音声対話システムにおける Web 検索タスクの発話分析と Web 検索のための大規模単語コーパスの検討

三宅 純平†, 竹内 翔大†, 川波 弘道†, 猿渡 洋†, 鹿野 清宏†

† 奈良先端科学技術大学院大学 情報科学研究科 情報処理学専攻 〒 630-0221 奈良県生駒市高山町 8916-5

E-mail:{junpei-m,shota-t,kawanami,sawatari,shikano}@is.naist.jp

あらまし 本稿では、当研究室で開発を行っている音声対話システム「たけまるくん」における Web 検索タスクの認識率向上を目指し、Web 検索時の発話データの分析および言語モデル構築のための様々な学習コーパスの比較検討を行なった。その結果、発話形式として単語発声の割合が高く、発話内容として流行語・新出語など流行情報とユーザが利用する土地の観光情報など地域情報への考慮が必要であるという知見を得た。流行情報は、「Wikipedia」や「検索キーワードランキング」などのオンラインデータベースサイトからキーワードのみを言語モデルの学習コーパスとした。地域情報の対応としては従来の「たけまるくん」言語モデルとの融合によるドメイン適応を行った。その結果では、従来の「たけまるくん」の音声認識率より、最大で子供発話にて23.94%の改善が確認できた。キーワード 音声対話システム、言語モデル、Web 検索、大規模単語コーパス

Utterance Analysis of Web Search Task for the Spoken Dialogue System and the Consideration of the Large Word Corpora for Web Search

Jumpei MIYAKE † Shota TAKEUCHI † Hiromichi KAWANAMI † Hiroshi SARUWATARI † and Kiyohiro SHIKANO †

[†] Graduate School of Information Science, Nara Institute of Science and Technology

Abstract This paper presents a analysis of utterance and a comparing investigation of a training corpora for language model of web search task in a speech-oriented dialog system. A dialogue system called "Takemaru-kun" has been developed in our laboratory and has been operated at a public community center since November 2002. In result of the utterance analysis, we confirmed a finding that word utterance accounts for a high rate in an utterance form. Additionally, It is also important to consider both trendy information and local information of the area where the system is placed. For trendy information, the training corpus for language models were built by online database site in web, for example, "Wikipedia" or keyword ranking of web search. On the other hand, for local information, we used the conventional language model used in "Takemaru-kun" system. Finally we merged the two language models. They result in the performance that speech recognition rate improve 23.94% in child.

Keywords Spoken Dialogue System, Language Model, Web Search, Large Word Copora

1 はじめに

近年、音声を用いたデータベース検索システムの研究に注目が集まっている。その例として、音声認識を用いたカーナビや携帯電話によるナビシステム・日英翻訳などの製品化が挙げられる。また、その中でも携帯電話の音声認識を用いて、サーバ上のデータベースから情報出出を行なうことと高齢者は、子供から高齢者はで様々な情報機器をシームレスに使いこなすための人に優しいユーザインタフェース「ナチュラの研究とは、

ルユーザインタフェース(NUI)」を実現するものであり、今後ますます需要が高まる技術である. 我々が開発・運営を行なっている音声情報案内システム「たけまるくん」¹⁾も質問応答データベースを用いたデータベース検索システムの一種であり、音声認識結果と類似度の高いデータベースの最前のテキストを抽出する方法やデータベースの最適化²⁾など様々な対話処理技術の研究が成されている.

音声情報案内システムとは、ある環境や施設への訪問者に適切な情報やサービスを提供するシス

テムである.人間と機械の自然なコミュニケーションを実現するために、ユーザが要求発声に対して、システムがマルチモーダルな応答、即ち応答をとエージェントの動きを柔軟に返すことが求められる.目指すべき最終的な音声情報案内システムをは、ユーザのどんな質問に対しても最適な応答を返すシステムであるが、音声情報案内システムであるが、音声情報案内システムであるが、音が情報とると、エ光情報というにというである。というにというである。

タスクを絞った設計方針をとるのは、タスクを 絞った言語モデル構築により、高い音声認識率が得 られるためである。また、「たけまるくん」の場合、 高い応答正解率を確保するため、質問応答データ ベースを用いた実装を行なっていることから、人手 による質問応答の作成に非常に膨大なコストがか かることもタスクを絞る大きな理由の一つである。

しかし、近年、w3voice³⁾などWeb上で音声入力・音声認識を可能にするWebアプリケーションの開発やWebテキストマイニングによる質問応答システムの研究も盛んに見られる。そのため今後、より幅広い分野の内容に適応した音声認識システムが求められ、大規模な未知語収集や効率よい大規模コーパス収集の手法の提案が求められる。

本稿では、幅広い内容への対応が必要なWeb検索に注目し、我々が開発を行っている音声情報案内システム「たけまるくん」のWeb検索タスクにおける発話データの分析および、流行語・新出語に対応した言語モデル構築のため最適な学習コーパスについて検討する.

2 音声情報案内システム「たけまるくん」 2.1 概要

音声情報案内システム「たけまるくん」は、奈良県生駒市にある生駒市北コミュニティセンターISTAはばたき(以下 北コミュニティセンター)に常設してあり、実環境下で頑健に動作する音声対話システムである.対話戦略は一問一答形式であり、対話タスクとしては、天気やニュースの一般的な情報提供の他、エージェントに関する質問と、た設とその周辺の案内が質問応答ドメインとなっている.

たけまるくんのシステム構成は、入力音声に対して、GMMによる雑音の棄却および、音響大度による大人か子供の年齢層識別を行なった後、音いまる大人か子供の年齢層識別を行なった後、付ける、おいた質問に答データベースから質問用例テキストをの形態素マッチングにより最も類似度の高になりまりまれた1つのの形態をマッチンがによりまりました。その月例テキストを抽出し、その月例テキストでが対した。という構成に対ったいる(図1参照)、応答は大人と合成ター「たけるの種類ほどあり、テキストラのター「たけまなどの画像情報とともにキャラクる、

「たけまるくん」で用いられている言語モデルと音響モデルは、2002年11月から2004年の10月で収集した約2年分の発話データから構築されている。また、大人と子供で別々の言語・音響モデルを構築しており、入力音声に対して並行に音

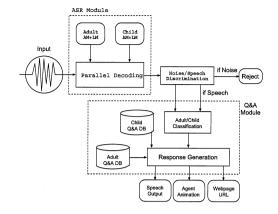


Fig. 1 たけまるくんのシステム構成.

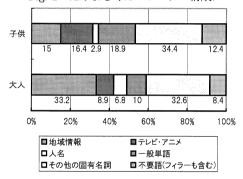


Fig. 2 たけまるくんにおけるユーザの発話ドメイン.

声認識を行い,音響尤度の高い認識結果を用いることで高い音声認識率を示している.

2.2 「たけまるくん」のタスクドメイン

音声対話システムを構築する上で, ユーザの発 話がどのようなタスクの発話が多く、どのような ドメインに属しているかを理解し、それに適した 言語モデルを構築することが重要である. 幅広い ドメイン対応のために無闇に大規模な言語モデル を構築しても、パープレキシティの増加により、認 識率が低下する可能性が考えられ、逆にドメイン を絞り過ぎると音声認識に可能なタスクの制限が 大きくなり, 汎用性のない対話システムになってし まう、また、システムに対するユーザの発話は、新 聞読み上げの文法に則した発話の音声認識とは違 い,口語調の自由発話がほとんどであり、フィラー も多く含まれている.「たけまるくん」ではユーザ の発話ドメインや自由発話の問題に対して、「たけ まるくん」で収集した実際の発話データの書き起 こしを基に、言語モデル構築および音響モデル構 築を行うことで、認識率の大幅な改善を行っている.「たけまるくん」タスクにおけるユーザの発話 ドメインを表 2 に示す.

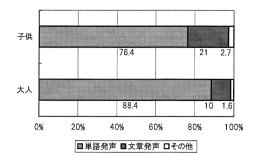


Fig. 3 Web 検索タスクにおける発話形式.

「たけまるくん」ではタスク外発話の割合が高く、ユーザにとって「たけまるくん」に要求するタスクは、まだ十分に満足のいくものではないことが分かる。また、Web検索タスクについても発話のほとんどが言語モデルの学習コーパスに含まれていないドメイン外発話である。音声対話システムにとって、タスク外やドメイン外の発話への対応も非常に大きな課題であり、様々な研究がなかなれている。特にWeb検索は、現在利用の割合は少ないが、今後、ドメイン外発話への対応が求められる。

3 Web 検索タスクの発話データの分析 3.1 発話の形式および内容

Web 検索のためのコーパスの収集を行うに当たり、Web 検索時のユーザの発話形式や発話内容の特徴を得るため「たけまるくん」における Web 検索タスクで収集された発話データの分析を行った.発話データは、2005 年 11 月から 2007 年 10 月の約2年間で発話されたデータである.

Web 検索タスクにおけるユーザの発話形式では、図3より大人と子供の両方で単語発声の割合が非常に高く、特に大人において著しいことがわかる。これは、ユーザが普段のキーボード入力による Web 検索において、キーワード検索を行うことが発発になっているためだと考えられる。一方、文ですか」「Yahoo! Japan を開いて」など、従来のエージェントへの質問応答タスクの発話と検索意思のでは、「Yahoo! Japan を開いて」など、従来のエージェントへの質問応答タスクの発話と検索意思のでは、シトへの質問応答タスクの発話と検索意思のでは、ドないもの声音である。既存のシステムには、アインには、記載によりである。とのでは、意味の持たないを記述している。このは、意味の持たない発話やフィラー語である。

次に図4に発話内容による分類を示す.この分類より奈良・生駒に関する地域情報と幅広い分野の固有名詞が大人・子供の共に高い割合で含んでいることが考察できる.地域情報には,付近の中学校・施設・地名などが含まれている.特に大人では地域情報に関する検索が非常に高い.また、大人と子供で共に「ケーム」「サッカー」「遊ぶこと」など分類を示す一般単語の発話も多い.この理が、発話の検索意思の低さも挙げられるが、カテゴリ検索による大雑把な情報提示を要求する

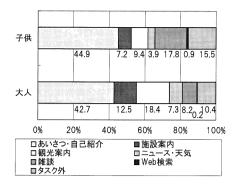


Fig. 4 Web 検索タスクにおける発話内容.

ユーザが多いということも大きな理由の一つであると考えられる。その他に、子供ではアニメやゲームのタイトルもしくはその登場人物に関する発話が多いことも考察できる。地域情報以外の発話内容については、一般ユーザがその時に興味を持っている事や時事の内容など流行語に関する発話が多い。

発話分析より Web 検索での音声認識では、幅広い分野の内容はもちろんのこと、地域情報と流行情報への対応が求められる。また、発話のほとんど単語発声であるため、コーパス構築には「しかし」「でも」および「て」「に」「を」「は」などの接続詞や助詞などの不要語を除いた複合語中心のテキストコーパスの収集および整形が求められる。

4 Web 検索のための言語モデル

4.1 学習コーパスの収集

対話システムにおける Web 検索に適したコーパ スは、前節でも述べたように流行情報と地域情報 への対応が求められる.流行情報の対応について は、一日に膨大な量の追加と更新が行われる Web テキストの収集を主に行うことを検討する。インターネットのブロードバンド化や携帯電話からの インターネット接続の利用者の急増を考慮すると, Web テキストはユーザの関心の動向を最も反映し ているテキストコーパスであると言える. その中 でも、近年、ユーザの動向を反映した Web サービ スとして「Wikipedia」などのユーザ参加型オンラ インデータベースや検索・ブログ日記のキーワード ランキングなどの統計的データベースが公開され るようになった. 本稿では、これらのユーザ参加 型オンラインデータベースや統計的データベース から Web テキスト収集を行った. また, Web テキ スト収集では文章ではなく、単語のみを学習コー パスとして収集した、それぞれ詳細について以下 に示す (表1参照)

4.1.1 ユーザ参加型オンラインデータベース

Wikipedia¹ オンライン百科事典と呼ばれ、歴史・経済・著名人・学問・地域名など普遍的な内容を非常に多く含んでいる. ユーザの手動によって現在、約20万語が登録されており、登録キーワー

¹ Wikipedia(日本語版) - http://ja.wikipedia.org

Table 1 Web から収集した大規模単語コーパス.

コーパス名	収集時期	キーワード数
はてな	2008.5.7	222852
Wikipedia	2008.5.7	196282
goo	2006.3.20 - 2008.5.7	92335
Yahoo	2007.11.12 - 2008.5.7	15207
@With	2006.10.26 - 2008.5.7	166126
blog360	2005.4.1 - 2008.5.7	34245

ドの詳細な説明や幅広い分野の内容を含んでいることから、言語モデルの学習コーパスとして用いられることが多々ある.一部のキーワードに対して読みが付与されており、未知語率の低減への利用も可能であると考えられる.

はてなキーワード2 ユーザが興味のあるキー ワードを手動で登録および説明や読み付与、分類 を行った共有辞書である. これは, ブログや日記 を書く中で共通の興味を持つ人との繋がりを提供 するために用いられている.20 万語以上キーワー ド登録がされており、「Wikipedia」と比べ、より頻 繁に登録キーワードが追加されている (2007年 11 月から 2008 年 5 月で約 2 万語のキーワードの追 加). 人名や時事など流行語・新出語の内容が多く, 普遍的な内容も含んでいる.「Wikipedia」は百科 事典の仕様であるため、キーワード登録だけでは なく詳細な説明を付加しなければならないのに対 し、「はてなキーワード」はキーワード登録と読み 付与だけでも良いという登録の簡易性が更新率の 高さに繋がっていると考えられる。そのため、「は てなキーワード」は流行情報に非常に強い. また, 全ての登録キーワードに対して読み付与がされて いる点も, 音声認識辞書更新に非常に有用である と考えられる.

4.1.2 統計的オンラインデータベース

検索・ブログのキーワードランキング これは, ユーザの日々の検索やブログで使用された言葉の 頻度を統計し、今話題になっている度合を算出し た後、その上位をランキング形式で表示するサー ビスである. 流行語・重要語の抽出には、「今日」 「私」「遊ぶ」などの一般語の棄却や名詞続きの単 語の複合語の処理などが行われており、過去の情 報と比較を行うことで話題となっているキーワー ド抽出を行われている.そのため、普遍的な内容 のデータは少ないが流行語・新出語が非常に多く 含まれている. 検索キーワードランキングとして 収集したサイトは以下のとおりである。検索ラン キングの特徴としては、キーワードの総合ランキ ングだけではなく、様々な分野の視点からランキ ングを計算している事である. (数字) は、上位何 位まで表示されているかを表す.

Yahoo! 検索ランキング³ 急上昇ワード (30), 総合 (20), 人名 (20), テレビ (20), ゲームアニメ (20), スポーツ (20)

goo ウェブ検索ランキング⁴ 上昇ワード (50), 女性有名人 (20), 男性有名人 (20), エンタメ (20),

総合 (50)

プログキーワードランキングとして収集したサイトは以下のとおりである. ブログキーワードランキングの特徴としては、TV ドラマや芸能人など関する内容が非常に多いことが挙げられる.

図With 人気プログランキング⁵ 約300 位までの 総合ランキングを表示している。そのため、検索 や他のランキングでは収集できないような比較的 頻度の低いキーワードを多く含んでいる

頻度の低いキーワードを多く含んでいる. BLOG 360⁶ 30 位までの急上昇ランキングを表 示している. TV ドラマや人名がメインとなってい

地域情報のコーパス収集については、ドメイン適応したテキスト収集としてWebからドメインに合った学習コーパスを収集する手法が多数提案されている $^{4)}$ 5). しかしながら、本稿では地域情報の学習コーパスとして、ドメインに応じたWebテキストの収集は行わず、たけまるくん」で2002年11月から2004年10月の2年間で収集した発話データの書き起こしを使用する.

4.2 統計的言語モデルの構築

流行情報と地域情報を反映させた言語モデルを構 築するために、Web テキストコーパスで学習した 言語モデルと「たけまるくん」言語モデルの融合を 行う. また, 形態素解析には ChaSen ver2.4.2⁷を用 いており、形態素辞書には、ipadic-2.7.0 (239,631 語彙) と「たけまるくん」用に追加した辞書 (598 語彙)を用いている. 言語モデル融合において、言 語モデルレベルでの融合より、コーパスレベルでの 融合の方が高い認識率であるいう先行研究がある 6). しかしながら,本稿ではコーパス融合を行わ ず、言語モデルレベルでの融合を行う、その理由 としては, 今回, 10万語彙近くの Web テキスト コーパスと1万強語彙の「たけまるモデル」との 融合を行なうので、コーパスレベルでの融合により、「たけまるくん」言語モデル(すなわち地域情 報に特化したモデル)の特徴が弱まることを避け るためである. 実際, 第5節の評価実験において, 予備実験としてコーパス融合と言語モデル融合の 両方の実験を行ったところ、どの評価データにお いても言語モデル融合の方がコーパス融合よりも 良い結果が得られている. 言語モデル構築は、前 向き 2-gram・後ろ向き 3-gram であり、平滑化手 法は Witten Bell 法を用いて構築している. また, 言語モデル融合にはSRILM ツール8を用いた.言 語モデルの融合重みは 0.5 である.

それぞれのWebテキストコーパスで学習した言語モデルと「たけまるくん」言語モデルへドメイン適応前と適応後した言語モデルの総単語数と異なり語彙数を表2に示す.Web検索・プログのキーワードランキングのコーパスにおいては、「はてなキーワード」や「Wikipedia」と比較するとコーパスサイズが極端に小さいので、コーパス融合を行い、「Keyword ranking」として言語モデルの構築を行なった。

² はてなキーワード - http://www.hatena.ne.jp

³ Yahoo!検索ランキング - http://searchranking.yahoo.co.jp

⁴ goo ランキング - http://ranking.goo.ne.jp/keyword

⁵ @With - http://blog.with2.net/trend_words.php

⁶ BLOG 360 - http://blog360.jp/

⁷ ChaSen - http://chasen.naist.jp/hiki/ChaSen/

⁸ SRILM - http://www.speech.sri.com

Table 2 ドメイン適応前と適応後の言語モデルの 総単語数と異なり語彙数.

適応前言語モデル	総単語数	異なり語彙数
たけまるくん	396741	10407
はてなキーワード	408170	92165
Wikipedia	1057803	79292
Keyword ranking	434213	23457
適応後言語モデル	-	異なり語彙数
はてなキーワード	-	100719
Wikipedia	-	89031
Keyword ranking	-	30212

5 評価実験

5.1 実験条件

実験では、それぞれの Web テキストコーパスか ら構築した言語モデルと「たけまるくん」言語モ デルのそれぞれの性能評価を単語正解率によって 比較する. また, 評価データは全て, 学習データと異なるデータであり, 言語モデルにおいてオー プンテストである. 評価データは、北コミュニティ センターに常設している「たけまるくん」で Web 検索タスク時に発話された大人と子供のそれぞれ の発話データ "Adult", "Child" と、今回の評価実験用に、学生十数名に一週間の期間の中で自由に Web 検索発話を行わせた発話データ "Student" を 用いる.後者は、北コミュニティセンター内では ない我々の研究室内で運営している「たけまるく ん」で収集している。"Student"では環境下が異な る場所での入力音声であるが、ユーザがマイクと 至近距離で発話している点や指向性の高いマイク を用いていることから、入力音声に影響する雑音の違いはほとんどないと考えられる。また、音響 モデルは JNAS データベースで構築した PTM モ デルを初期値として、「たけまるくん」で発話され た大人と子供のそれぞれの発話データで再学習を 行ったものである. それぞれの評価データと実験 条件を表3に示す. また, 評価データに対するそ れぞれの言語モデルの未知語率と補正パープレキ シティを表 4 に示す.

Table 3 実験条件

Table 6 XXXIII:			
評価データ	収集時期	データ数	
Adult	2005.0108~2007.1006	163	
Child	2005.0108~2007.1006	423	
Student	2008.0111~2008.0116	157	
音響モデル	2002.11~2004.	10	
	2003.8 は除く		
音声認識	Julius Ver. 4.0.1		
エンジン			

5.2 実験結果

5.2.1 ドメイン適応前言語モデルの性能評価

表5にそれぞれの評価データにおける「たけまるくん」言語モデルへの適応前の言語モデルの性能を示す.また,「たけまるくん」言語モデルの性

Table 4 ドメイン適応前と適応後の評価データの 未知語率とパープレキシティ

Adult	Before		After	
	OOVs	ppl	OOVs	ppl
たけまるくん	16.0%	21.9	-	-
はてなキーワード	5.6%	78.4	4.1%	44.7
Wikipedia	5.1%	127.46	2.6%	63.5
Keyword ranking	12.8%	47.9	7.1%	36.4
Child	Bet	fore	Aft	er
	OOVs	ppl	OOVs	ppl
たけまるくん	5.2%	47.7	-	-
はてなキーワード	2.1%	68.9	1.4%	39.9
Wikipedia	4.0%	92.0	2.8%	44.3
Keyword ranking	7.3%	36.7	3.5%	31.4
Student	Bet	fore	Aft	er
	OOVs	ppl	OOVs	ppl
たけまるくん	5.3%	6.0	-	-
はてなキーワード	3.4%	78.9	2.2%	84.6
Wikipedia	7.6%	6.9	5.5%	114.8
Keyword ranking	10.8%	41.2	6.6%	51.7

能の比較対象として、一般的に音声認識実験で用 いられる CSRC ⁹の毎日新聞記事の言語モデルに ついても示す. 全ての評価データにおいて,「はてな キーワード」を学習コーパスとした言語モデルの 単語正解率が高い.「はてなキーワード」は語彙サ イズの大きさや未知語率の低さから、非常に幅広 い内容に対して高い性能を示すことが考察できる. また、非常に興味深いのは、「Wikipedia」や「はて なキーワード」と比較すると語彙サイズが5万語 以上小さい「Keyword ranking」の単語正解率が 他の二つと健闘している点である。また、"Child" の評価データに関しては,一番良い単語正解率を 示している. これは, ユーザが普段の PC 上での Web 検索・ブログで頻繁に用いるキーワードのド メインと、音声対話システムにおける Web 検索発 話のドメインが類似しているためではないかと考 えられる.

5.2.2 ドメイン適応後言語モデルの性能評価

それぞれの Web テキストコーパスで学習した言 語モデルを「たけまるくん」言語モデルにドメイン 適応したモデルの単語正解率を図5,6,7に示す。 地域情報に対する検索発話の正解が増えたことで、 全体的に大幅な認識率の改善が見られる.一番良 い性能を示したのは、実験結果5.2.1の表5と同様 に「はてなキーワード」とのドメイン適応後言語モ デル「Hatena+Take」である. また、ドメイン適 応前と同様に、「Keyword ranking」のドメイン適 応の言語モデル「Keyword ranking+Take」は、他 と比べて語彙サイズが非常に小さいにもかかわら ず、全て評価データにおいて「Hatena+Take」と 同等の性能を示しており、「Wikipedia+Take」より 高い性能を示している. これは、前述したように 「Keyword ranking」がユーザの日々の動向を反映 しているため、音声対話システムにおける Web 検 索タスクにおいても、ドメインの類似性のため高 い認識率を示しているのではないかと考えられる.

音声対話システムにおける Web 検索のための言語モデルの学習コーパスについては、大規模であれ

⁹ CSRC - http://www.lang.astem.or.jp/CSRC/

Table 5	ドメイン適応前の言語モデルの音声認識
率性能.	

Adult	Word Cor.	Word Acc.
たけまる	45.9 %	16.3 %
MNP60K(CSRC)	45.2 %	22.0 %
はてなキーワード	58.9 %	44.6 %
Wikipedia	51.4 %	28.7 %
Keyword ranking	50.1 %	36.2 %
Child	Word Cor.	Word Acc.
たけまる	43.6 %	8.1 %
MNP60K(CSRC)	37.4 %	18.1 %
はてなキーワード	50.4 %	38.7 %
Wikipedia	47.6 %	22.0 %
Keyword ranking	51.5 %	40.8 %
Student	Word Cor.	Word Acc.
Student たけまる	Word Cor. 38.1 %	Word Acc. 10.5 %
たけまる	38.1 %	10.5 %
たけまる MNP60K(CSRC)	38.1 % 48.5 %	10.5 % 33.3 %

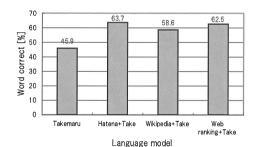


Fig. 5 様々な言語モデルによる大人発話での単語 正解率.

ばある程良い性能が得られるわけではない. Web 検索では、キーワードランキングのようなユーザの関心事がよく反映されたコーパスを効率よく収集するということも大事な要因の1つである.

6 おわりに

本稿では、Web 検索のためのテキストコーパスの検討のため、世界での出来事やユーザの関心を対が反映されている「はてなキーワード」や「コースを開いた言語モデルの性能の比較積計をといった。今回、コーでリードランキング」などの大規模単語コーパった。今回、コーでは一般をでは一次のでは一次のでは、正しい読みの付与などを表示をしている形態素解析辞書のみをで実験低ができる。また、日に変化するユーザの検索キーリードに対応望ましい。特にWeb 検索ではキーフードに対応望ましい。特にWeb 検索ではキーワードに対応できるとが望ましい。特にWeb 検索ではカードに対応では、随時言語モデルをリードに対応望ましい。特にWeb 検索ではカーマーマーンキングに浮上するキーワードの表もしい。対応表がある。今後は動的な読み付与の手法について検討する。

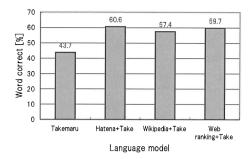


Fig. 6 様々な言語モデルによる子供発話での単語 正解率.

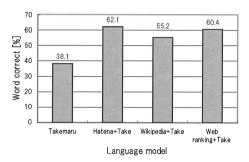


Fig. 7 様々な言語モデルによつ学生十数名の発話 における単語正解率.

参考文献

- 1) Ryuichi Nishimura Akinobu Lee, Hiroshi Saruwatari, Kiyohiro Shikano, "Public Speech-Oriented Guidanced System with Adult and Child Discrimination Capability", Proc. of ICASSP 2004, vol.1, pp.443-436, 2004.
- 2) 竹内翔大 他, "音声案内システムにおける質問応答データベースの最適化 -音声認識結果を用いた質問応答データベースによる応答生成と最適化-,"電子情報通信学会技術研究報告, SP2007-146, vol. 107, no. 406, pp. 295-300, Dec. 2007.
- 3) 西村竜一,三宅純平 他, "音声入力・認識機能を有する Web システム w3voice の開発と運用," 情報処理学会研究報告,2007-SLP-68-3,pp. 13-18, Oct. 2007.
- 4) Motoyuki Suzuki et al, "Unsupervised language model adaptation based on automatic text collection from WWW," in Proc. INTER-SPEECH, 2006, pp. 2202-2205.
- 5) 西村竜一 他, "Web からの音声認識用言語モデル自動生成ツールの開発", 情処研報 SLP-35-8, pp.49-54(2005)
- 6) 長友健太郎 他, "相補的バックオフを用いた 言語モデル融合ツールの構築,"情報処理学会 論文誌、Vol.43、No.9、pp.2884-2893、2002.