(19)日本国特許庁(JP)　(12)公表特許公報(A)　(11)特許出願公表番号
特表2011-524991
(P2011-524991A)
(43)公表日　平成23年9月8日(2011.9.8)

(51)Int.Cl.		F I			テーマコード（参考）
G10L 15/06	(2006.01)	G10L	15/06	300J	5B091
G10L 15/00	(2006.01)	G10L	15/00	200C	5D015
G10L 15/22	(2006.01)	G10L	15/22	470F	
G06F 17/28	(2006.01)	G06F	17/28	P	
		G06F	17/28	V	

審査請求 未請求　予備審査請求 有　（全 36 頁）

(21)出願番号　特願2011-505169 (P2011-505169)
(86)(22)出願日　平成21年4月15日 (2009.4.15)
(85)翻訳文提出日　平成22年12月6日 (2010.12.6)
(86)国際出願番号　PCT/US2009/040677
(87)国際公開番号　WO2009/129315
(87)国際公開日　平成21年10月22日 (2009.10.22)
(31)優先権主張番号　61/092,581
(32)優先日　平成20年8月28日 (2008.8.28)
(33)優先権主張国　米国 (US)
(31)優先権主張番号　61/045,079
(32)優先日　平成20年4月15日 (2008.4.15)
(33)優先権主張国　米国 (US)
(31)優先権主張番号　61/093,818
(32)優先日　平成20年9月3日 (2008.9.3)
(33)優先権主張国　米国 (US)

(71)出願人　510273640
モバイル　テクノロジーズ，エルエルシー
ＭＯＢＩＬＥ　ＴＥＣＨＮＯＬＯＧＩＥＳ，ＬＬＣ
アメリカ合衆国　１５２２１　ペンシルベニア，ピッツバーグ，ウインザー　アベニュー　６１９
(74)代理人　100066728
弁理士　丸山　敏之
(74)代理人　100100099
弁理士　宮野　孝雄
(74)代理人　100111017
弁理士　北住　公一
(74)代理人　100119596
弁理士　長塚　俊也

最終頁に続く

(54)【発明の名称】現場にて音声－音声翻訳をメンテナンスするシステム及び方法

(57)【要約】
【解決手段】方法及び装置は、第１言語を書き言葉及び話し言葉を含む第２言語に翻訳する音声翻訳システムの語彙を更新する。この方法は、第１言語における新語を第１言語の第１認識辞書に加えるステップと、発音及び語クラス情報を含む説明を新語と関連付けるステップとを含む。次に、これらの新語及び説明は、第１言語と関連付けられた第１機械翻訳モジュールにおいて更新される。この第１機械翻訳モジュールは、第１タグ付けモジュール、第１翻訳モデル及び第１言語モジュールを含み、新語を第２言語において対応する翻訳語に翻訳するように構成される。選択的に、本発明は、双方向の又は多方向の翻訳に利用されてもよい。
【選択図】図１

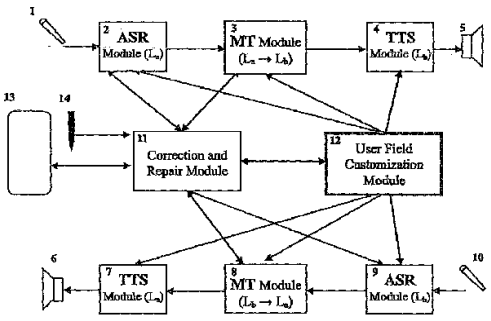


FIG. 1

【特許請求の範囲】
【請求項１】
第１言語を書き言葉及び話し言葉を含む第２言語に翻訳する音声翻訳システムの語彙を更新する方法であって、
前記第１言語における新語を前記第１言語の第１認識辞書に加えるステップと、
発音及び語クラス情報を含む説明を前記新語と関連付けるステップと、
前記第１言語と関連付けられた第１機械翻訳モジュールにおいて、前記新語及び前記説明を更新するステップと、を含み、
前記第１機械翻訳モジュールは、第１タグ付けモジュール、第１翻訳モデル及び第１言語モジュールを含み、前記新語を前記第２言語において対応する翻訳語に翻訳するように構成されている方法。
【請求項２】
前記新語及び前記新語の説明を前記第１言語の前記認識辞書に加えるステップと、前記第１言語と関連付けられたクラスベース型言語モデルのクラス内において前記新語の相対的語確率を更新するステップとが、音声認識コンポーネントを再初期化又は再起動せずに実行される、請求項１又は２に記載の方法。
【請求項３】
前記新語及び前記新語の説明を前記第１言語の翻訳辞書に加えるステップと、前記新語の翻訳確率及び前記新語の翻訳を更新するステップとが、機械翻訳コンポーネントを再初期化又は再起動せずに実行される、請求項１又は２に記載の方法。
【請求項４】
前記第２言語から前記第１言語の前記新語に前記翻訳語を翻訳し戻し、前記新語を前記第２言語において対応する翻訳語と関連付け、前記翻訳語及び前記翻訳語の説明を前記第２言語の第２認識辞書に加えるステップと、
前記第２言語と関連付けられた第２機械翻訳モジュールであって、第２タグ付けモジュール、第２翻訳モデル及び第２言語モジュールを含む第２機械翻訳モジュールを、前記翻訳語及び前記説明で更新するステップとを更に含む、請求項１乃至３の何れかに記載の方法。
【請求項５】
前記第２言語と関連付けられたテキスト－音声発音辞書に前記翻訳語を入力するステップを更に含む、請求項１乃至４の何れかに記載の方法。
【請求項６】
前記第１言語と関連付けられたテキスト－音声発音辞書に前記新語を入力するステップを更に含む、請求項４又は５に記載の方法。
【請求項７】
前記システムは、前記第１言語の前記新語を、前記第２言語及び他の１つ以上の言語において対応する語に翻訳し、
関連付けるステップは、
前記新語を、他の１つ以上の言語において夫々対応する第３又はそれ以上の語と関連付けるステップと、他の１つ以上の言語における前記第３又はそれ以上の語を、前記他の１つ以上の言語の各々と関連付けられた認識辞書に加えるステップと、前記他の１つ以上の言語と関連付けられた機械翻訳モジュールであって、タグ付けモジュール、翻訳モデル及び言語モジュールを含む複数の機械翻訳モジュールを更新するステップとを含む、請求項１乃至６の何れかに記載の方法。
【請求項８】
前記新語は、前記システムによって生成された認識エラーを修正する過程において入力され、
前記修正された新語は、前記第１認識辞書、前記第１言語モデル及び前記第１翻訳モデルからなるモデル群のうちの１つには無い、請求項１乃至７の何れかに記載の方法。
【請求項９】

新語を加えるステップは、信頼基準及び新語モデルを用いて前記新語を推定するステップを含み、
前記新語は、前記第１認識辞書、前記第１言語モデル及び前記第１翻訳モデルからなるモデル群のうちの１つには無い、請求項１乃至７の何れかに記載の方法。
【請求項１０】
追加の１つ以上の新語が語リストとして入力され、
前記リストは、前記第１言語及び前記第２言語において前記新語及び追加の１つ以上の語をもたらし、前記リスト上の前記新語の各々の発音、前記説明、及び翻訳情報をもたらす、請求項８に記載の方法。
【請求項１１】
前記第１言語における新語リストは、外部情報、外部ソース又は外部イベントの発生によって選択される、請求項１０に記載の方法。
【請求項１２】
前記外部イベントは、前記システムを地理的位置決めに適用することで検出される位置の変化を含み、
前記新語は、前記位置の変化の結果としてもたらされる、請求項１１に記載の方法。
【請求項１３】
前記新語を前記第１言語の前記認識辞書に加えるステップは、前記第１言語と関連付けられたクラスベース型言語モデルのクラス内において、前記新語の相対的語確率を増加させるステップを更に含み、それにより、前記新語を引き続いてより確実に認識するように尤度を改善する、請求項１乃至１２の何れかに記載の方法。
【請求項１４】
前記翻訳語を前記第２認識辞書に加えるステップは、前記第２言語と関連付けられたクラスベース型言語モデルのクラス内において、前記新語の相対的語確率を増加させるステップを更に含み、それにより、前記新語を引き続いてより確実に認識するように尤度を改善する、請求項４乃至１２の何れかに記載の方法。
【請求項１５】
他の１つ以上の言語における前記第３又はそれ以上の語を前記他の１つ以上の言語のそれぞれと関連付けられた前記認識辞書に加えるステップは、前記第３又はそれ以上の言語と関連付けられたクラスベース型言語モデルのクラス内において、前記新語の相対的語確率を増加させるステップを更に含み、それにより、前記第３又はそれ以上の語を引き続いてより確実に認識するように尤度を改善する、請求項７乃至１２の何れかに記載の方法。
【請求項１６】
前記第１言語と関連付けられた前記新語の相対的語確率を増加させるステップは、前記新語を未知のクラスと関連付けることによって、且つ未知語のクラス内において前記新語の確率を増加させることによって、既知のクラス以外で実行される、請求項１３乃至１５の何れかに記載の方法。
【請求項１７】
前記新語を前記第１言語の前記翻訳辞書に加えるステップは、前記新語及び前記新語の翻訳における翻訳確率を増加させるステップを更に含み、ユーザにより適切に引き続いて前記新語を翻訳するように尤度を改善する、請求項１３に記載の方法。
【請求項１８】
説明を関連付けるステップは、ユーザによってもたらされる前記説明を受け取るステップを含む、請求項１乃至１７の何れかに記載の方法。
【請求項１９】
説明を関連付けるステップは、
辞書から、可能性がある１つ以上の説明を選択するステップと、
ユーザの承認用に前記可能性がある１つ以上の説明を表示するステップとを含み、
前記辞書は、ローカルに存在する、又はインターネットを介してアクセス可能である、請求項１乃至１８の何れかに記載の方法。

【請求項２０】
説明を関連付けるステップは、ユーザ現場カスタマイズモジュールを用いて仮説を自動的に生成するステップを含む、請求項１乃至１９の何れかに記載の方法。
【請求項２１】
前記生成は、手書きか又は既存の音声辞書から抽出されるかの何れかの変換ルールを用いて行われる、請求項２０に記載の方法。
【請求項２２】
前記生成は、字訳された音声データから学習される、請求項２０に記載の方法。
【請求項２３】
前記システムは、更に、前記新語用に最も高い可能性の語クラスを、既知のクラスを有する類似の他の語の共起統計に基づいて自動的に選択する、請求項２０に記載の方法。
【請求項２４】
説明を関連付けるステップは、インターネットを介してアクセス可能な人間言語のインフォーマントの外部ソースから、可能性がある１つ以上の説明を特定するステップを含む、請求項１乃至１７の何れかに記載の方法。
【請求項２５】
前記新語は、ユーザによってインターフェースを介して加えられ、
前記システムは、前記ユーザの言語における正書法の入力に基づき、前記第２言語の翻訳語と、前記説明と、言葉の予測発音とを仮定する、請求項１乃至２４の何れかに記載の方法。
【請求項２６】
前記第１言語における新語の説明内の発音表記及び他の言語において対応する語の説明内の発音表記は、音声再生によって前記インターフェースにおいて再検討及び確認される、請求項２５に記載の方法。
【請求項２７】
前記第１言語における新語の説明の発音表記は、前記ユーザによる例の口頭入力に基づいて選択される、請求項２５に記載の方法。
【請求項２８】
第１言語を第２言語に翻訳する機械翻訳の方法であって、
前記第１言語における文を第２言語における文に翻訳するステップを含み、
第１機械翻訳モジュールは、第１タグ付けモジュール、第１翻訳モデル及び第１言語モジュールを含み、前記文を前記第２言語において対応する翻訳文に翻訳するように構成されており、
語又は語句は、翻訳の間、１つのクラスとして取り扱われ、
クラスは、意味クラス、統語的クラス及び等価な語又は語句からなる群から選択されるクラスを含む方法。
【請求項２９】
請求項２８において適用された翻訳モデル及び言語モデルをトレーニングするために使用されるコーパス用に、クラスラベルを生成するステップを更に含み、
翻訳文ペア内の両方の文は、前記クラス－タグの集合が前記文ペアにわたって等価でなければならないという制約を適用している間、一緒にラベリングされる、請求項２８に記載の方法。
【請求項３０】
トレーニングコーパス内の複数の文ペア全部にわたって、ラベルが既知の第１言語から注釈が付けられていない言語にラベルを投影することによって、クラスラベルが生成される、請求項２９に記載の方法。
【請求項３１】
音声認識仮説に対するユーザ修正が、前記言語モデルの確率を更新するために使用され、前記修正された語系列における前記言語モデルの確率を増加させることによって、且つ似た競合仮説における前記言語モデルの確率を低減することによって、同一エラーが発生

する尤度を低減する、請求項２８乃至３０の何れかに記載の方法。
【請求項３２】
機械翻訳仮説に対するユーザ修正が、前記翻訳モデル及び言語モデルの確率を更新するために使用され、前記修正された語系列における前記翻訳モデル及び言語モデルの確率を増加させることによって、且つ似た競合仮説における前記翻訳モデル及び言語モデルの確率を低減することによって、同一エラーが発生する尤度を低減する、請求項２８乃至３１の何れかに記載の方法。
【請求項３３】
現場でメンテナンス可能なクラスベース型翻訳装置であって、
第１言語の辞書であって、前記第１言語における新語を含む音声を受け取り、発音及び語クラス情報を含む説明を前記新語と関連付けるように構成される認識辞書と、
前記第１言語と関連付けられた第１機械翻訳モジュールであって、タグ付けモジュール、翻訳モデル及び言語モジュールを含み、前記新語及び前記説明で更新されるように構成された第１機械翻訳モジュールとを含む装置。
【請求項３４】
前記第１機械翻訳モジュールは、前記新語を前記第２言語に対応する翻訳語に翻訳し、
出力は、第２テキストを表示するディスプレイを含む、請求項３３に記載の装置。
【請求項３５】
前記第２言語と関連付けられた第２機械翻訳モジュールであって、タグ付けモジュール、翻訳モデル及び言語モジュールを含み、前記第２言語における第２新語を第２翻訳語で翻訳する第２機械翻訳モジュールを更に含む、請求項３３又は３４に記載の装置。
【請求項３６】
前記第１言語と関連付けられた辞書であって、前記新語を受け取るように構成されたテキスト－音声発音辞書を更に含む、請求項３３乃至３５の何れかに記載の装置。
【請求項３７】
前記第２言語と関連付けられた辞書であって、前記翻訳語を受け取るように構成されたテキスト－音声発音辞書を更に含む、請求項３３乃至３６の何れかに記載の装置。
【請求項３８】
前記認識辞書は、前記第１言語における前記新語に対応する第１テキストを生成する、請求項３３乃至３７の何れかに記載の装置。
【請求項３９】
前記新語が前記口頭の第１言語に対応し、前記翻訳語が前記第２言語における前記第１テキストの翻訳に対応することを確認するモジュールであって、前記確認に基づいて低い又は高い信頼評価を生成する、相関及び修理モジュールを更に含み、
修正及び修理モジュールが、前記確認モジュールからの前記低い信頼評価によって特定されたエラーを認識し、前記エラーをユーザ挿入の修理テキストと置き換える、請求項３３乃至３８の何れかに記載の装置。
【請求項４０】
前記説明は、インターネットを介してアクセス可能な人間言語のインフォーマントの外部ソースから、可能性がある１つ以上の説明を特定することによって、前記新語と関連付けられる、請求項３３乃至３９の何れかに記載の装置。
【請求項４１】
ユーザ入力によって前記新語を受け取るように構成されるインターフェースを更に含み、
前記入力は、前記ユーザの言語における正書法の入力を含み、前記第２言語における翻訳語と、前記説明と、言葉の予測発音とを仮定する基準をもたらす、請求項３３乃至４０の何れかに記載の装置。
【請求項４２】
前記第１言語における前記新語の前記説明内の発音表記及び他の言語において対応する語の説明内の発音表記は、音声再生によりインターフェースにおいて再検討及び確認され

る、請求項３３乃至４１の何れかに記載の装置。
【請求項４３】
トレーニングコーパス内の複数の文ペア全部にわたって、ラベルが既知の第１言語から注釈が付けられていない言語にラベルを投影することによって、クラスラベルが生成される、請求項３３乃至４２の何れかに記載の装置。
【請求項４４】
前記第１言語モデルは、音声認識仮説に対して行われた修正に基づいて更新され、
前記更新は、前記言語モデルの確率を更新し、前記修正された語系列の前記言語モデルの確率を増加させることによって、且つ似た競合仮説の前記言語モデルの確率を低減することによって、同一エラーが発生する尤度を低減する、請求項３３乃至４３の何れかに記載の装置。
【請求項４５】
前記第１機械翻訳モジュールは、機械翻訳仮説に対する修正に基づいて更新され、
前記更新は、前記翻訳モデル及び前記言語モデルの確率を更新し、前記修正された語系列の前記翻訳モデル及び前記言語モデルの確率を増加させることによって、且つ似た競合仮説の前記翻訳モデル及び前記言語モデルの確率を低減することによって、同一エラーが発生する尤度を低減することを含む、請求項３３乃至４４の何れかに記載の装置。
【発明の詳細な説明】
【技術分野】
【０００１】
（関連出願の相互参照）
本願は、２００８年４月１５日に出願された米国仮特許出願Ｎｏ．６１／０４５，０７９、２００８年８月２８日に出願された米国仮出願Ｎｏ．６１／０９２，５８１、及び２００８年９月３日に出願された米国仮特許出願Ｎｏ．６１／０９３，８９８の優先権を主張する。
【０００２】
本発明は、一般に異言語間コミュニケーション用の音声－音声翻訳システムを対象とし、より詳細には、ユーザが、現場で、言語学若しくは技術的知識又は専門的知識を必要とせずに、新語彙（新たな語彙）項目を加え、ユーザのシステムの内容及び使用法を改善し修正することを可能にする、現場メンテナンス用の方法及び装置を対象とする。
【背景技術】
【０００３】
自動音声認識（ＡＳＲ）及び機械翻訳（ＭＴ）技術は十分発達してきており、限定された領域用及び非限定の領域用として、ラップトップ又はモバイルデバイス上に実用的な音声翻訳システムを開発することが実現可能となるまでになってきた。特に、領域限定の音声－音声変換システムは、研究分野において及び研究所において、観光事業、医療展開及び軍用を含むさまざまな適用領域向けに開発されてきた。このようなシステムは、以下に示す研究の中で以前から知られていた：A. Waibel, C. Fugen, "Spoken language translation" in Signal Processing Magazine, IEEE May 2008; 25(3):70-79, In Proc. HLT, 2003、及び Nguyen Bach, Matthias Eck, Paisarn Charoenpornsawat, Thilo Kohler, Sebastian Stuker, ThuyLinh Nguyen, Roger Hsiao, Alex Waibel, Stephan Vogel, Tanja Schultz and Alan W. Black, による例えば、"The CMU TransTac 2007 eyes-free and hands-free two-way speech-to-speech translation system," In Proc. of the IWSLT, Trento, Italy, October 2007。しかしながらこれらの研究は、前もってシステムの開発者によって定義されており、適用領域で決定された限定語彙と、システムが使用される想定される場所とで動作するように制限されている。それ故に語彙及び言語の使用法は、用例のシナリオに基づいて、且つこのようなシナリオにおいて収集又は推定されたデータによって、大きく決定される。
【０００４】
しかしながら、現場の状況では、実際の言葉及び言語の使用法は、研究所における予想

のシナリオから逸脱している。観光事業などの単純な領域であっても、ユーザが種々の場所へと移動し、種々の人々とやり取りし、種々の目標及びニーズを追求するにつれて、現場では言語の使用法は劇的に変化する。それ故に、新語（新たな語）及び新たな言い回しが常に生じるようになる。このような新語、つまり音声認識用語「未知語」（ＯＯＶ）は、語彙内語として誤認識され、そして、不正確に翻訳される。ユーザは言い換えを試みてもよいが、しかし（人名や都市名などの）決定的な語若しくは概念を入力する又はコミュニケーションできない場合には、語又は言い回しの不足により、コミュニケーションが断絶される場合もある。

【０００５】

ユーザが修正可能な音声－音声翻訳システムの必要性があるにもかかわらず、実際の解決法は今まで提案されてなかった。システムに語を加えることは容易に見えるかもしれないが、このような修正を行うことは極めて困難なことが分かっている。適切な修正は、システム全体を通して多くのコンポーネントモジュールに対して行わなければならず、大部分のモジュールは再トレーニングされて、コンポーネントのバランス及び統合化された機能を修復しなければならない。実際には、大略２０個の種々のモジュールが、新語を学習するために修正され又は再最適化されなければならないだろう。このような修正は、音声翻訳システムのコンポーネントに関する専門的知識及び経験を必要とする。その結果として、発明者の理解ではこのような修正は、これまでのところ専門家によって研究所内だけで行われてきており、人間の専門的知識、時間及びコストを必要としていた。

【０００６】

例えば、欧州のユーザ向けに設計されたシステムが、その語彙内に名称「Ｈｏｎｇ　Ｋｏｎｇ」を含んでいない場合を想定する。話し手が「Ｌｅｔ's　ｇｏ　ｔｏ　Ｈｏｎｇ　Ｋｏｎｇ」という文を話すと、システムは、辞書内における最も音声が似ている類似語を認識して、「Ｌｅｔ's　ｇｏ　ｔｏ　ｈｏｍｅ　ｃａｌｌ」と生成することになる。この時点では、エラーが、認識エラーの結果かどうか、又は音声－音声翻訳システム全体におけるこの語の欠落の結果かどうかは明らかでない。それ故に、ユーザは、次にシステムを修正し始める。これは、幾つかの修正技術のうちの１つによって行うことができる。最も単純には、再度話す又はタイプすることがあるが、代わりとして、他の開示及び従来技術（Ｗａｉｂｅｌ等による、米国特許第５，８５５，０００号明細書）に記載されるクロスモーダルによるエラー修正技術によって、より効果的に行える。所望の語系列による修正スペリングが確立される（「Ｌｅｔ's　ｇｏ　ｔｏ　Ｈｏｎｇ　Ｋｏｎｇ」）と、システムは翻訳を実行する。「Ｈｏｎｇ　Ｋｏｎｇ」が辞書内にある場合、システムはそこから正常に続行し、翻訳及び合成を実行することになる。しかしながら、認識及び翻訳辞書には無い場合、システムは、この語が固有表現か否かを確証する必要があるだろう。最後に、そして最も重要なことだが、ユーザが介在することによって名前又は語が学習されずに出力言語へと適切に翻訳できたとしても、ユーザが次回に同一の語を話す場合に、システムは再び故障することになる。

【発明の概要】

【発明が解決しようとする課題】

【０００７】

残念ながら、新語の学習は、語リスト内の新語をただ単にタイプ入力することだけでは対処することができず、大略２０個の種々の点において且つ音声翻訳システムの全てのレベルにおいて変更を必要とする。現在、新語の学習は、エントリについての手動によるタグ付け（ｔａｇｇｉｎｇ）及び編集、必要となる語を含む広範囲にわたるデータベースの収集、言語モデル及び翻訳モデルの確率の再トレーニング、並びに全体システムの再最適化を必要とし、それによって、全てのコンポーネントとコンポーネントの辞書との間の一貫性を再確立しようとし、且つシステム内の語、句及び概念間の統計的バランスを修復しようとする（確率は合計で１になる必要があり、それ故に全ての語は単一の語の追加の影響を受けることになる）。

【０００８】

結果として、現行の音声翻訳システムでは、一般に、小修正であっても、研究所内に見出される先端的な計算ツール及び言語学のリソースを使用する必要があった。しかしながら、実際の現場で使用するに際しては、研究所で行われるあらゆる修正を必要とすることは、多大な時間、努力及びコストがかかりすぎるので、受け入れ可能なことではない。代わりに、全ての複雑性をユーザから隠蔽し、全ての決定的な動作及び言語処理ステップを、裏で半自律的に又は自律的に実行すると共に、単純で直感的なインターフェースを介してできる限り混乱を起こさないやり方で人間ユーザとやり取りすることで、現場で言語学的又は技術的専門的知識の必要性を完全に取り除く学習及びカスタマイズモジュールが必要となる。本発明において、我々は、学習及びカスタマイズモジュールがこれらのニーズを満足するという詳細な説明を提供する。

【０００９】

残念ながら翻訳システムの多くは、非常に複雑であり、ユーザによるアクセスは実行可能でなく又は利用されない。それ故に、機械翻訳技術を利用すると共に、ユーザ修正機能が、言語学若しくは技術的知識又は専門的知識を必要とせずに異言語間コミュニケーションを提供することを可能にし、それによって、言語障壁を克服し、人々をより親密にさせることを可能にする、システム及び方法が必要である。

【課題を解決するための手段】

【００１０】

種々の実施形態では、本発明は、音声翻訳システムの語彙を更新する方法及び装置を提供することで、上述した課題を解決する。種々の実施形態では、書き言葉及び話し言葉を含む第１言語を第２言語に翻訳する音声翻訳システムの語彙を更新する方法が提供される。この方法は、第１言語における新語を第１言語の第１認識辞書（recognition lexicon）へと加えるステップと、発音及び語のクラス情報を含む説明を新語と関連付けるステップとを含む。次いで、これらの新語及び説明は、第１言語と関連付けられた第１機械翻訳モジュールにおいて更新される。この第１機械翻訳モジュールは、第１タグ付けモジュール、第１翻訳モデル及び第１言語モジュールを含み、新語を第２言語において対応する翻訳語へと翻訳するように構成される。

【００１１】

選択的に、双方向の翻訳用として、この方法は、更に加えて、翻訳語を第２言語から第１言語の新語に翻訳し返すステップと、新語を第２言語の対応する翻訳語と関連付けるステップと、翻訳語及びこの翻訳語の説明を第２言語の第２認識辞書へと加えるステップとを含む。次いで、第２言語と関連付けられた第２機械翻訳モジュールは、この翻訳語及び説明を用いて更新される。この第２機械翻訳モジュールは、第２タグ付けモジュール、第２翻訳モデル及び第２言語モジュールを含む。

【００１２】

複数の実施形態では、この方法は、第１言語と関連付けられたテキスト－音声発音辞書に第１語を入力し、第２言語と関連付けられたテキスト－音声発音辞書に第２語を入力するステップを更に含む。これらの入力信号は、異なるモダリティ（ｍｏｄａｌｉｔｙ）（例えば音声によるスペリング及び言葉によらないスペリング、音声によるスペリング及び言葉によるスペリング、書面及び音声など）（本明細書では「クロスモーダル」と称する）であってもよく、又は同一のモダリティ（音声及びやり直し音声、書面及び書き直し書面など）であってもよい。

【００１３】

本発明のある実施形態は、現場でメンテナンス可能なクラスベース型システムであって、第１言語と第２言語との間でコミュニケーションする音声－音声翻訳システムを対象とする。このシステムは、２つの音声認識ユニットであって、各音声認識ユニットは第１言語又は第２言語の口頭語を含む音声を受け取り、口頭の言語に対応するテキストを生成するように構成される音声認識ユニットと、２つの対応する機械翻訳ユニットであって、各機械翻訳ユニットは音声認識ユニットのうちの一方からテキストを受け、他方の言語のテキストへとテキストの翻訳を出力するように構成される機械翻訳ユニットとを含む。この

システムは、システムがユーザと連携して新語を学習可能にするユーザ現場カスタマイズモジュールも含む。このユーザ現場カスタマイズモジュールは、言語のうちの一方若しくは両方に対応する音声又はテキストを含むユーザ選択入力を受け取るように構成され、機械翻訳ユニットをこのユーザ選択入力で適切に更新する。

【００１４】

ある実施形態では、現場でメンテナンス可能なクラスベース型音声－音声翻訳システムを提供するために、４つの主要な特徴がシステムに備えられる。第１の特徴は、アクティブなシステム語彙に新語を加えることを可能にし、又は場所若しくはタスク固有の複数の語彙間で切り替えることを可能にする音声翻訳フレームワークを含む。この音声翻訳フレームワークは、音声認識モジュールに語を、このモジュールの再起動を必要とせずに動的に加えることをもたらす。このシステムは、音声－音声翻訳デバイス内の全てのシステムコンポーネントにわたる多言語システム辞書及び言語独立の語－クラスと、クラスベース型機械翻訳（句ベース型統計的ＭＴ、統語的ＭＴ、実例ベース型ＭＴなど）と、モデルトレーニングの間に単一言語のタガー（tagger）による組み合わせに基づいてなされる多言語の語－クラスのタグ付けと、既知のタグ付き言語からパラレルコーパスを介してアラインメントを行うために新たな言語でなされる語－クラスのタグ付けとを使用する。第２の特徴は、マルチモーダルなインタラクティブインターフェースにより、非専門家が新語をシステムへと加えることが可能となることである。第３の特徴は、このシステムが、ユーザによって提供されたマルチモーダルなフィードバックを用いて、ＡＳＲ及びＳＭＴモデル適応を適応させるように設計されていることである。そして第４の特徴は、このシステムが、修正又は語の共有を可能にするネットワーク機能を有することである。

【００１５】

別の実施形態では、現場で且つ技術的専門知識なしに、ユーザが音声－音声翻訳デバイスに新語を加えることを可能にするマルチモーダルなインタラクティブインターフェースが開示される。これには次のような例が含まれる：（１）システムへと加えられる語又は語句のクラスを自動的に分類し、語の発音及び翻訳を自動的に生成する方法、（２）話すこと、タイピング、スペリング、手書き、ブラウジング、言い換えのうちの１つ以上によって、クロスモーダル的に新語を入力する方法、（３）言語的に未トレーニングのユーザが、音声の字訳及び翻訳が適正かどうかを決定することを支援するマルチモーダルなフィードバックであって、多数のテキスト形式（即ちローマ字形式及び他の言語の文字による書面形式）と、テキスト－音声（ＴＴＳ、即ち適切に聞こえるかどうか）経由の音声形式とを含むフィードバック、（４）新語用に言語モデル及び翻訳の確率を設定する方法、並びに（５）ユーザの行動、関心事及び使用履歴との関連性に基づいて学習した新語に対して、言語モデル及び翻訳の確率を引き上げる又は割り引くこと。

【００１６】

別の実施形態では、現場においてユーザによるマルチモーダルなフィードバックを介して修正するオンラインシステムが開示される。これには次のような例が含まれる：（１）ユーザが自動音声認識結果を修正し、このフィードバック情報を使用して音声認識コンポーネントを適応させることを可能にするインターフェース及び方法、（２）ユーザが機械翻訳の仮説を修正し、このフィードバック情報を使用して機械翻訳コンポーネントを改善することを可能にするインターフェース及び方法、並びに（３）ユーザ修正に基づいて正しい語又は修正された語に対する言語モデル、辞書及び翻訳モデルの確率を自動的に調整する（高める又は減少させる）方法。

【００１７】

別の実施形態では、現場で行われる修正又は新語の追加を、複数のデバイスにわたってユーザが共有することを可能にするインターネット応用が開示される。これらには次のような例が含まれる：（１）ワールドワイドウェブを介して音声－音声翻訳デバイスに用いるモデルをアップロードする、ダウンロードする及び編集する方法、（２）ユーザのコミュニティ全体にわたる新語の追加及び修正を現場で照合する方法、及び（３）音声－音声翻訳デバイスに用いる場所又はタスク固有の語彙をアップロードする、ダウンロードする

及び編集する方法。
【図面の簡単な説明】
【００１８】
添付の図面は、本発明の実施形態における例を図示する。このような図面の簡単な説明は以下の通りである。
【００１９】
【図１】図１は、本発明の一実施形態に従って構成される音声－音声翻訳システムを図示するブロック図である。
【００２０】
【図２】図２は、タブレットインターフェースを介してユーザへと表示されるグラフィック・ユーザ・インターフェースの一例を図示する。
【００２１】
【図３】図３は、図１における本発明の一実施形態に従って実行される音声－音声翻訳のステップを図示するフローチャートである。
【００２２】
【図４】図４は、ユーザによって行われた修正からシステムが学習するステップを図示するフローチャートである（修正及び修理モジュール）。
【００２３】
【図５】図５は、ユーザが新語をシステムへと加えることができるステップを図示するフローチャートである（ユーザ現場カスタマイズモジュール）。
【００２４】
【図６】図６は、ユーザがシステムに加えることを望む新語について、装置が翻訳及び発音を自動的に生成する方法の一例を図示するフローチャートである。
【００２５】
【図７】図７は、マルチモーダルなインターフェースを介して入力する新語を検証する方法の一例を図示するフローチャートである。
【００２６】
【図８】図８は、自動的に生成される語情報を表示するビジュアルインターフェースの一例を図示する。
【００２７】
【図９】図９は、クラスベース型ＭＴモデルをトレーニングするために必要なステップを図示するフローチャートである。
【００２８】
【図１０】図１０は、クラスベース型ＭＴを入力文へと適用するステップを図示するフローチャートである。
【００２９】
【図１１】図１１は、統計的又は機械学習アプローチによって語－クラスのタグ付けの間に使用される可能性がある特徴を図示する図である。
【発明を実施するための形態】
【００３０】
本発明の種々の実施形態は、音声－音声翻訳用の方法及びシステムを説明する。実施形態は、モデル適応によって、ユーザの音声及び話し方に適応させるために使用されてもよい。さらなる実施形態において、ユーザは認識エラーを修正でき、システムは、ユーザが修正したエラーから明確｛めいかく｝に学習でき、それによってこれらのエラーが将来再び発生する可能性を小さくする。新語をシステムに加えるか又は固有の場所若しくはタスク用に最適化された事前定義型の辞書を選択するかの何れかによって、本発明は、ユーザが語彙を彼又は彼女の個々のニーズ及び環境にカスタマイズ可能にする。新語を加える場合に、マルチモーダルなインターフェースは、自動的に生成された翻訳及び発音をユーザが修正し検証することを可能にする。これにより、ユーザが他の言語についての知識を少しも有しないとき、ユーザが新語をシステムへと加えることが可能となる。ある実施形態

では、システムは、更に、ユーザによって入力される任意の新語彙をユーザのコミュニティへと送るように構成される。このデータは照合され、辞書は自動的に生成され、次いで任意のユーザによってダウンロードできる。

【００３１】

図１は、本発明に従っており、現場でメンテナンス可能な音声－音声翻訳システムの一例のブロック図の概要を図示する。この例は、システムは２つの言語L_a及びL_bの間で動作する。これは、音声－音声変換ダイアログシステムの典型的な実装例であって、ＬａからＬｂへ及びＬｂからＬａへの両方向の音声－音声翻訳を含む。しかしながら、この構成における双方向性は、本発明にとって必要条件ではない。ＬａからＬｂへの一方向システム、又は幾つかの言語L_1．．．L_nを含む多方向システムが、本発明から等しく利益を得られるだろう。システムは、２つのＡＳＲモジュール２及び９を有し、ＡＳＲモジュール２及び９は、音声モデル１８、ＡＳＲクラスベース型言語モデル１９及び認識辞書モデル２０（図３に示される）を用いて、L_a及びL_bの音声を夫々認識し、L_a及びL_bに対応するテキストを夫々生成する。この例では、我々は、モバイルテクノロジーズＬＬＣ（Mobile Technologies, LLC）で開発された音声認識システム「Ｎｉｎｊａ」を使用した。使用されてもよい他のタイプのＡＳＲモジュールは、ＩＢＭコーポレーション（IBM Corporation）、ＳＲＩ、ＢＢＮによって、又はケンブリッジ（Cambridge）若しくはアーヘン（Aachen）において開発された音声認識装置を含む。

【００３２】

システムは、２つの機械翻訳モジュール３及び８も含み、機械翻訳モジュール３及び８は、ＬａからＬｂへ及びＬｂからＬａへテキストを夫々翻訳する（モジュール１１）。この例で使用されるＭＴモジュールは、モバイルテクノロジーズＬＬＣで開発された「ＰａｎＤｏＲＡ」システムであった。他のＭＴモジュールでは、ＩＢＭコーポレーション、ＳＲＩ、ＢＢＮなどによって、又はアーヘン大学などにおいて開発されたモジュールが使用可能である。

【００３３】

２つのテキスト－音声変換エンジン４及び７は、各々が機械翻訳モジュール３及び８のうちの一方に対応し、対応するＡＳＲユニットから生成されたテキストを受け取るように構成される。出力テキストは、対応するＭＴモジュール３又は８へと転送され、ＭＴモジュール３及び８はテキストをＬａからＬｂへ及びテキストをＬｂからＬａへ夫々翻訳する。ＴＴＳモジュールは音声出力を生成し、Ｌａの少なくとも１つのテキスト語を、スピーカーなどの出力デバイス５を介して音声へと、及びＬｂの少なくとも１つのテキスト語を、デバイス５又はスピーカー６などの別の出力デバイスを介して音声へと夫々変換する。この例では、ケプストラル（Cepstral）ＴＴＳモジュールが使用された。ウィンドウズ（登録商標）ＳＡＰＩ（speech application programming interface）規約をサポートする任意のＴＴＳモジュールを、同様に用いることができる。

【００３４】

修正及び修理モジュール１１は、音声、ジェスチャー、書面、触角、タッチセンサー式及びキーボードインターフェースを含む多数のモダリティを介して、ユーザがシステムの出力を修正することを可能にし、システムがユーザの修正から学習することを可能にする。修正及び修理モジュールは、米国特許第５，８５５，０００号明細書において開示されるようなタイプであってもよい。ユーザ現場カスタマイズモジュール１２は、ユーザが新語彙をシステムへと加えるインターフェースを設けており、現在の状況に適切なシステム語彙を選択することもできる。例えば、デバイスの現在の場所を示すＧＰＳ座標によって決定されるような、場所の変更によって始動される、又はユーザによるタスク若しくは場所の明示の選択によって始動されるインターフェースがある。

【００３５】

ユーザは、ユーザ現場カスタマイズモジュール１２にアクセスし、デバイス１３のスクリーン（又はアクティブなタッチスクリーン）上に表示されるグラフィック・ユーザ・インターフェース・ディスプレイ、及びマウス又はペンを含むポインティングデバイス１４

を介して、システムとやり取りできる。グラフィック・ユーザ・インターフェースの一例が図２に示される。この例では、デバイス１３は、L_aの音声入力のテキスト及び対応するテキストをウィンドウ１５に表示する。テキストL_aの第２言語L_bにおける機械翻訳は、ウィンドウ１６に表示される。

【００３６】

ある実施形態では、両方の言語で同一のマイクロホン及びスピーカーが利用可能である。それ故に、マイクロホン１及び１０は、単一の物理的デバイスであり得、スピーカー５及び６は単一の物理的デバイスであり得る。

【００３７】

本発明の方法における一例の動作を図示するフローチャートが、図３に示される。最初にステップ１５ｂにおいて、音声認識システムはユーザによって起動される。例えば、グラフィック・ユーザ・インターフェース（図２、項目１５ｂ）上のボタン又は外部の物理的ボタン（図示されていない）を選択できる。次いでユーザの音声（項目２５）は、ステップ２７内のＡＳＲモジュールのうちの１つによって、即ち、ユーザがL_aを話している場合はモジュール２によって、ユーザがL_bを話している場合にはモジュール９によって認識される。ＡＳＲモジュール２及び９は、３つのモデル、即ち音声モデル１８、ＡＳＲクラスベース型言語モデル１９及び認識辞書モデル２０を適用する。これらのモデルは言語固有のものであり、各ＡＳＲモジュールはそのＡＳＲモジュール独自のモデル集合を含む。ステップ２８において、ユーザの音声の結果としてもたらされるテキストが、デバイススクリーン１３上のＧＵＩを介して表示される。

【００３８】

次いで、入力言語に基づきＭＴモジュール３又は８を用いて、翻訳がなされる（ステップ２９）。ＭＴモジュール３及び８は、３つの主なモデル、即ち語クラスを特定するタグ付け又は構文解析［Ｃｏｌｌｉｎｓ０２］モデル（モデル２２）、クラスベース型翻訳モデル（モデル２３）、及びクラスベース型言語モデル（モデル２４）を適用する。タグ付けモデル２２は、以下の先行技術文献に記載されたタイプなどの、適切などんなタイプのタグ付け又は構文解析モデルであってもよい：J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data, "In Proceedings of 18th International Conference on Machine Learning, pages 282-289, 2001 ("Lafferty01") 又は Michael Collins, "Parameter estimation for statistical parsing models: Theory and practice of distribution-free methods" (2004) In Harry Bunt, John Carroll, and Giorgio Satta, editors, New Developments in Parsing Technology, Kluwer。機械翻訳の間に適用される他のモデルには、翻訳の間に語が並べ替えられるやり方を制約する歪みモデル、及び文長モデルがある。クラスベース型機械翻訳の詳細な説明が、以下に与えられる。その結果としてもたらされる翻訳が、ステップ３０に示されるようにデバイス１３上にＧＵＩを介して表示される。

【００３９】

翻訳出力が適正かどうかをユーザが決定するのを支援するために、自動的に生成された翻訳（図２、項目１６）は、ＭＴモジュール３又は８を介して、入力言語へと翻訳し返されて、例えば図２、項目１５ａにおいて図示されるように、元の入力の真下に括弧付きで表示される。ＡＳＲモデル２又は９及びＭＴモジュール３又は８によって決定されるように、音声認識及び翻訳の両方の信頼性が高い場合（ステップ３１）、ＴＴＳモジュール４又は７を経由し（ステップ３３）、スピーカー５又は６を介して、音声出力（項目２６）が生成される。そうでない場合、システムは、ＧＵＩ、音声及び／又は周到なフィードバックを介して、翻訳が間違っている可能性があることを示す。ステップ３３において使用される特定のＴＴＳモジュールは、出力言語に基づいて選択される。

【００４０】

その後、ユーザが生成された翻訳に不満である場合、ステップ２７から３３までの何れかにおける音声－音声翻訳処理の間に、又は処理が完了した後で、ユーザが介入してもよい。これにより、修正及び修理モジュール１１が呼び出される（ステップ３５）。修正及

び修理モジュール１１は、ユーザが行える如何なる修正を記録しログをとる。この修正は後でＡＳＲモジュール２及び９並びにＭＴモジュール３及び８を更新するために使用でき、このことはこの文書の更に下方で詳細に説明される。修正が新語彙項目を含む場合（ステップ３６）、ステップ１５ｃにおいてユーザが現場カスタマイズモードを入力し、新語をシステムに明確に加える場合、又は、ステップ１５ｄにおいて、Thomas Schaaf, "Detection of OOV words using generalized word models and a semantic class language model," in Proc. of Eurospeech, 2001に記載された方法などの、信頼基準又は新語モデルを用いて入力音声内で新語が自動的に検出される場合、ユーザ現場カスタマイズモジュール１２が呼び出される。このモジュール１２は、マルチモーダルなインターフェースをもたらし、ユーザが新語をアクティブなシステム語彙へと加えることを可能にする。ユーザによって新語又は新句（新たな句）が加えられると、ＡＳＲ、ＭＴ及びＴＴＳモデル（項目１７、２１及び３３ａ）は要求に応じて更新される。このモジュールの機能は、両方の言語について更に下方で説明される。

【００４１】

複数のクラスの共通集合（例えば人名、地名、及び組織名）が、両方の言語におけるＡＳＲ及びＭＴの両方において使用される。この共通集合は、語義スロット（semantic slots）のシステム全体にわたる集合をもたらし、これにより新語がシステムに加えられることが可能となる。これらのクラス内に発生する名前、特別用語及び言い回しは、種々のユーザの配置、場所、行動様式、習慣及びタスクに依存して大いに変わりやすい語であり、それ故に最もユーザによるカスタマイズが必要となる。

【００４２】

好ましい例では、使用される特定のクラスは、システムの適用領域に依存する。クラスには、固有表現用の意味クラス、即ち人名、地名及び組織名、又はタスク固有の名詞句、例えば食品名、病名、薬名、及びどんな事前定義型のクラスにも適応しない語又は句用に開かれているもう１つのクラスが含まれてもよい。類義語などの統語的クラス又は語等価クラスも同様に使用できる。適用領域の例には、旅行者、医療行為、平和維持及び同類のものが含まれるが、これらには限定されない。一例では、旅行者の適用領域において必要となるクラスには、人々、都市、食品及び同類のものの名前が含まれる。別の例では、医療の専門家が必要とする適用クラスには、疾患名、薬剤名、解剖学上の名前、及び同類のものが含まれる。別の例では、平和維持適用において必要となるクラスには、武器、車両、及び同類のものの名前が含まれる。現場でカスタマイズできる音声翻訳を可能にするために、システムは、ユーザ現場カスタマイズモジュール１２と組み合わせた修正及び修理モジュール１１の動作を通じて、エラーを修正するとともに後でこれらのエラーから学習することを可能にする。

【００４３】

[修正及び修理モジュール]

修正及び修理モジュール１１は、ユーザが音声－音声翻訳処理中に任意の時点で介入可能にする。ユーザは、エラーを特定してそのログをとるか、又は彼／彼女が望む場合、音声認識若しくは翻訳出力内のエラーを修正してもよい。このようなユーザの介入は、人と人のコミュニケーション過程において即時に修正することをもたらし、システムがユーザのニーズ及び関心事に順応するとともに誤りから学習する機会をもたらすので、かなり価値がある。このようなエラーをフィードバックする機能を図示するフロー図が、図４に示される。ユーザが発言の翻訳に不満である場合（即ちエラーの発生）、ユーザは現在の入力のログをとり得る（ステップ４０）。システムは、現在の発言の音声及び他の情報をログファイルへと保存する。これにより、後の時点でユーザがアクセスして修正でき、又はユーザがコミュニティのデータベースへとアップロードすることができ、専門家のユーザがエラーを特定し修正可能にする。

【００４４】

ユーザは、相当数のモダリティを介して、音声認識又は機械翻訳出力を修正することもできる。ユーザは、再度話すことによって、又はキーボード若しくは手書きインターフェ

ースを介して文を入力して、発言全体を修正できる。代わりの構成としては、ユーザは、タッチスクリーン、マウス又はカーソルキーを介してアウトプット仮説の誤りのある部分をハイライトすることができ、キーボード、手書き、若しくは音声を用いて、又は語を一字一字はっきりとつづることで、その句又は語だけを修正する。ユーザは、タッチスクリーンを介してアウトプット仮説内の誤りのある部分を選択し、自動的に生成されるドロップダウンリスト内の競合仮説を選択することによって、若しくは音声で再度入力することによって、又は他の任意の補完的モダリティ（例えば、手書き、スペリング、言い換えなど）によっても修正できる。これらの方法、及び補完的修理行為を適切に組み合わせるやり方は、マルチモーダルな音声認識修正及び修理について米国特許第５，８５５，０００号明細書においてＷａｉｂｅｌ等によって提案された方法に基礎を置く。ここで、この提案された方法はインタラクティブな音声翻訳システムの音声認識及び翻訳モジュールへと適用される。

【００４５】

ユーザが音声認識出力を修正する場合（ステップ４３）、システムは最初にこの修正が新語を含むか否かを決定する（ステップ４４）。言語L_a及びL_bと夫々関連付けられた認識辞書モデル２０の語を照合して、この決定が行われる。語が見つけられない場合、システムは、ユーザが、所望ならば新語をアクティブなシステム語彙へと加えるように促す（図５、ステップ５０）。そうでない場合、ＡＳＲモデルにおける確率（図３、項目１７）は、再び同一エラーが発生する尤度を低減するように更新される。これは、修正された語系列の確率が増加し、似ている競合仮説の確率が低減される特徴的なやり方で実行できる。

【００４６】

ユーザは、言語の専門的知識を十分に有する場合、機械翻訳出力を修正できる。ＡＳＲの場合に使用されるのと同一のモダリティが使用可能である。機械翻訳出力がユーザによって修正される場合（ステップ４５）、且つ修正が新語を含む場合、ユーザは、新語をアクティブなシステム語彙へと加えるようにダイアログに促される（図５、ステップ５０）。修正が、アクティブなシステム語彙内にすでにある語を含むだけの場合、機械翻訳モデル（図３、項目２１）は更新される。具体的には、修正された文ペアから句が抽出されて、翻訳モデルに納められるように実施され得る。使用される目的言語モデルは、ＡＳＲの場合と類似のやり方で更新できる。

【００４７】

[ユーザ現場カスタマイズモジュール]

ユーザ現場カスタマイズモジュール１２は、システムがユーザと連携して新語を学習可能にする。従来のシステムは、ユーザが音声－音声翻訳システム内の語彙を修正するのを可能にしていない。従来のシステムとは違って、ユーザ現場カスタマイズモデル１２は、ユーザが実行システムにおいて逐次の修正を可能にし、この逐次の修正は、コンピューター音声及び言語処理技術の知識又は言語学の知識が最小限しかない又は少しもないような非専門家が実行するのに比較的容易である。モデル１２は、このような現場カスタマイズを提供するが、それは、ユーザからのある程度理解しやすいフィードバックをもたらすとともに受け取ることによって、並びにこのフィードバックに基づいて必要な全パラメータ及びシステム構成を自律的に導き出すことによってなされる。現場カスタマイズモジュール１２は、以下のことを通じてこれを成し遂げる：１）ユーザ－カスタマイズ用の直感的なインターフェース、及び２）ユーザカスタマイズのために必要となる内部の全パラメータ及び全設定を自動的に見積もり、それによってユーザの負担を軽減する内部ツール。

【００４８】

一方向の翻訳用に、システムは、語又は句に関する最低で４つの情報を処理し、新語又は新句をアクティブなシステム語彙へと加える。これら情報は以下を含む：

－　クラス（即ち新たなエントリの意味的又は統語的クラス）

－　言語L_aにおける語（即ちL_a内の書面形式）

－　L_aにおける語の発音

－　L_bにおける語の翻訳（即ちL_bにおける書面形式）
双方向の翻訳用に、システムは、L_bにおける新語の発音の入力も必要とする。L_bは、逆に、ＴＴＳが音声出力を生成し、L_b用のＡＳＲモジュールが新語を認識することを可能にする。

【００４９】
ユーザ現場カスタマイズモデル１２の動作ステップを図示するフローチャートが、例えば図５において示される。システムが新語に直面すると、前節に記載の修正及び修理モデル１１を介してなされる修正介入に基づいて、システムは、ユーザ（図５、ステップ５０）が、この語を「学習する」べきかどうか、即ちアクティブなシステム語彙へと加えるべきかどうかを決定するように促すことになる。その場合、語学習モードが起動され、現場カスタマイズモジュール１２が作動し始める。ここで、現場カスタマイズ又は新語の学習が、エラー修正ダイアログの結果だけを必要としているのではないことに留意すべきである。ユーザは、具体的にはプルダウンメニューから語学習モードに入り、新語又は先験的な新語のリストを加えることを選択してもよい。外部のイベントによって専門用語、名前、場所などの種々の語が急に必要となっても、新語の学習は同様に始動することができる。しかしながら全てのこのような場合において、システムは上述した情報を収集しなければならない。

【００５０】
ユーザが新語をシステム語彙へと加えたい旨を示した後（ステップ５０）、システムは最初に大きな外部辞書を調べる。この辞書は、デバイス上にローカルに含まれるか、若しくはインターネットを介してアクセスできる辞書サービスであるか、又は両方の組み合わせであるかの何れかである。外部辞書は、語翻訳ペアのエントリで構成される。各エントリは発音及び語－クラス情報を含み、この発音及び語－クラス情報によって、アクティブなシステム語彙に新語を容易に加えられる。各エントリは、両方の言語における語ペアの各々の説明も含む。これにより、目的の言語の知識が無くても、ユーザが語の適切な翻訳を選択可能となる。新語が外部の辞書に含まれる場合（ステップ５１）、システムは、各々の説明を含む語の代替的な翻訳リストを表示する（ステップ５２）。ユーザが辞書から事前定義型の翻訳のうちの１つを選択する場合（ステップ５３）、ユーザは、辞書によってもたらされる発音及び他の情報を検証し（ステップ５３ａ）、もし必要なら編集できる。次いで新語は、アクティブなシステム語彙へと加えられる。

【００５１】
アクティブなシステム語彙に新語を加えるために、３つのステップが必要となる（ステップ５９、５９ａ、５９ｂ）。最初に、語及びその翻訳が、モジュール２及び９のＡＳＲ認識辞書へと加えられる（ステップ５９）。この語は、外部の辞書によって与えられる発音（単数又は複数）とともに、この認識辞書２０へと加えられる。ユーザがこの語を単に入力すると、この語の発生確率は、ＡＳＲクラスベース型言語モデル１９内の同一クラスの競合する要素よりも大きく設定される。これにより、ユーザが加えた語の可能性が高くなる。次に、語及びその翻訳がＭＴモデルへと加えられ（図３、項目２１）、システムが新たな語を両方の翻訳方向に翻訳可能となる。最後に、この語は、ＴＴＳ発音モデルに登録され（図３、モデル３３ａ）、システムが両方の言語でこの語を正確に発音可能となる。

【００５２】
ユーザによって入力された新語が外部の辞書に含まれていない場合、システムは、語をアクティブなシステム語彙へと登録するために必要となる情報を自動的に生成して、この情報をユーザと検証する。最初に、周囲の語の文脈が利用可能ならばそれを用い（ステップ５４）、タグ付けモデルを介して（図３、モデル２２）、新語のクラスが推定される。次に、ルールベース型モデル又は統計的モデルのうちの何れかを介して、新語の発音及び翻訳が自動的に生成される（ステップ５５）。次いで、その結果としてもたらされる情報が、マルチモーダルなインターフェースを介してユーザへと示される（ステップ５８）。システムは、ユーザが自動的に生成される翻訳若しくは発音を検証する（ステップ５８）

又は修正する（ステップ５７）ように促す。最後に、ユーザがこの情報を検証した後で、新語はアクティブなシステム語彙へと加えられる（ステップ５９、５９ａ、５９ｂ）。新語（具体的には、「語＋発音＋語クラス」）をＡＳＲ語彙へと動的に加える（５９）ために、認識辞書２０（それは典型的にはＡＳＲモジュール２又は９内にツリー構造として記憶される）は検索され、次いで新語を含むように更新される。これにより、新語が認識語彙へと動的に加えられることが可能となり、それ故に口頭で次の発言があれば直ちに認識することができる。ＡＳＲシステムにとっては、従来のシステムにあるような再初期化又は再起動は必要としない。

【００５３】

同様に、新語（具体的には「語＋翻訳＋語クラス」）はＭＴ翻訳モデルに付け足すことができ（５９ａ）、翻訳モデル２３（それはＭＴモジュール３及び／又は８内にハッシュ－マップとして記憶することができる）は検索され、新語及びその翻訳を含む新たな翻訳ペア（対訳）、並びに語クラスが付加される。これにより、ＭＴモジュール３及び／又は８に動的に新語が加えられることが可能となり、新語は続く発言において正確に翻訳されることになる。ＭＴシステムにとっては、従来の研究にあるような再初期化又は再起動は必要としない。

【００５４】

この情報をすべて自動的に推定することは、現場における非専門家のユーザがカスタマイズの作業を行えるようにするのに不可欠である。以下では、語に関するこの決定的な情報が如何に自動的に推定されるか、次いで、それが如何にユーザから得られるか又は直感的に検証できるかを詳細に説明する。

【００５５】

[新語の発音及び翻訳の生成]

音声－音声翻訳システムのユーザは、通常は、音声学、言語学、言語技術に関して、限定された知識を有する又は少しの知識もなく、しばしば他の言語における語及びその使用法について少しの知識さえもないので、システムへと加えることを望む各新語についての翻訳及び関連のある全情報（発音、正書法、語の使用法など）を、ユーザがもたらすと期待することはできない。それ故に、ユーザが新語を入力するとき、システムは語－クラスを推定し、両方の言語における語の翻訳及び発音情報を自動的に生成する。

【００５６】

アクティブなシステム語彙に新語を登録するために、この語の翻訳並びにこの語及びその翻訳の両方の発音が必要となる。この情報の生成は、例えば図６に示されるように、３つのステップ処理として実行できる。最初に、語の発音が生成される（ステップ６０）。語及びその発音の文字系列に基づいて、翻訳が生成される（ステップ６１）。次に、目的の言語における新語の発音が、先行するステップにおいて生成される情報を用いて生成される（ステップ６２）。日本語－英語現場メンテナンス可能Ｓ２Ｓ翻訳システムにおける種々の技術を用いてこの情報を生成する２つの例が、図６の左側に示される。英語の新語「Ｗｈｅｅｌｉｎｇ」（項目６１）をシステムへと加えるために、まず機械学習を介して英語の発音が生成される（ステップ６５）。機械学習は、以下の先行技術文献に記載されるような適切な任意の技術によって行われてもよい：Damper, R. I. (Ed.), Data-Driven Techniques in Speech Synthesis. Dordrecht, The Netherlands: Kluwer Academic Publishers (2001)。次に、この語の日本語における字訳が統計的機械字訳を介して自動的に生成され（ステップ６６）、次いで手動で定義されたルールを介して日本語の発音が生成される（ステップ６７）。字訳は、適切な任意の統計的機械字訳エンジンを用いて成し遂げられてもよい。例えば、以下の先行技術文献には例が記載されている：K. Knight and J. Graehl, Machine transliteration. Computational Linguistics 24 4 (1998), pp. 599-612、及び HLT/NAACL-2007に見られるBing Zhao, Nguyen Bach, Ian Lane, and Stephan Vogel, "A Log-linear Block Transliteration Model based on Bi-Stream HMMs"。次いでその結果としてもたらされる情報（項目６８）は、アクティブなシステム語彙に語を登録する前に、ユーザによって音響再生機を介して及び音声ストリングで検証される。

【００５７】
同様に、日本語の新語「ワカヤマ」（項目７０）をシステムへと加えるために、まず手動で定義されたルールを介して日本語の発音が生成される（ステップ７１）。次に、日本語におけるこの語の字訳がルールベース型字訳を介して自動的に生成され（ステップ７２）、次いで手動で定義されたルールを介して英語の発音が生成される（ステップ７３）。ルールベース型字訳は、以下の先行技術文献の方法を用いて実行されてもよい：Mansur Arbabi, Scott M. Fischthal, Vincent C. Cheng, and Elizabeth Bar, "Algorithms for Arabic name transliteration, "IBM Journal of research and Development, 38(2):183-193, 1994。次いでその結果としてもたらされる情報（項目７４）は、アクティブなシステム語彙に語を登録する前に、ユーザによって検証される。

【００５８】
ユーザは、生成された翻訳及び音声出力による発音を検証できる。代わりの構成としては、ユーザに更に適切に考慮されるなら、母語を与える（即ちユーザが英語の話し手の場合、中国語に対しては「Ｈａｎｙｕ　Ｐｉｎｙｉｎ」で、又は日本語に対しては「Ｒｏｍａｊｉ」で）記載形式が使用されてもよい。ユーザは、もし必要なら、翻訳及び／又は発音を編集してもよい。ユーザによって承認された時点で、語及び語の特徴が多言語システム辞書に加えられる。

【００５９】
システムはまた、インタラクティブなユーザ入力の助けを借りて、必要となる情報を自動的に生成することによって、辞書に加えられる各新語の翻訳の必要性を取り除く。ユーザインターフェースの一例が図３に示される。

【００６０】
[インタラクティブなユーザインターフェース]
その後、システムは、推定された言語学の情報を確認し検証するために、ユーザに相談する。これは、どんな特別な言語学の又は技術的な知識でも推定しないように、直感的なやり方で行われる。それ故に、適切なインターフェースが使用される。以下にて、新語の学習の間におけるユーザとのやり取りを説明する。

【００６１】
このインターフェースにおいて、ユーザは、メニューから「新語」モードを選択してもよく、又はユーザ修正によって新／未知語が生成された後で、新語学習モードを呼び出すことができる。現れるウィンドウ枠において、彼／彼女は、所望の新語、名前、特別用語、概念、言い回しをタイプできる。これはユーザの言語における正書法の入力に基づいている（これは英語とは異なる、例えば中国語、日本語、ロシア語などの文字集合であり得る）。次に、システムは、ローマ字の字訳と、言葉の予測発音（words predicted pronunciation）とを生成する。これは、手書きの若しくは既存の音声辞書から抽出された又は字訳された音声データから学習された変換ルールによって行われる。次に、ユーザは自動変換を検査し、ＴＴＳを介して生成された発音の音声を再生できる。ユーザは、これらの表現（何れかの言語における文字、ローマ字で書かれた字訳、発音表記、及び発音表記の音声）のうちの何れかを反復し修正してもよく、他の対応するエントリも同様に再生成される（それ故に一方の言語における修正された発音表記は、他方の言語における発音表記を修正してもよい）。

【００６２】
システムは、更に、新語が属する最も高い可能性の語クラスを、類似の文脈内の（既知のクラスを有する）他の複数の語の共起統計（co-occurrence statistics）に基づいて自動的に選択する。新語のウィンドウ枠は、このクラス特定の手動選択（及び／又は修正）も可能にするが、ユーザが、このように推定された任意のクラス判断を覆せるようにされている。

【００６３】
要約すると、ユーザから新語／句が与えられたとすると、システムは、次のように動作する。

－（ＡＳＲ及びＭＴコンポーネントによって使用される）エントリの意味クラスを自動的に分類する。
－（ＡＳＲ及び$Ｌ_1$用のＴＴＳで使用される）語の発音を自動的に生成する。
－（両方のＭＴコンポーネントで使用される）語の翻訳を自動的に生成する。
－（ＡＳＲ及び$Ｌ_2$用のＴＴＳで使用される）翻訳の発音を自動的に生成する。
－　要求に応じて自動的に生成されたデータをユーザが修正／編集可能にする。
－　自動的に生成された翻訳が適正であるかどうかをユーザが検証する他のモダリティを設ける（即ちＴＴＳを介して語の発音を聞く）。

【００６４】

システムにおいて事前に定義されたクラスの何れとも一致しない語をユーザが入力する場合、ユーザはこの語を「未知の」クラスへと割り当てることができる。ＡＳＲに対しては、「未知の」クラスは、トレーニングデータに見出され、且つ認識辞書内には見出されない語によって定義される。ＳＭＴに対しては、翻訳辞書内に見出されない二言語のエントリが、目的言語モデル内の未知のタグに設定される。

【００６５】

[クラス内確率及び関連性の引き上げ]

これらの入力方法のいずれも、言語学のトレーニングを必要とせず、新語が適切に表されているかどうかをユーザが判断する直感的なやり方をもたらす。次いでユーザは、この新語を「多言語システム－辞書」即ちユーザ個別の辞書へと加えることによって、新語のエントリを受け取ってもよい。全体のシステムは、標準の辞書とカスタマイズした辞書とをユーザのランタイム辞書へと統合する。

【００６６】

上述した５つのエントリに加えて、クラス内確率Ｐ（ｗ｜Ｃ）も定義される。このやり方では、システムが、同一のクラスに属する複数の語を識別可能となる。それ故に、ユーザのタスク、好み及び習慣により近接した語が、好ましくてより高いクラス内確率が割り当てられる。ユーザへの関連性に基づいて、クラス内確率をより高くするこの引き上げが決定され、この場合に、以下の点を観察して関連性が評価される：

－　新語エントリ及びその最新性
　＊　新語を入力することによってユーザが新語を望んでいたことを彼／彼女が示したので、入力された新語は、当然ながら近い将来使用される可能性が高く、それ故にクラス内確率が、代わりの既存のクラスエントリを超えて引き上げられる（増加する）
－　新語と、ユーザ行動、関心事及びタスクとの間の相関関係、以下を含む
　＊　都市名、ランドマーク、興味深い場所などの場所に対する距離
　＊　過去の使用履歴
　＊　共起統計（スシはボゴタよりもトウキョウとより関連している）
－　新語の一般的顕著性、以下を含む。
　＊　都市人口
　＊　メディアにおける最近の言及

【００６７】

このような観察及び関連性の統計は、ユーザの観察した場所、履歴若しくは行動に基づいて、及び／又は代わりにインターネットなどの広いバックグラウンドの言語資源においてシステムの新語の発生を観察することによって、収集される。このような統計は、データが豊富な言語において単一言語で収集され、翻訳辞書及び翻訳言語モデルへと適用されてもよい。

【００６８】

ユーザの新たな行動及びタスクが、時間をかけてこのような複数の語の類似性を少なくするように、及び／又は（種々の都市で出現するような）新たな情報が複数の語のサブクラスの関連性を少なくする場合、引き上げられた複数の語の関連性は、時間をかけて減衰する場合もある。

【００６９】

[クロスモーダルのエントリ]
選択的に、新語は、次のうちの１つによって入力される：
－　スピーキング：ユーザは新語を話す。発音及び字訳などの全ての情報は、音声入力に基づく以外、上述と同様に、新語モデル、翻訳モデル、バックグラウンド辞書によって推定される。システムは、言葉による対話を行い、クラスの同一性及び他の関連情報を選択してもよい。
－　スペリング：ユーザは新語を音声でつづる。この入力方法は、一般に話している間にわたって正確な字訳の尤度を改善する。スペリングは、話すやり方及び他の入力モダリティに対して補完的に使用されてもよい。
－　ハンドライティング：ユーザは新語を手書きで入力する。この入力方法は、一般に話している間にわたって正確な字訳の尤度を改善する。手書きは、スピーキング、スペリング、又は他の入力モダリティに対して補完的に使用されてもよい。
－　ブラウジング：新語は、インタラクティブなブラウジングによって選択されてもよい。ここでシステムは、ユーザの最近の使用法の履歴及び／又は最近、選択され入力された新語と同様な統計的プロファイルを含むテキストを求めてインターネットを検索することによって、関連した、関係のある新語を提案してもよい。
【００７０】
[インターネットを介する遠隔型新語学習及び共有型辞書開発]
前節において説明された方法の全ては、個々のユーザが音声翻訳システムを現場で彼／彼女のニーズ及びタスクにカスタマイズ可能にすることを目指している。しかしながら、このようなユーザカスタマイズの多くは、他のユーザにも同様に役立つ可能性がある。ある実施形態では、ユーザカスタマイズはコミュニティの大きなデータベースへとアップロードされ、名前、特別用語、又は言い回しは、利害関係者間で共有される。語彙エントリ、翻訳及びクラスタグは収集され、類似の利害関係があるコミュニティへと関連付けられる。引き続いてユーザは、これらの共有されたコミュニティ資源をダウンロードし、ユーザ独自のシステムへの資源として加えることができる。
【００７１】
代わりの構成としては、ユーザは、翻訳文を不完全な状態でアップロードするだけであり、手動翻訳をコミュニティに要求することを選択してもよい。このような不正確な若しくは不完全な翻訳前の語又は文、及びこれらの不足している若しくは不正確な翻訳に対して、他のユーザはボランティア（又は有料）基準でオンライン修正及び翻訳を提供することができる。その結果としてもたらされる修正及び翻訳は、更新され共有されるコミュニティ翻訳データベースへともう一度提出される。
【００７２】
[管理されていない適応]
修正、修理及び新語学習の後、最後に、我々は修正された仮説、及びそれ故に口頭文についての真の筆記録又は翻訳を得る。音声－音声翻訳デバイス又はシステムは、このようなグランドトルース（ground truth）がもたらされ、ＡＳＲモジュール（図１、モジュール２又は９）をこのデバイスの主要なユーザに更に適応させる、という事実を自動的に使用できる。このような適応が、デバイスの正確さ及び有用性を改善するように設計される。適応についての２つの固有の方法が実行される。第１は、ユーザの音声をより良好に認識するシステムの適応；音声モデル及び発音モデル適応、並びに第２は、言語モデル適応のためにユーザの音声スタイルへの適応。特定のユーザ用の適応データを記憶するために、プロファイルが使用され、現場で切り替えることができる。
【００７３】
[クラスベース型機械翻訳]
前節において、我々はエラー修理及び新語学習を説明した。これらのモジュールでは、基準はクラスベース型機械翻訳としていた。次に我々は、このようなクラスベース型機械翻訳の詳細な機能を説明する。
【００７４】

[アプローチ]
最先端の機械翻訳システムは、語－レベル上で翻訳を実行する。このことは、次の３つのドキュメントにおいて記載されたシステムを含む先行の翻訳システムから明らかである;(1) P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B. Cowan, W. Shen, C. Moran, R. Zens, C. Dyer, O. Bojar, A. Constantin, and E. Herbst, 'Moses: Open source toolkit for statistical machine translation', In Proc. ACL, 2007 ("[Koehn07"); (2) D. Chiang, A. Lopez, N. Madnani, C.Monz, P. Resnik and M. Subotin, "The Hiero machine translation system: extensions, evaluation, and analysis,", In Proc. Human Language Technology and Empirical Methods in Natural Language Processing, pp. 779-786, 2005 ("Chiang05")；及び (3) K. Yamada and K. Knight" A decoder for syntax-based statistical MT". In Proc. Association for Computational Linguistics, 2002 ("Yamada02")。アラインメントはワードツーワード（word-to-word）で実行される；翻訳例、即ち句ペアは語レベルで一致処理が行われる；及び語ベース型言語モデルが適用される。Ｃｈｉａｎｇ０５などに記載の階層的翻訳モジュール、及びＹａｍａｄａ０２などに記載の構文ベース型翻訳モデルは、中間の構造を導入することによってこれを拡張する。しかしながら、これらのアプローチは、依然として語の正確な一致処理を必要としている。各語が別々のエンティティとして処理されるので、これらのモデルは初見の語に対して一般化されていない。

【００７５】
クラスベース型機械翻訳の一実施形態は、次のようなクラスベース型統計的機械翻訳であり、外国語文$f^J_1 = f_1、f_2、．．．、f_J$は、最大尤度を有する仮説であって次式で与えられる仮説$\hat{}\ e^I_1$を検索することによって別の言語$e^I_1 = e_1、e_2、．．．、e_I$へと翻訳される：

$$\hat{e}^I_1 = \mathrm{argmax}\, P(e^I_1 | f^J_1) = \mathrm{argmax}\, P(f^J_1 | e^I_1) \cdot P(e^I_1)$$

クラスは、固有表現、統語的クラス又は等価な語又は語句で構成されるクラスなどの意味クラスであり得る。一例として我々は、固有表現クラスがシステムへと組み込まれる事例を説明する。

【００７６】
翻訳の間に適用される最も大きな２つの情報モデルは、目的言語モデル$P(e^I_1)$及び翻訳モデル$P(f^J_1 | e^I_1)$である。クラスベース型統計的機械翻訳フレームワークにおいて、$P(f^J_1 | e^I_1)$はクラスベース型翻訳モデル（図３、モデル２３）、及び$P(e^I_1)$はクラスベース型言語モデル（図３、モデル２４）である。

【００７７】
統計的機械翻訳フレームワーク用のクラスベース型モデルは、図１０に示されるやり方を用いてトレーニングできる。最初に、文ペアのコーパスをトレーニングすることが正規化され（ステップ１００）、コーパスにタグ付けするためにタグ付けモデル（図３、モデル２２）が使用される（ステップ１０１）。これを実行する１つのアプローチがＬａｆｆｅｒｔｙ０１に記載されている。このステップにおいて、トレーニングペアを形成するように組み合わせる文は、独立してタグ付けすることができ、一緒にタグ付けすることができ、又は一方の言語からのタグが他方の言語へと推定できる。トレーニングコーパス全体がタグ付けされた後、文ペア内の複数の語がアライメントされる（ステップ１０２）。アラインメントは、以下の先行技術文献などに記載される現在のアプローチを用いて成し遂げられる：Franz Josef Och, Christoph Tillmann, Hermann Ney: "Improved Alignment Models for Statistical Machine Translation" ;pp. 20-28; Proc. of the Joint Conf. of Empirical Methods in Natural Language Processing and Very Large Corpora; University of Maryland, College Park, MD, June 1999、及び Brown, Peter R, Stephen A. Delia Pietra, Vincent J. Delia Pietra, and R. L. Mercer. 1993." The mathematics of statistical machine translation: Parameter estimation, "Computational Linguistics, vol 19(2):263-311。このステップにおいて、タグ付きエンティティ（即ち「ニューヨーク」）内にある連語の句は、単一のトークンとして処理される。次に、Ｋｏｅｈ

ｎ０７などの方法を用いて句が抽出され（ステップ１０３）、クラスベース型翻訳モデルを生成する（図３、モデル２３）。タグ付きコーパスは、クラスベース型目的言語モデルをトレーニングするためにも使用される（図３、モデル２４）。トレーニングは、以下の先行技術文献などに記載のやり方を用いて成し遂げられてもよい:B. Suhm and W. Waibel, "Towards better language models 'for spontaneous speech' in Proc. ICSLP-1994, 1994 (Suhm94)（ステップ１０４）。

【００７８】

入力文を翻訳するために、図１１において図示される方法が適用される。最初に、入力文が正規化され（ステップ１０５）、トレーニングコーパスへと適用されたようなやり方を用いてタグ付けされる（ステップ１０６）。入力文は、単一言語タガーを用いてタグ付けされる（図３、モデル２２）。次に、入力文は、クラスベース型ＭＴモデルを用いて復号化される（図３、モデル２３及び２４）。クラスベース型統計的機械翻訳に対して、標準統計的機械翻訳において使用されるやり方と同じやり方を用いて、復号化が実行されるが、しかしながら、句ペアは、以下の例に示されるように、語ではなくクラス－レベルで一致され得る。

【００７９】

タグ付き入力文を以下のように仮定する：

the train to @PLACE.city[Wheeling] leaves at @TIME[4:30]

次の句は、一致され得る：

【数１】

train	#	□□
leaves	#	□□
@PLACE.city	#	@PLACE.city
@TIME	#	@TIME
at @TIME	#	@TIME□
the train to @PLACE.city	#	@PLACE.city□□ □ □□
leaves at @TIME	#	□ @TIME□ □□ □□□

クラス内の語又は句（即ち：@PLACE.city[Wheeling]、 @TIME[4:30]）は、数／時間の事例の場合、直接通過するか、翻訳が翻訳モデルから決定される。ユーザは、「ユーザ現場カスタマイズモジュール」を介して翻訳モデルに新語を加えることができる（図１、モジュール１２）。次いで、（図６の例で詳述しているように）ユーザがあらかじめ都市名「Ｗｈｅｅｌｉｎｇ」を加えた場合、翻訳モデルは次の句も含むことになる：

【数２】

@PLACE.city	#	Wheeling	#	□□□□□□

【００８０】

翻訳モデル確率Ｐ（f^{J}_{1}｜e^{I}_{1}）（図３、モデル２３）及びＭＴクラスベース型言語モデル確率Ｐ（e^{I}_{1}）（図３、モデル２４）を仮定すると、検索は、最大尤度Ｐ（f^{J}_{1}｜e^{I}_{1}）・Ｐ（e^{I}_{1}）を有する翻訳仮説を見つけるように実行される。

【００８１】

上述した入力文及び句を仮定すると、結果的にもたらされる翻訳は次のようになる：

【数３】

@PLACE.city{□□□□□□}□□□□□ □ @TIME{4:30}□□□□□

これは入力文の正確な翻訳である。

【００８２】

この例では、語「Ｗｈｅｅｌｉｎｇ」はトレーニングコーパスに現れなかったにもかか

わらず、「ユーザ現場カスタマイズモジュール」（図１、モジュール１２）を介してユーザが語を入力した後、システムは、語を正確に翻訳できる。更にその上、語－クラスが既知であるので（この例では「@ＰＬＡＣＥ．ｃｉｔｙ」）、システムは、周囲の複数の語に対してより良好な翻訳を選択でき、翻訳出力の複数の語を正確に順に並べることになる。

【００８３】

[多言語コーパスのパラレルタグ付け]

ある実施形態では、単一言語タガーを有するトレーニングコーパスの各側を独立にタグ付けし、次いで各文ペアから一貫性がないラベルを取り除くことで、ラベリングされたパラレルコーパスが得られる。このアプローチにおいて、各文ペア（Ｓａ，Ｓｂ）に対して、最大条件付きの確率Ｐ（Ｔａ，Ｓａ）及びＰ（Ｔｂ，Ｓｂ）を有するラベル－系列ペア（Ｔａ，Ｔｂ）が選択される。任意のクラス－タグが発生する総数が、Ｐ（Ｔａ，Ｓａ）とＰ（Ｔｂ，Ｓｂ）との間で異なる場合、そのクラス－タグはラベル－系列ペア（Ｔａ，Ｔｂ）から取り除かれる。Ｐ（Ｔａ，Ｓａ）及びＰ（Ｔｂ，Ｓｂ）を推定する１つの方法は、条件付き確率場ベース（conditional random field-based）のタグ付けモデルＬａｆｆｅｒｔｙ０１を適用することによる。単一言語タグ付けの間に使用される機能セット（feature set）の一例が、図１１に示される。

【００８４】

ある実施形態では、単一言語の特徴に加えて、語－アラインメント（図１１内のｗｂ，ｊ）から抽出される目標語を用いることによって、複数の文ペア全部にわたるラベリングの一貫性を、更に改善することができる。

【００８５】

別の実施形態では、クラス－タグの集合が等価でなければならないという制約を適用している間、翻訳ペア内の両方の文が一緒にラベリングされる。具体的には、文ペア（Ｓａ，Ｓｂ）に対して、我々は結合最大条件付き確率を最大化するラベル－系列ペア（Ｔａ，Ｔｂ）を検索する。

【数４】

$$\lambda_a P(Ta, Sa) \cdot \lambda_b P(Tb, Sb) \quad \text{ここで、} 1 \leqq i \leqq M \text{に対して} \quad Oi(Ta) = Oi(Tb)$$

Oi(Ta)　ラベル系列Ta内のクラスータグiの発生総数
（エンティティの数であり、語の総数ではない）
Ｍ　クラスの全数
λ_a、λ_b　倍率

単一言語モデルの性能が著しく異なる場合、λａ及びλｂを最適化して、二言語タグ付け性能を改善できる。

【００８６】

ある実施形態では、手動で注釈を付けられたコーパスが、特定の言語に対して入手できない場合、トレーニングコーパス内の複数の文ペア全部にわたって、ラベルが既知の第１言語から注釈が付けられていない言語へとラベルを投影すること（projecting）によって、ラベルを生成できる。これを実行する１つのアプローチが、以下の先行技術文献に記載される：D. Yarowsky, G. Ngai and R. Wicentowski, "Inducting Multilingual Text Analysis Tools via Robust Projection across Aligned Corpora," In Proc. HLT, pages 161-168, 2001 ("Yarowsky01")。

【００８７】

[クラスベース型機械翻訳のシステム及び評価の例]

実験的評価を通じて、詳細を説明したクラスベース型機械翻訳が、従来のアプローチと比較して翻訳性能を改善することを示す。更に、セクション２．２．２に記載のパラレルタグ付けアプローチを用いることにより、翻訳の正確さが更に改善されることを示す。

【００８８】

旅行者領域用に開発された日本語と英語の間の翻訳用システムが評価された。トレーニングデータ及びテストデータの説明は、表１に示される。

【表１】

	英語	日本語
パラレルトレーニングコーパス		
文ペアの数	400k	
トークンの数	3，257k	3，171k
平均の文長	8．7	8．5
手動でタグ付けされたトレーニングデータ（上述したデータの部分集合）		
トレーニング（文ペア無し）	12600	
ヘルドアウトテスト（文ペア無し）	1400	
テスト集合		
文ペアの数	600	
トークンの数	4393	4669
平均の文長	7．3	7．8
OOVレート	0．3%	0．5%

表１：トレーニングデータおよびテストデータ

【００８９】

効果的なクラスベース型ＳＭＴを実現するために、複数の文ペアにわたる正確な及び一貫性のあるタグ付けが極めて重要である。我々はタグ付け品質を改善する以下の２つのアプローチを研究した；第１は、語－アラインメントからの二言語の特徴の導入；及び第２は、文ペアの両方の側が一緒にタグ付けされる二言語タグ付け。パラレルトレーニングコーパスからの１４，０００文ペアが、表２に示される１６クラスラベルを用いて手動でタグ付けされた。

【表２】

クラス	クラスラベル
数	基数、序数、系列、文字
時間	時間、日付、日、月
人	姓名の名、姓
場所	都市、国、ランドマーク
組織	航空会社、ホテル、企業名

表２：評価システム内で使用されるクラス

【００９０】

手動でラベリングされたこの集合から、我々は、タグ付けの正確さを評価するヘルドアウト（held-out）データとして、１つ以上のタグを収容する１０％（１４００文ペア）を選択した。

【００９１】

最初に、ベースラインの単一言語ＣＲＦベースのタガーの性能が評価された。ヘルドアウト集合の各側は、言語依存性モデルを用いて独立にラベリングされた。次いで出力は、手動の基準と比較された。種々の測定基準に対するタグ付けの正確さを表３に示す。

【表３】

タグ付け法	英語			日本語			二言語			%正確にタグ付けされた文ペア
	P	R	F	P	R	F	P	R	F	
単一言語	0.95	0.89	0.92	0.94	0.88	0.91	0.88	0.80	0.84	80%
＋アラインメント特徴	0.97	0.85	0.91	0.98	0.93	0.95	0.95	0.82	0.88	82%
＋一貫性がないタグを取り除く	0.99	0.83	0.90	0.99	0.82	0.90	0.99	0.81	0.89	82%
二言語タグ付け	0.98	0.92	0.95	0.98	0.92	0.95	0.97	0.90	0.93	92%
＋アラインメント特徴	0.98	0.93	0.96	0.98	0.93	0.96	0.98	0.92	0.95	92%

表３:ヘルドアウトトレーニング集合における単一言語および二言語タグ付けの正確さ

【００９２】

二言語のタグ付けでは、エンティティがコーパスの両方の側に正確にラベリングされる場合、タグは正確であると考えられる。右側列は、両方の側が正確にタグ付けされた文ペアのパーセンテージを示す。Ｆの得点が個々の言語に対して０．９０を上回っているにもかかわらず、二言語タグ付けの正確さは０．８４に著しく低下し、８０％の文ペアだけが正確にタグ付けされたにすぎない。アラインメント特徴を単一言語タガーへと組み込むことにより、両方の言語に対する精度が改善され、日本語側に対する再現率が著しく改善されたが、しかしながら正確にタグ付けされた文ペアのパーセンテージは僅かだけ増加したにすぎない。複数の文ペア全部にわたって一貫性がないタグを取り除くことにより、精度が改善されたが、しかし正確にタグ付けされた文ペアの数は改善されなかった。

【００９３】

次に、二言語タグ付けの有効性が、上述したアプローチを用いて評価された。このアプローチのタグ付けの正確さ、及び語アラインメント特徴が組み込まれた場合が、表３の下側の２行に示される。単一言語の事例と比較すると、二言語タグ付けはタグ付けの正確さを著しく改善した。タグ付けの一貫性が改善された（二言語タグ付けに対するＦの得点は、０．８４から０．９５へと増加した）だけでなく、英語及び日本語の両方の側におけるタグ付けの正確さも改善された。語－アラインメント特徴を組み込むことにより、全ての測定に対するタグ付けの正確さにおいて更に小さい改善が得られた。

【００９４】

更に、３つのクラスベース型システムとクラスモデルを使用しないベースラインシステムとの性能を比較することによって、システムの有効性が評価された。

【００９５】

ベースラインシステムに対して、句－ベースの翻訳モデルが、Ｋｏｅｈｎ０５及びＧＩＺＡ＋＋（これは以下の先行技術文献などに使用されている）などに記載のＭｏｓｅｓツールキットを用いてトレーニングされた：Franz Josef Och, Hermann Ney. "A Systematic Comparison of Various Statistical Alignment Models", Computational Linguistics, volume 29, number 1, pp.19-51 March 2003。３－グラム言語モデルは、以下の先行技術文献に記載のＳＲＩＬＭツールキットを用いてトレーニングされた:A. Stolcke "SRILM - an extensible language modeling toolkit", In Proc. of ICSLP, pp. 901-904, 2002。復号化は、ＰａｎＤｏＲＡ復号化器を用いて実行された。この復号化器は、以下の先行技術文献に記載されている:Ying Zhang, Stephan Vogel, "PanDoRA: A Large-scale Two-way Statistical Machine Translation System for Hand-held Devices, "In the Proceedings of MT Summit XI, Copenhagen, Denmark, Sep. 10-14 2007。システムは、両方の翻訳方向Ｊ→Ｅ（日本語から英語へ）及びＥ→Ｊ（英語から日本語へ）に対して、表１

に記載のトレーニング集合を用いて作成された。目的言語モデルをトレーニングするために使用されるデータは、このコーパスに限定された。ベースラインシステムの翻訳品質は、６００文のテスト集合上で評価された。評価の間、１つの基準が使用された。Ｊ→Ｅ及びＥ→Ｊシステムに対するＢＬＥＵ得点は、それぞれ０．４３８１及び０．３９４７であった。ＢＬＥＵ得点は、以下の先行技術文献に記載されている:Kishore Papineni, Salim Roukos, Todd Ward and Wei-Jing Zhu "BLEU: a Method for Automatic Evaluation of Machine Translation," In Proc. Association for Computational Linguistics, pp. 311-318, 2002。異なる３つのタグ付け法を用いる翻訳品質が評価された：

＋ｎｕｍ　　：数、時間に関連した８クラス

＋ＮＥ－クラス　：上述クラス＋固有表現用の８クラス

＋Ｂｉ－タグ付け：上述の１６クラス、二言語でタグ付けされたトレーニングコーパス

【００９６】

単一言語タグ付けは、＋ｎｕｍ及び＋ＮＥ－クラスの事例に対して適用され、文ペアにわたって一貫性がないタグは取り除いた。＋Ｂｉ－タグ付けの事例では、語アラインメント特徴を組み込む二言語タグ付けが使用された。タグ付け法のそれぞれに対して、トレーニングコーパス全体が、クラス－ラベルの適切な集合でタグ付けされた。次いでクラスベース型翻訳モデル及び言語モデルは、ベースラインシステムにおいて使用されるのと等価なやり方を用いてトレーニングされた。テストの間、入力文は単一言語タガーを用いてタグ付けされた。テスト集合内の全ての固有表現は、翻訳の間に使用するために、ユーザ辞書へと入力された。

【００９７】

Ｊ→Ｅ及びＥ→ＪシステムのＢＬＥＵ得点について、ベースライン及びクラスベース型システムに対する６００文のテスト集合における性能を表４に示す。

【表４】

システム	翻訳品質(BLEU[Papineni02])	
	J→E	E→J
ベースライン	0.4381	0.3947
＋num	0.4441	0.4104
＋NE－クラス	0.5014	0.4464
＋Bi－タグ付け	0.5083	0.4542

表４:クラスベース型ＳＭＴの翻訳品質

【００９８】

数及び時間タグ（＋ｎｕｍ）を用いるクラスベース型ＳＭＴシステムが、ベースラインシステムと比較して、両方の翻訳方向に対して改善された翻訳品質を得た。これらのモデルについて、０．４４４１及び０．４１０４というＢＬＥＵ得点が得られた。数及び時間タグに加えて固有表現を用いるクラスベース型システムが適用される場合、翻訳品質は著しく改善された。Ｊ→Ｅシステムについて０．５０１４、及びＥ→Ｊの事例について０．４４６４というＢＬＥＵ得点が得られた。トレーニングコーパスをタグ付けするために二言語タグ付けが使用されたとき（＋Ｂｉ－タグ付け）、ＢＬＥＵにおいて両方の翻訳方向に対して更に０．８点の増加が得られた。１つ以上の固有表現を含むテスト集合内の１４％の文において、（＋Ｂｉ－タグ付け）システムは、単一言語でタグ付けされたシステム（「＋ＮＥ－クラス」）を、最大３．５ＢＬＥＵ点まで優れていた。

【００９９】

上述したように非常に詳細な内容が説明されたが、図面及び詳細な実施形態は、説明するために提供されたものであり、限定するためではないことが理解されるべきである。設計及び構成の変形が、本発明の原理内でなされる場合もある。当業者は、本発明のこのよ

うな変更若しくは修正若しくは要素の組み合わせ、変形、均等物、又は改善が、添付の請求項において定められるような本発明の範囲内に、依然として存在することを理解することになる。

【図１】

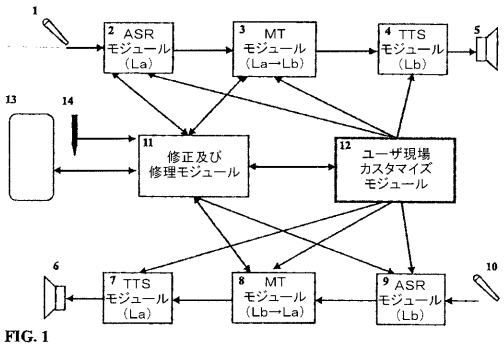


FIG. 1

【図２】



FIG. 2

【図３】

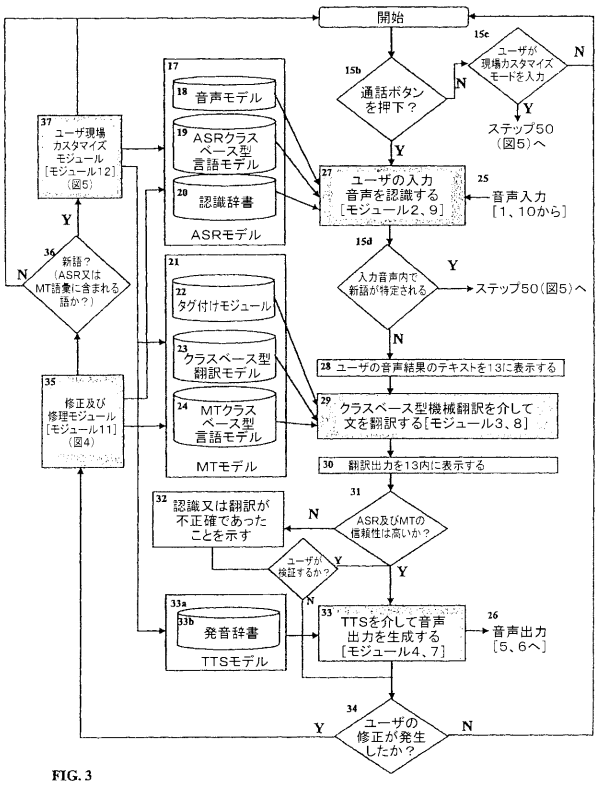


FIG. 3

【図４】

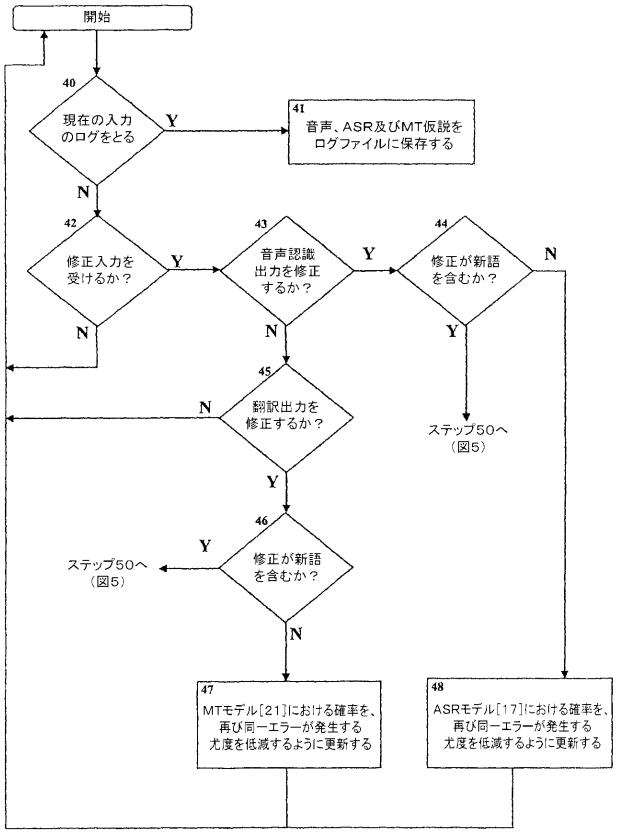


FIG. 4

【図５】

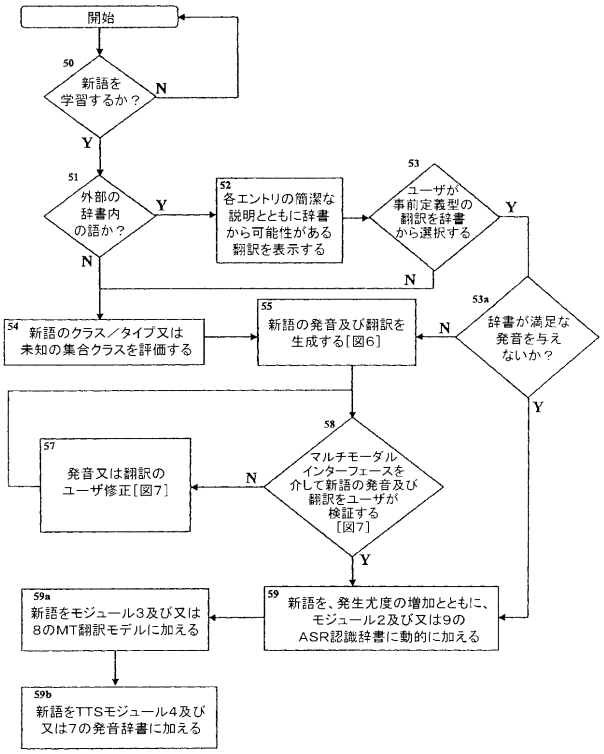


FIG. 5

【図６】

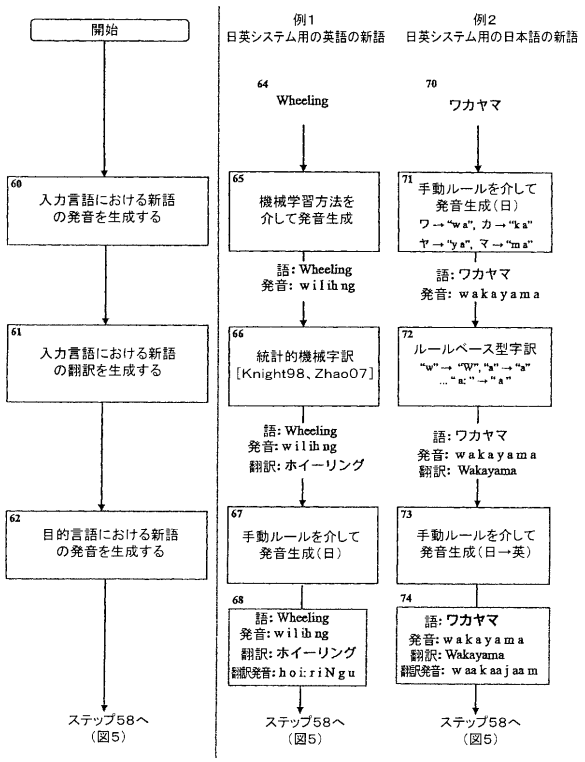


FIG. 6

【図７】

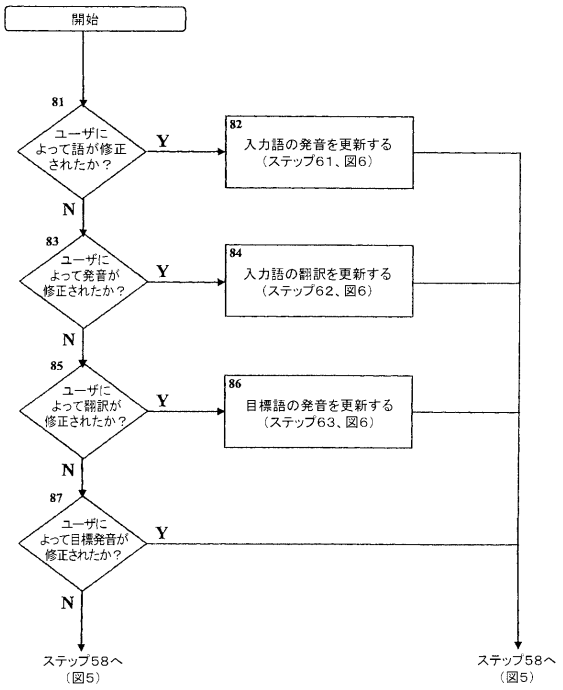
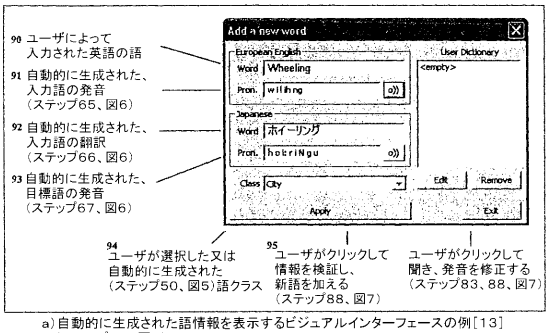
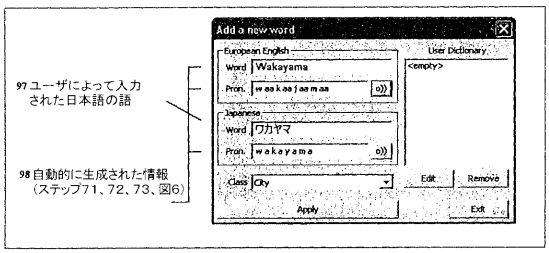


FIG. 7

【図８】



a）自動的に生成された語情報を表示するビジュアルインターフェースの例[13]
（ステップ80、図6）
ユーザは英語の語「Wheeling」（都市名）を加える



b）自動的に生成された語情報を表示するビジュアルインターフェースの例[13]
（ステップ80、図6）
ユーザは日本語の語「ワカヤマ」（都市名）を加える

FIG. 8

【図９】

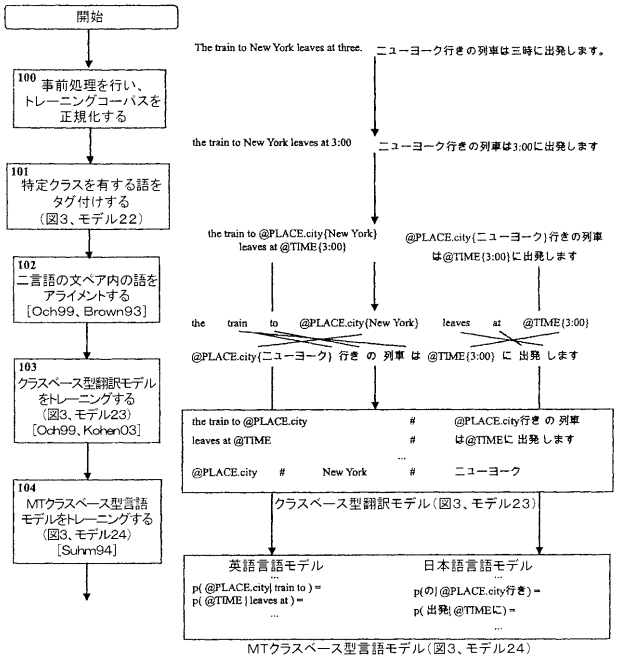


FIG. 9

【図１０】

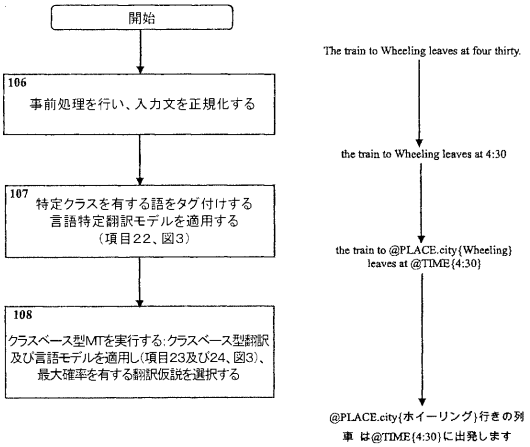
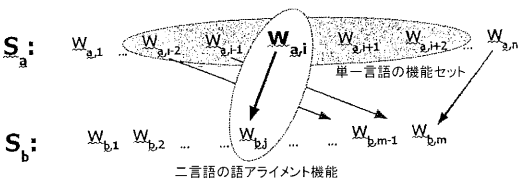


FIG. 10

【図１１】



語 $w_{a,i}$ 用の単一言語の機能セット $= \{ w_{a,i}, stem(w_{a,i}), w_{a,i-2}, w_{a,i-1}, w_{a,i+1}, w_{a,i+2} \}$

語 $w_{a,i}$ 用の二言語の機能セット $= \{ w_{a,i}, stem(w_{a,i}), w_{a,i-2}, w_{a,i-1}, w_{a,i+1}, w_{a,i+2}, w_{b,j} \}$

FIG. 11

【手続補正書】
【提出日】平成22年2月15日(2010.2.15)
【手続補正１】
【補正対象書類名】特許請求の範囲
【補正対象項目名】全文
【補正方法】変更
【補正の内容】
【特許請求の範囲】
【請求項１】
第１言語を書き言葉及び話し言葉を含む第２言語に翻訳する音声翻訳システムの語彙を更新する方法であって、
前記第１言語における新語を前記第１言語の第１自動音声認識辞書に加えるステップと、
発音、翻訳及び語クラス情報を含む説明であって、前記語クラスは、第１機械翻訳モジュールのタグ付けモデルを介して推定され、前記発音及び前記翻訳はルールベース型モデル又は統計的モデルを介して生成される説明を、前記新語と関連付けるステップと、
マルチモーダルなインターフェイスを介して前記説明を表示するステップと、
前記説明の検証を、又は、前記説明の修正及び検証をユーザに促すステップと、
前記第１言語と関連付けられた第１機械翻訳モジュールにおいて、前記新語及びユーザが検証した前記説明を更新するステップと、を含み、
前記第１機械翻訳モジュールは、第１翻訳モデル及び第１言語モジュールを含み、前記新語を前記第２言語において対応する翻訳語に翻訳するように構成されている方法。
【請求項２】
前記新語及び前記新語の説明を前記第１言語の前記認識辞書に加えるステップと、前記第１言語と関連付けられたクラスベース型言語モデルのクラス内において前記新語の相対的語確率を更新するステップとが、音声認識コンポーネントを再初期化又は再起動せずに実行される、請求項１又は２に記載の方法。
【請求項３】
前記新語及び前記新語の説明を前記第１言語の翻訳辞書に加えるステップと、前記新語の翻訳確率及び前記新語の翻訳を更新するステップとが、機械翻訳コンポーネントを再初期化又は再起動せずに実行される、請求項１又は２に記載の方法。
【請求項４】
前記第２言語から前記第１言語の前記新語に前記翻訳語を翻訳し戻し、前記新語を前記第２言語において対応する翻訳語と関連付け、前記翻訳語及び前記翻訳語の説明を前記第２言語の第２認識辞書に加えるステップと、
前記第２言語と関連付けられた第２機械翻訳モジュールであって、第２タグ付けモジュール、第２翻訳モデル及び第２言語モジュールを含む第２機械翻訳モジュールを、前記翻訳語及び前記説明で更新するステップとを更に含む、請求項１乃至３の何れかに記載の方法。
【請求項５】
前記第２言語と関連付けられたテキスト－音声発音辞書に前記翻訳語及びその発音を登録するステップを更に含む、請求項１乃至４の何れかに記載の方法。
【請求項６】
前記第１言語と関連付けられたテキスト－音声発音辞書に前記新語を入力するステップを更に含む、請求項４又は５に記載の方法。
【請求項７】
前記システムは、前記第１言語の前記新語を、前記第２言語及び他の１つ以上の言語において対応する語に翻訳し、
関連付けるステップは、
前記新語を、他の１つ以上の言語において夫々対応する第３又はそれ以上の語と関連付

けるステップと、他の１つ以上の言語における前記第３又はそれ以上の語を、前記他の１つ以上の言語の各々と関連付けられた認識辞書に加えるステップと、前記他の１つ以上の言語と関連付けられた機械翻訳モジュールであって、タグ付けモジュール、翻訳モデル及び言語モジュールを含む複数の機械翻訳モジュールを更新するステップとを含む、請求項１乃至６の何れかに記載の方法。

【請求項８】

前記新語は、前記システムによって生成された認識エラーを修正する過程において入力され、

前記修正された新語は、前記第１認識辞書、前記第１言語モデル及び前記第１翻訳モデルからなるモデル群のうちの１つには無い、請求項１乃至７の何れかに記載の方法。

【請求項９】

新語を加えるステップは、信頼基準及び新語モデルを用いて前記新語を推定するステップを含み、

前記新語は、前記第１認識辞書、前記第１言語モデル及び前記第１翻訳モデルからなるモデル群のうちの１つには無い、請求項１乃至７の何れかに記載の方法。

【請求項１０】

追加の１つ以上の新語が語リストとして入力され、

前記リストは、前記第１言語及び前記第２言語において前記新語及び追加の１つ以上の語をもたらし、前記リスト上の前記新語の各々の発音、前記説明、及び翻訳情報をもたらす、請求項８に記載の方法。

【請求項１１】

前記第１言語における新語リストは、外部情報、外部ソース又は外部イベントの発生によって選択される、請求項１０に記載の方法。

【請求項１２】

前記外部イベントは、前記システムを地理的位置決めに適用することで検出される位置の変化を含み、

前記新語は、前記位置の変化の結果としてもたらされる、請求項１１に記載の方法。

【請求項１３】

前記新語を前記第１言語の前記認識辞書に加えるステップは、前記第１言語と関連付けられたクラスベース型言語モデルのクラス内において、前記新語の相対的語確率を増加させるステップを更に含み、それにより、前記新語を引き続いてより確実に認識するように尤度を改善する、請求項１乃至１２の何れかに記載の方法。

【請求項１４】

前記翻訳語を前記第２認識辞書に加えるステップは、前記第２言語と関連付けられたクラスベース型言語モデルのクラス内において、前記新語の相対的語確率を増加させるステップを更に含み、それにより、前記新語を引き続いてより確実に認識するように尤度を改善する、請求項４乃至１２の何れかに記載の方法。

【請求項１５】

他の１つ以上の言語における前記第３又はそれ以上の語を前記他の１つ以上の言語のそれぞれと関連付けられた前記認識辞書に加えるステップは、前記第３又はそれ以上の言語と関連付けられたクラスベース型言語モデルのクラス内において、前記新語の相対的語確率を増加させるステップを更に含み、それにより、前記第３又はそれ以上の語を引き続いてより確実に認識するように尤度を改善する、請求項７乃至１２の何れかに記載の方法。

【請求項１６】

前記第１言語と関連付けられた前記新語の相対的語確率を増加させるステップは、前記新語を未知のクラスと関連付けることによって、且つ未知語のクラス内において前記新語の確率を増加させることによって、既知のクラス以外で実行される、請求項１３乃至１５の何れかに記載の方法。

【請求項１７】

前記新語を前記第１言語の前記翻訳辞書に加えるステップは、前記新語及び前記新語の

翻訳における翻訳確率を増加させるステップを更に含み、ユーザにより適切に引き続いて前記新語を翻訳するように尤度を改善する、請求項１３に記載の方法。
【請求項１８】
説明を関連付けるステップは、ユーザによってもたらされる前記説明を受け取るステップを含む、請求項１乃至１７の何れかに記載の方法。
【請求項１９】
説明を関連付けるステップは、
辞書から、可能性がある１つ以上の説明を選択するステップと、
ユーザの承認用に前記可能性がある１つ以上の説明を表示するステップとを含み、
前記辞書は、ローカルに存在する、又はインターネットを介してアクセス可能である、請求項１乃至１８の何れかに記載の方法。
【請求項２０】
説明を関連付けるステップは、ユーザ現場カスタマイズモジュールを用いて仮説を自動的に生成するステップを含む、請求項１乃至１９の何れかに記載の方法。
【請求項２１】
前記生成は、手書きか又は既存の音声辞書から抽出されるかの何れかの変換ルールを用いて行われる、請求項２０に記載の方法。
【請求項２２】
前記生成は、字訳された音声データから学習される、請求項２０に記載の方法。
【請求項２３】
前記システムは、更に、前記新語用に最も高い可能性の語クラスを、既知のクラスを有する類似の他の語の共起統計に基づいて自動的に選択する、請求項２０に記載の方法。
【請求項２４】
説明を関連付けるステップは、インターネットを介してアクセス可能な人間言語のインフォーマントの外部ソースから、可能性がある１つ以上の説明を特定するステップを含む、請求項１乃至１７の何れかに記載の方法。
【請求項２５】
前記新語は、ユーザによってインターフェースを介して加えられ、
前記システムは、前記ユーザの言語における正書法の入力に基づき、前記第２言語の翻訳語と、前記説明と、言葉の予測発音とを仮定する、請求項１乃至２４の何れかに記載の方法。
【請求項２６】
前記第１言語における新語の説明内の発音表記及び他の言語において対応する語の説明内の発音表記は、音声再生によって前記インターフェースにおいて再検討及び確認される、請求項２５に記載の方法。
【請求項２７】
前記第１言語における新語の説明の発音表記は、前記ユーザによる例の口頭入力に基づいて選択される、請求項２５に記載の方法。
【請求項２８】
第１言語における文を第２言語に翻訳する機械翻訳の方法であって、
第１機械翻訳モジュールは、第１機械翻訳モジュールは、第１タグ付けモジュール、第１クラスベース型翻訳モデル及び第１クラスベース型言語モジュールを含み、前記第１機械翻訳モジュールに現在登録されていない前記第１言語における新語を前記第１機械翻訳モジュールに加えるステップと、
前記新語を説明と関連付けるステップであって、前記説明は、語クラス情報と、前記第２言語における前記新語の翻訳とを含んでおり、前記語クラスは、前記第１タグ付けモデルを介して推定され、前記第２言語における前記翻訳はルールベース型モデル又は統計的モデルを介して仮定されるステップと、
マルチモーダルなインターフェイスを介して前記説明を表示するステップと、
前記説明の了承又は修正をユーザに促すステップと、

前記第１記載翻訳モジュールに前記ユーザが了承又は修正した前記説明を登録するステップと、
新語を含む前記第１言語における前記文を前記第２言語において対応する翻訳文に翻訳するステップとを含んでおり、
語又は語句は、翻訳の間、１つのクラスとして取り扱われ、
クラスは、意味クラス、統語的クラス及び等価な語又は語句からなる群から選択されるクラスを含む方法。
【請求項２９】
請求項２８において適用された翻訳モデル及び言語モデルをトレーニングするために使用されるコーパス用に、クラスラベルを生成するステップを更に含み、
翻訳文ペア内の両方の文は、前記クラス－タグの集合が前記文ペアにわたって等価でなければならないという制約を適用している間、一緒にラベリングされる、請求項２８に記載の方法。
【請求項３０】
トレーニングコーパス内の複数の文ペア全部にわたって、ラベルが既知の第１言語から注釈が付けられていない言語にラベルを投影することによって、クラスラベルが生成される、請求項２９に記載の方法。
【請求項３１】
音声認識仮説に対するユーザ修正が、前記言語モデルの確率を更新するために使用され、前記修正された語系列における前記言語モデルの確率を増加させることによって、且つ似た競合仮説における前記言語モデルの確率を低減することによって、同一エラーが発生する尤度を低減する、請求項２８乃至３０の何れかに記載の方法。
【請求項３２】
機械翻訳仮説に対するユーザ修正が、前記翻訳モデル及び言語モデルの確率を更新するために使用され、前記修正された語系列における前記翻訳モデル及び言語モデルの確率を増加させることによって、且つ似た競合仮説における前記翻訳モデル及び言語モデルの確率を低減することによって、同一エラーが発生する尤度を低減する、請求項２８乃至３１の何れかに記載の方法。
【請求項３３】
現場でメンテナンス可能なクラスベース型翻訳装置であって、
第１言語の辞書であって、前記第１言語における新語を含む音声を受け取り、前記新語の発音、翻訳及び語クラス情報を含む説明を前記新語と関連付けるように構成される認識辞書と、
前記第１言語と関連付けられた第１機械翻訳モジュールであって、第１タグ付けモジュール、第１クラスベース型翻訳モデル及び第１クラスベース型言語モジュールを含み、前記語クラスは、前記第１タグ付けモデルを介して推定され、前記発音及び前記翻訳はルールベース型モデル又は統計的モデルを介して生成される第１機械翻訳モジュールと、
前記第１言語における前記新語の前記説明をユーザに出力し、前記新語の前記発音及び前記翻訳の前記ユーザによる検証又は修正を受け入れるように構成されたマルチモーダルインターフェイスとを備えており、
前記第１機械翻訳モジュールは、前記新語及び前記説明で更新されると共に、前記第１言語における文を前記第２言語において対応する翻訳文に翻訳するように構成されている装置。
【請求項３４】
前記第１機械翻訳モジュールは、前記新語を前記第２言語に対応する翻訳語に翻訳し、
出力は、第２テキストを表示するディスプレイを含む、請求項３３に記載の装置。
【請求項３５】
前記第２言語と関連付けられた第２機械翻訳モジュールであって、タグ付けモジュール、翻訳モデル及び言語モジュールを含み、前記第２言語における第２新語を第２翻訳語で翻訳する第２機械翻訳モジュールを更に含む、請求項３３又は３４に記載の装置。

【請求項３６】
前記第１言語と関連付けられた辞書であって、前記新語を受け取るように構成されたテキスト－音声発音辞書を更に含む、請求項３３乃至３５の何れかに記載の装置。
【請求項３７】
前記第２言語と関連付けられた辞書であって、前記翻訳語を受け取るように構成されたテキスト－音声発音辞書を更に含む、請求項３３乃至３６の何れかに記載の装置。
【請求項３８】
前記認識辞書は、前記第１言語における前記新語に対応する第１テキストを生成する、請求項３３乃至３７の何れかに記載の装置。
【請求項３９】
前記新語が前記口頭の第１言語に対応し、前記翻訳語が前記第２言語における前記第１テキストの翻訳に対応することを確認するモジュールであって、前記確認に基づいて低い又は高い信頼評価を生成する、相関及び修理モジュールを更に含み、
修正及び修理モジュールが、前記確認モジュールからの前記低い信頼評価によって特定されたエラーを認識し、前記エラーをユーザ挿入の修理テキストと置き換える、請求項３３乃至３８の何れかに記載の装置。
【請求項４０】
前記説明は、インターネットを介してアクセス可能な人間言語のインフォーマントの外部ソースから、可能性がある１つ以上の説明を特定することによって、前記新語と関連付けられる、請求項３３乃至３９の何れかに記載の装置。
【請求項４１】
ユーザ入力によって前記新語を受け取るように構成されるインターフェースを更に含み、
前記入力は、前記ユーザの言語における正書法の入力を含み、前記第２言語における翻訳語と、前記説明と、言葉の予測発音とを仮定する基準をもたらす、請求項３３乃至４０の何れかに記載の装置。
【請求項４２】
前記第１言語における前記新語の前記説明内の発音表記及び他の言語において対応する語の説明内の発音表記は、音声再生によりインターフェースにおいて再検討及び確認される、請求項３３乃至４１の何れかに記載の装置。
【請求項４３】
トレーニングコーパス内の複数の文ペア全部にわたって、ラベルが既知の第１言語から注釈が付けられていない言語にラベルを投影することによって、クラスラベルが生成される、請求項３３乃至４２の何れかに記載の装置。
【請求項４４】
前記第１言語モデルは、音声認識仮説に対して行われた修正に基づいて更新され、
前記更新は、前記言語モデルの確率を更新し、前記修正された語系列の前記言語モデルの確率を増加させることによって、且つ似た競合仮説の前記言語モデルの確率を低減することによって、同一エラーが発生する尤度を低減する、請求項３３乃至４３の何れかに記載の装置。
【請求項４５】
前記第１機械翻訳モジュールは、機械翻訳仮説に対する修正に基づいて更新され、
前記更新は、前記翻訳モデル及び前記言語モデルの確率を更新し、前記修正された語系列の前記翻訳モデル及び前記言語モデルの確率を増加させることによって、且つ似た競合仮説の前記翻訳モデル及び前記言語モデルの確率を低減することによって、同一エラーが発生する尤度を低減することを含む、請求項３３乃至４４の何れかに記載の装置。

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2009/040677

A. CLASSIFICATION OF SUBJECT MATTER
INV. G10L15/18 G06F17/28

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)
G10L G06F

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

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	H OKUMA ET AL: "Introducing Translation Dictionary Into Phrase-based SMT" PROC. MACHINE TRANSLATION SUMMIT XI, 10 September 2007 (2007-09-10), - 14 September 2007 (2007-09-14) pages 361-368, XP002535595 Retrieved from the Internet: URL:http://www.mt-archive.info/MTS-2007-Okuma.pdf> [retrieved on 2009-07-06] page 361, left-hand column, paragraph 1 page 361, right-hand column, paragraph 1 first paragraph of section 'Baseline Method'; page 362 sections 'Basic idea' and 'Algorithm'; page 362 - page 363 last paragraph of section 'Conclusion and future work'; page 365, right-hand column	1-45

-/--

[X] 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 document 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	Date of mailing of the international search report
7 July 2009	20/07/2009
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax (+31-70) 340-3016	Authorized officer Ramos Sánchez, U

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2009/040677

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NARAYANAN S ET AL: "Transonics: a speech to speech system for English-Persian interactions" AUTOMATIC SPEECH RECOGNITION AND UNDERSTANDING, 2003. ASRU '03. 2003 I EEE WORKSHOP ON ST. THOMAS, VI, USA NOV. 30-DEC. 3, 2003, PISCATAWAY, NJ, USA,IEEE, 30 November 2003 (2003-11-30), pages 670-675, XP010713230 ISBN: 978-0-7803-7980-0 page 670 page 672, right-hand column - page 673, left-hand column, paragraph 1 page 674, left-hand column, last paragraph	1-45
X,P	I R LANE ET AL: "Class-based Statistical Machine Translation for Field Maintainable Speech-To-Speech Translation" INTERSPEECH 2008, 22 September 2008 (2008-09-22), - 26 September 2008 (2008-09-26) pages 2362-2365, XP002535596 Retrieved from the Internet: URL:http://isl.ira.uka.de/fileadmin/publication-files/851_IS081131.pdf> [retrieved on 2009-07-06] the whole document	1-45

Form PCT/ISA/210 (continuation of second sheet) (April 2005)

フロントページの続き

(81)指定国 AP(BW,GH,GM,KE,LS,MW,MZ,NA,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,RO,SE,SI,SK,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BR,BW,BY,BZ,CA,CH,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,IS,JP,KE,KG,KM,KN,KP,KR,KZ,LA,LC,LK,LR,LS,LT,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PG,PH,PL,PT,RO,RS,RU,SC,SD,SE,SG,SK,SL,SM,ST,SV,SY,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,ZA,ZM,ZW

(74)代理人 100141841
弁理士 久徳 高寛
(72)発明者 ワイベル, アレキサンダー
アメリカ合衆国 １５２２１ ペンシルベニア, ピッツバーグ, ウインザー アベニュー ６１９
(72)発明者 レーン, イアン アール.
アメリカ合衆国 １５２３２ ペンシルベニア, ピッツバーグ, エルスワース ５４００, アパートメント ５

Ｆターム(参考) 5B091 CB12 CB32 CD15
5D015 GG01 KK02