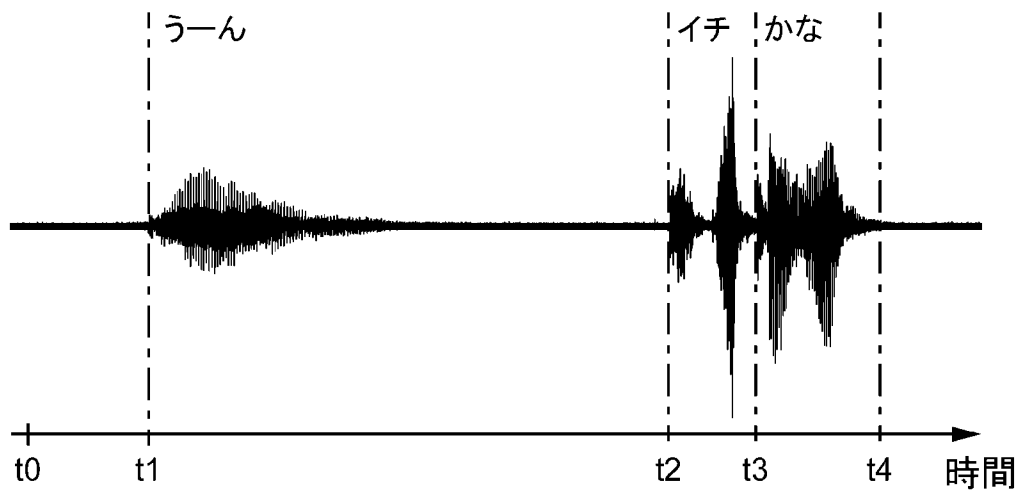
[図2]



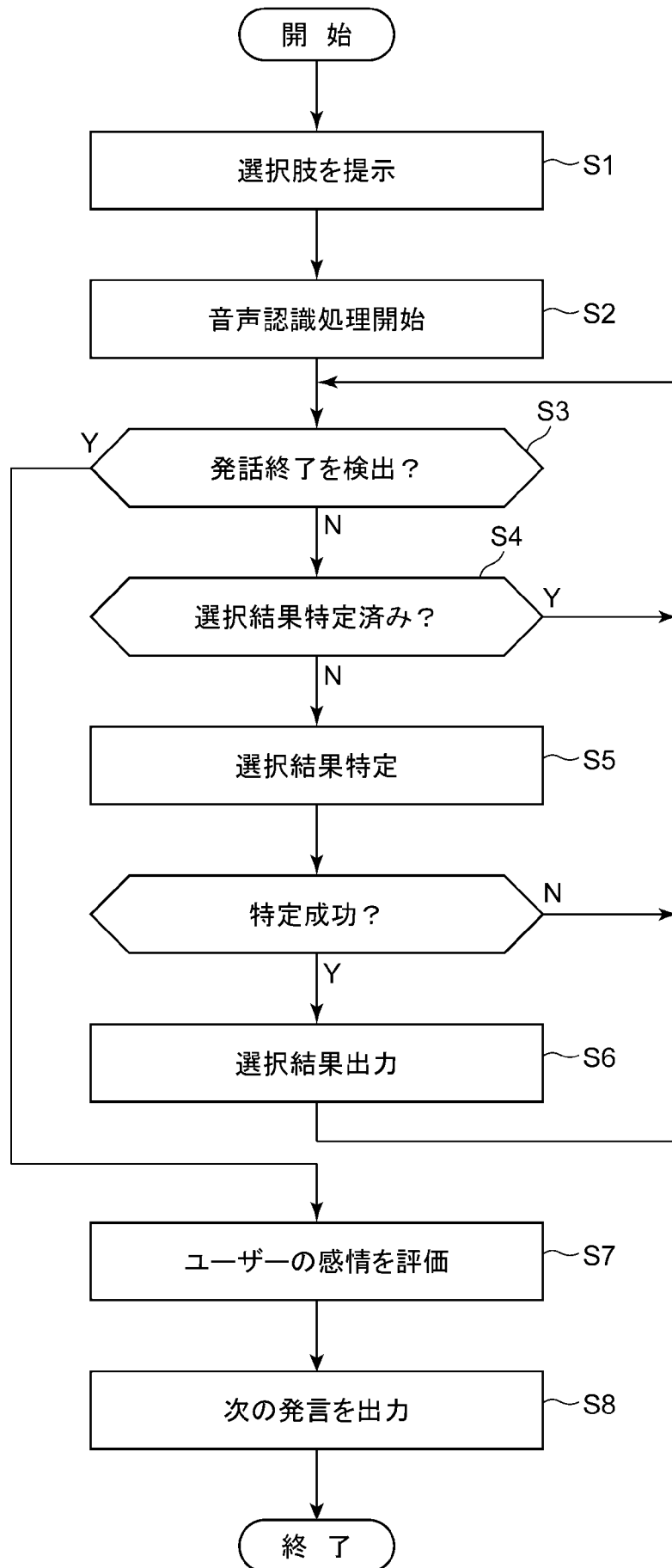
[図3]



[図4]



[図5]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2016/061385

A. CLASSIFICATION OF SUBJECT MATTER

G10L15/22(2006.01)i, G10L13/00(2006.01)i, G10L15/10(2006.01)i, G10L15/28(2013.01)i

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)

G10L15/22, G10L13/00, G10L15/10, G10L15/28

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

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2016
Kokai Jitsuyo Shinan Koho	1971-2016	Toroku Jitsuyo Shinan Koho	1994-2016

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y A	JP 2013-114020 A (Seiko Epson Corp.), 10 June 2013 (10.06.2013), paragraphs [0029] to [0031], [0061] to [0068]; fig. 1, 2, 8 (Family: none)	1, 16-18 2-8, 10-11, 13-14 9, 12, 15
Y	JP 2006-343661 A (Sanyo Electric Co., Ltd.), 21 December 2006 (21.12.2006), paragraphs [0020], [0029]; fig. 1, 3 (Family: none)	2
X Y	JP 2010-54549 A (Gifu University), 11 March 2010 (11.03.2010), paragraphs [0029] to [0031], [0038]; fig. 2, 3 (Family: none)	1, 16-18 6-7

☒ 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
10 June 2016 (10.06.16)

Date of mailing of the international search report
21 June 2016 (21.06.16)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2016/061385

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2009-271258 A (Toshiba Corp.), 19 November 2009 (19.11.2009), paragraph [0004] (Family: none)	3
Y	JP 2011-171948 A (Advance Co., Ltd.), 01 September 2011 (01.09.2011), paragraph [0027] (Family: none)	4-5
Y	JP 2007-65226 A (Advanced Telecommunications Research Institute International), 15 March 2007 (15.03.2007), paragraph [0028]; fig. 1 & US 2009/0089051 A1 paragraph [0069]; fig. 1 & WO 2007/026436 A1	8
Y	JP 8-146991 A (Canon Inc.), 07 June 1996 (07.06.1996), paragraphs [0003], [0011] to [0017]; fig. 1 (Family: none)	10-11, 13-14

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G10L15/22(2006.01)i, G10L13/00(2006.01)i, G10L15/10(2006.01)i, G10L15/28(2013.01)i

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））

Int.Cl. G10L15/22, G10L13/00, G10L15/10, G10L15/28

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	1971-2016年
日本国実用新案登録公報	1996-2016年
日本国登録実用新案公報	1994-2016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）

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 Y A Y X Y	JP 2013-114020 A（セイコーエプソン株式会社）2013.06.10, [0029] - [0031]、[0061] - [0068]、第1, 2, 8図（ファミリーなし） JP 2006-343661 A（三洋電機株式会社）2006.12.21, [0020]、[0029]、第1, 3図（ファミリーなし） JP 2010-54549 A（国立大学法人岐阜大学）2010.03.11, [0029] - [0031]、[0038]、第2, 3図 （ファミリーなし）	1, 16-18 2-8, 10-11, 13 -14 9, 12, 15 2 1, 16-18 6-7

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、一般的技術水準を示すもの	「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矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「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	「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
「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）	「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
「O」 口頭による開示、使用、展示等に言及する文献	「&」 同一パテントファミリー文献
「P」 国際出願日前で、かつ優先権の主張の基礎となる出願	

国際調査を完了した日

10.06.2016

国際調査報告の発送日

21.06.2016

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

大野 弘

5Z

9175

電話番号 03-3581-1101 内線 3591

C (続き) . 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
Y	JP 2009-271258 A (株式会社東芝) 2009. 11. 19, [0004] (ファミリーなし)	3
Y	JP 2011-171948 A (株式会社アドバンス) 2011. 09. 01, [0027] (ファミリーなし)	4-5
Y	JP 2007-65226 A (株式会社国際電気通信基礎技術研究所) 2007. 03. 15, [0028]、第1図 & US 2009/0089051 A1, [0069]、第1図 & WO 2007/026436 A1	8
Y	JP 8-146991 A (キヤノン株式会社) 1996. 06. 07, [0003]、[0011] - [0017]、第1図 (ファミリーなし)	10-11, 13-14