社団法人 情報処理学会 研究報告
IPSJ SIG Technical Report

2005－FI－80 (9)
2005－NL－169 (9)
2005／9／29

機械翻訳と翻字を併用した英日言語横断質問応答

川岸 将実
横浜国立大学 大学院 環境情報学府

森 辰則
横浜国立大学 大学院 環境情報研究院

E-mail: {kawagisi,mori}@forest.eis.ynu.ac.jp

機械翻訳と翻字を併用し，既存の日本語依存の質問応答を用いた英日言語横断質問応答の手法を提案する．機械翻訳の対訳辞書の不備を補うために，Webを用いて固有名詞の対訳を行なう．これにより機械翻訳の対訳辞書にない語に対応することが可能と考える．また，パタンに基づいた原言語での質問文タイプの決定を行ない，翻訳誤りに対処する手法を提案する．

評価実験によれば，機械翻訳のみを用いた場合のMRR値が0.062であるのに対し，上記二つの提案手法を同時に用いることでMRR値が0.125に向上した．

A Method of English-Japanese Cross Language Question-Answering Based on Both Machine Translation and Transliteration

Masami KAWAGISHI and Tatsunori MORI
Graduate School of Environment and Information Sciences,
Yokohama National University

We propose a method of English-Japanese cross language question-answering (E-J CLQA) that uses a machine translation, a transliteration, and an existing Japanese Q/A system that deeply depends on Japanese. In order to compensate the insufficiencies in the bilingual dictionary of a machine translation system, we utilize the documents on the Web to translate proper nouns. Using the method, some of words that do not appear in the dictionary can be translated. In addition, we introduce a pattern-matching-based question type detection in the source language in order to cope with translation errors. The experimental result shows that our proposed method improves the accuracy of E-J CLQA, that is, the MRR value is 0.125, while the MRR value is 0.062 in the case of a simple combination of a Japanese Q/A system and a machine translation system.

1 はじめに

近年，文書情報に対するアクセス技術として，質問応答が注目されている．質問応答は，利用者が与えた自然言語の質問文に対し，その答を知識源となる大量の文書集合から見つける技術である．

知識源となる文書は種々の言語で書かれているため，単言語だけではなく言語を横断するような検索技術が求められてきている．また，情報検索，言語横断情報アクセス，テキスト要約，質問応答などの「情報アクセス技術」研究の評価ワークショップシリーズであるNTCIR-5において[12]，言語横断質問応答のタスクであるCLQAタスクが初めて行なわれるなど，注目されてきている．ヨーロッパにおいては，CLEF(Cross Language Evaluation Forum)[6]において2003年より，イタリア語 - 英語，スペイン語 - 英語などの言語横断質問応答タスクが行なわれている．

言語横断質問応答は，利用者の与える質問文に対し，その答を言語の異なる文書集合から見つけるタスクである．本稿で扱うEJ-Taskは，英語の質問文に対し日本語の知識源から答えを抽出するものである．この時，答えは日本語のままである．

本稿では，上記の背景の下，英日言語横断質問応答の一手法について提案する．本手法では，機械翻訳と既存の日本語依存の質問応答手法を組み合わせる．同時に，機械翻訳の不備を補うために，Webに存在する対訳情報を用いて固有名詞の翻訳を行なう．これにより機械翻訳の対訳辞書にない語に対応することが可能と考える．また，パタンに基づいた原言語での質問文タイプの決定を行ない，翻訳誤りに対処する手法を提案する．

2 関連研究

佐々木は機言語横断質問応答に対する新しいアプローチとして，械学習を用いて，質問文の特徴，文書の特徴から質問文タイプを用いずに解答を抽出する手法を提案している[1]．この手法では質問文タ

イプを用いることをせず，かつ機械学習を用いるため，人出と時間が節約できるとしている．関根は英語とヒンディ語における言語横断質問応答について報告している [2]．短期間でシステムを構築することに主眼が置かれており，簡単な処理である程度の精度が得られているとしている．このシステムでは既存の質問解析部分を流用しているが，既存の質問応答システムを有効活用するという我々の方針と通ずるところがある．

CLEF では，ヨーロッパで使用される言語間での言語横断質問応答タスクが行なわれており，フランス語など 6 言語での単言語質問応答と，フランス語 - 英語など 50 種の言語横断質問応答（実際に参加があったのは 13 種）の結果について報告されている [7]．ここでは，質問として時間や場所などを聞く質問に加え，How 型の質問（`How did Hitler die?` のような手段を聞くもの），定義型質問（`Who is Kofi Annan?`のように人や物の定義を聞くもの）も質問文として採用されている（ただし定義型質問は人や組織について問うものに限定されている）．また評価基準として，正解精度と TREC-2002[8] で使われた CWS[1]が採用されている．ヨーロッパ言語と日本語では言語構造が異なるため比較は難しいが，言語横断質問応答では平均して 14.7% の精度，最もよかったものは質問が英語，知識源がオランダ語のシステムで 35.00% と報告されている．

翻字に関しては，辻らはサーチエンジンを利用した人名の翻字の手法を提案している [3]．これは，対訳辞書から翻字パタンを学習し，複数の日本語訳候補をサーチエンジンのヒット件数を元に決定する手法である．この手法では事前に学習が必要であるが，我々の提案手法では形態素解析・パタンマッチにより対訳語を抽出しており，この点で大きく異なる．

3 提案手法

言語横断質問応答では，質問文と知識源の言語が異なるためいずれかの段階で翻訳を行なう必要がある．翻訳には対訳辞書や機械翻訳を使うことが考えられるが，翻訳精度が全体の精度に大きく影響してくる．また，言語に依存したシステムにするか否かという問題がある．我々は日本語に依存した質問応答システムを開発してきており [4]，そのシステムと機械翻訳を組み合わせたシステムを提案する．

我々の日本語質問応答は，質問文解析，文書検索，パッセージ抽出，命題照合の 4 つのモジュールからなっており，図 1のような構成となっている．

質問文解析では，質問文の形態素解析・構文解析の結果から質問文の特徴となる語（以下，キーワードと呼ぶ）と質問が何について尋ねているのか（以下，質問文タイプと呼ぶ）を決定する．文書検索では，抽出されたキーワードを手がかりに質問に関連する文書を検索する．パッセージ抽出では，検索文書の中から，正解を含む可能性の高い連続する数文（パッセージ）を抽出する．そして命題照合において，質問文タイプ，質問文の解析結果を手がかりに解候補を決定する．この解候補は質問文解析の結果などからスコアリングを行ない，スコアの高い方から優先して出力される．

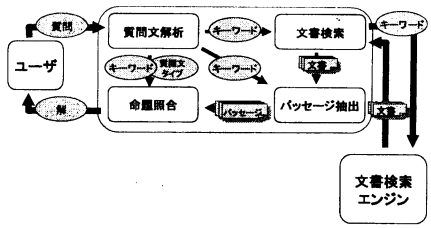

図 1: 日本語質問応答システムの例

このシステムでは，文書検索モジュール以外は日本語依存となっている．

EJ-Task は，質問文が英語で知識源となる文書群は日本語のタスクのため，パッセージ抽出，命題照合においては既存の日本語質問応答システムのものをそのまま利用することができる．そのため，質問文解析部を変更することで英日言語横断質問応答が可能と考える．

質問文解析部の変更にあたっては以下のような手法が考えられる．

- 質問文（英文）を機械翻訳して，質問文解析モジュールの入力とする．
- 英文のまま質問文解析し，質問文タイプとキーワードを決定する．キーワードは対訳辞書等を用いて翻訳する．

前者は，既存の機械翻訳システムを前提とすれば処理は簡単だが，翻訳できなかった語（以下未翻訳語）の存在や，翻訳誤りのために質問文を正しく解析できなくなり，質問文タイプが正しく決定できない場合がある．一方後者は，質問文の構文情報が使用できなくなり命題照合に支障をきたすため，既存のシステムが活かせなくなってしまう．

そこで，本稿では前者と後者を併用した質問文解析を提案する．具体的には，未翻訳語に関しては，Web 上の情報を用いて対訳語を抽出し，質問文タイプは原文から決定する．これにより，質問文タイプを正しく選択でき，かつ，質問文の構文情報を使用でき，既存のシステムを有効に活かせると考える．システムの概略を図 2に示す．

ユーザが英語の質問文を入力すると，機械翻訳システムを用いて日本語に翻訳すると同時に，英語の

[1]Confidence-weighted Score．解に対する確信度を考慮した評価方法 [11].

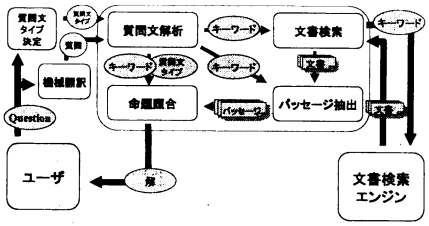

図 2: 提案する英日言語横断質問応答システム

質問文に基づいて質問文タイプを決定する．そして，翻訳された質問文と質問文タイプを質問文解析へと渡す．以後の処理は日本語質問応答と同じ流れになる．

翻字に関しては節 3.1で，原言語による質問文タイプ決定については節 3.2で詳しく述べる．

3.1 Web からの対訳情報抽出

例として，以下の文をあげる．

原 文： What is the title of the book written by Rogger Dingman the Awa Maru case?

翻訳文： 何を，本のタイトルは，Ａｗａ Ｍａｒｕ事例についてロジャーＤｉｎｇｍａｎによって書かれますか？

この例では，Ａｗａ ＭａｒｕとＤｉｎｇｍａｎという語が未翻訳のままキーワードとして判別されてしまい，日本語の知識源から，正しい対訳である「阿波丸」と「ディングマン」に関連する文書を検索することができなくなる．その問題を解決するために，Web から対訳情報を抽出する手法を提案する．

まず，質問文から固有名詞と思われる単語を以下の判別法に従い抽出する．この時，日本語（ローマ字列）と思われる単語と，それ以外の（英語の）固有名詞が区別される．

- ローマ字かな変換できるかどうかの判定を行ない，かなで 2 文字以上に変換できる場合，その単語を日本語とする．ただし，大文字で始まっておらず，かつ，翻訳辞書にその語がある場合は除外する．
- 大文字で始まっている単語が連続した場合，そのまとまりを固有名詞とする．
- 上記の条件を両方満たす場合は，日本語単語の翻字であると判断する．

条件に従い，Awa Maruが日本語，Rogger DingmanとAwa Maruが固有名詞と判別されるが，Awa Maruは両条件を満たすため，日本語と判別する．それぞれの詳しい翻字方法は 3.1.1以降で述べるが，日本語を優先させるのは，それ以外の固有名詞の翻字法よりも信頼性が高いためである．

また，翻字を行なう際の情報源として，Google Web APIs[2]を用い，Snippet（Web 文書の数文程度の要約）から対訳語（日本語表記）を抽出する．検索語には英文から抽出した語を使用する．

3.1.1 日本語と判別された文字列の処理

日本語（ローマ字列）と判別された語に対しては，以下の手順で処理を行なう．

1. 抽出された語を検索語（以下クエリとする）とし，Web 検索を行なう．
2. 取得した Snippet から，特殊文字・英記号を除去する．
3. 日本語形態素解析を行ない，形態素ごとに，読みの情報を元にローマ字変換を行なう．
4. クエリ全体とローマ字表記が一致した形態素列を対訳語として抽出する．表層表現が異なる語がある場合は，それらをすべて抽出する．

3.1.2 英語の固有名詞と判別された文字列の処理

Snippet は文の断片を結合したような構成になっており，また，取得してくる Snippet は英語と日本語が混在したものを想定しているので，構文解析により対訳語を選定するのは難しい．そのため英語の翻字はパタンマッチのみで行なう．この時，同じテキスト内に英語の対訳となる日本語がある場合，その語の近くに対訳語があるという仮定に基づいた処理を行なう．予備実験を行なった結果，近くに存在する場合は何らかの区切り文字 (() ;,/ など) で分けられていることが多いことがわかった．また，日本語列の場合区切り文字がない場合，どこで切れるか判別しにくいため，区切り文字があるもののみに限定した．実際の手順は以下のようになる．

1. 抽出された語をクエリとし，Web 検索を行なう．
2. 取得した Snippet から特殊記号を除去する．
3. 整形文中で，以下の条件にあてはまる語を対訳語としてすべて抽出する．
   - (D)?SWDTWD
   - DTWDSWD

ここで，D は区切り文字，SW は翻訳すべき語，TW は抽出すべき対訳語である．また，()? は括弧内が省略可能であることを表す．

[2] http://www.google.com/apis/

3.1.3 対訳語が複数ある場合の処理

複数の対訳語候補が得られた場合，それぞれの候補について（翻訳された）質問文を作成する．例えば，

原 文： When did the Battle of Sekigahara begin?

翻訳文： いつＳｅｋｉｇａｈａｒａの戦いは，始まりましたか？

という質問に対しては，Sekigaharaが日本語として判別され，対訳語として{関ケ原,関ヶ原,せきがはら,関が原}が抽出される．これに対してそれぞれ，

- いつ関ケ原の戦いは，始まりましたか．
- いつ関ヶ原の戦いは，始まりましたか．
- いつせきがはらの戦いは，始まりましたか．
- いつ関が原の戦いは，始まりましたか．

という質問文が作られる．この文全てにおいて質問応答を行ない，解候補のリストをスコアに従って併合する．このとき，違う質問文から同一の解候補が出てきた場合は，それぞれのスコアを比較し，高いスコアの方をその解候補のスコアとする．

3.2 質問文タイプの同定

我々の日本語質問応答システムでは，表1に示す17種類の質問文タイプを使用している．この質問文タイプは命題照合で使用しているため，原言語で決定する質問文タイプも17種類のどれかに属することとなる．

表1: 質問文タイプ一覧

人名	PERSON	長さ	length
地名	LOCATION	速度	speed
組織名	ORGANIZATION	面積	area
日付	date	重量	weight
時間	time	年齢	year
間隔	interval	金額	money
期間	period	数量	vol
割合	rate	数値	num
不明	none		

質問文タイプは，疑問詞に注目したパタンに基づく推定規則により決定する．これは，英語は疑問詞がはっきりしているため，表層から容易にタイプを決定できるからである．例えば，文頭にWhenがくれば日付を問うdateとなり，Whereがくれば場所を問うLOCATIONとなる．

表2: 推定規則一覧

正規表現	質問文タイプ
^When	date
^Where	LOCATION
^(Who|Whose)	PERSON
^How much	money
^How many	vol
^How long	period
^How tall	length
what percent	rate
what time	time
what *DAT*	date
WHAT LOC	LOCATION
WHAT company	ORGANIZATION
what *LEN*	length
population of	vol
percentage of	rate
name of *X*'s *PER*	PERSON
what && (rate|ratio)	rate
what && *MON*	money
what && *ORG*	ORGANIZATION
what && *VOL*	vol
上記以外	none

WHAT:what または which
DAT: 日付に関する表現
LOC: 場所に関する表現
LEN: 距離に関する表現
X: 人名などの表現
PER: 人に関する表現
MON: お金に関する表現
ORG: 組織に関する表現
VOL: 数量に関する表現

正規表現による推定規則の一覧を表2に示す．複数の条件を満たした場合は，表の上の優先度が高くなっている．

ここで*WHAT*はwhatまたはwhich，*DAT*はyearなどの日付に関する表現，*LOC*はcountryなど場所に関する表現，*LEN*はdistanceなど距離に関する表現，*X*はDianaなどの人名などの表現，*PER*はhusbandなど人に関する表現，*MON*はdollarなどお金に関する表現，*ORG*はorganizationなど組織に関する表現，*VOL*はnumber ofなど数量に関する表現である．

詳しい評価は節4.4にて述べるが，開発用セットでの質問文タイプの判定精度は86.7%であった．

4 評価実験

4.1 使用データ

英日言語横断質問応答全体と翻字の評価にはNTCIR5 CLQA1 EJ-Taskのサンプルテストセット300問を使用した．また，節4.2で指標として利用している日本語質問応答の質問文としては，CLQA1 JE-Task（日本語の質問に対し，英語の知識源から解答を抽出，提示するもの）のサンプルテストセット300問を使用した．なお，EJ-TaskとJE-Taskの問題文は対になっている．

質問文タイプについては前述の300問を開発用とし，評価用にはNTCIR5 CLQA1 EJ-Taskで使用された200問を使用した．

また，知識源は読売新聞の2000年，2001年の記事2年分，機械翻訳は市販の翻訳ソフトウェア[9]を，対訳辞書としてはEDR対訳辞書[10]を使用した．

4.2 システムの性能評価

英日言語横断質問応答の評価基準として，MRR値[3]と最上位に正解がきた割合（以下TOP1とする）を用いた．結果を表3に示す．

表3: システムの性能評価

	C_1	C_M	U_1	U_M
MT	0.017	0.023	0.053	0.062
+GS	0.020	0.026	0.067	0.078
+TYPE	0.037	0.044	0.090	0.108
+GS+TYPE	0.043	0.054	0.103	0.125
日本語 QA	0.170	0.204	0.243	0.296

C:Correct（解と記事番号が一致）
U:Unsupported（記事番号が異なる）
1:TOP1による評価
M:MRRによる評価
MT: 機械翻訳と対訳辞書のみ使用
+GS:GoogleのSnippetを使用して対訳
+TYPE: 英文により質問文タイプを決定

ここで，Cは Correct（解と記事番号が一致），UはCorrectにUnsupported（解は一致だが記事番号が異なる）を加えたもの，1はTOP1での評価，MはMRRでの評価である．また，MTは機械翻訳と対訳辞書のみを用いたもの，GSはGoogleのSnippetを使って未知語を翻訳したもの，TYPEは英文から質問文タイプを決定したものである．「日本語QA」は完全に翻訳が成功した場合の評価となる．

[3]Mean Reciprocal Rank．各問について最上位正解の順位の逆数を評価値とし，それを平均したもの．

機械翻訳と外部辞書のみを用いたものに比べ，提案手法のすべての機能を使用したものの性能がよくなっていることがわかる．また，翻字による未翻訳語の補完よりも，質問文タイプを原言語から決定した方が性能に大きく関わっていることがわかる．これは，機械翻訳した質問文を日本語で解析する際にうまくタイプが決定できないことが原因と考えられる．さらに詳しい解析は節4.4で述べる．

4.3 Webからの対訳情報抽出

節3.1の手法により取得した日本語列と英語列それぞれについて正しく翻訳できたかを評価した．対訳情報抽出処理を行なったのは300問中219問で，日本語の異なり語は112語，英語の異なり語は111語であった．結果を表4に示す．

表4: 対訳情報抽出性能評価

	MT	提案	併用	一部	不可
日本語	31	43	61	12	39
英語	44	7	49	0	63

MT: 機械翻訳のみ
() 内は提案手法でも抽出できた語数
提案: 提案手法のみ
併用: 機械翻訳に加え提案手法を適用
一部: 一部のみ正しく抽出できた語数
不可: 対訳語が抽出できなかった語数

ここで，「MT」は機械翻訳のみを使用したもの，「提案」は提案手法のみを使用したもの，「併用」は機械翻訳に加え提案手法を使用したもの，「一部」は{Okuma 講堂}のように全体の一部のみ正しく抽出できた語数，「不可」は抽出できなかった語数である．

また，評価尺度を以下のように定義し，結果を表5に示す．再現率で対訳辞書にない単語としているのは，翻字処理に対訳辞書を使用しているためである．

翻字対象の同定に関する尺度

$$\text{再現率}(R) = \frac{\text{正しく同定された翻字対象単語列数}}{\text{対訳辞書にない翻字すべき単語列数}}$$

$$\text{適合率}(P) = \frac{\text{正しく同定された翻字対象単語列数}}{\text{翻字対象として同定した単語列数}}$$

翻字に関する尺度

$$\text{ヒット率}(H) = \frac{\text{翻字候補が得られた語数}}{\text{翻字対象として同定した単語列数}}$$

$$\text{精度}(A) = \frac{\text{正しく同定された翻字対象単語列数}}{\text{翻字候補が得られた語数}}$$

表 5: 対訳語抽出の評価

	R	P	H	A
日本語	58.1%	38.4%	45.5%	84.3%
英語	12.3%	7.14%	26.1%	24.1%

R: 再現率　　P: 適合率
H: ヒット率　　A: 精度

4.3.1 日本語の対訳語抽出

処理時間は1キーワードにつき平均で約9秒であった[4]．表5より，日本語と判別された文字列については，ヒット率は小さいものの，精度が高いことから提案手法により機械翻訳の辞書にない語を適切に補えているといえる．

成功例を次に示す．

成功例

原文：`What was the name of the apartment where Takiko Mizunoe or Yoshiko Okada used to live?`
抽出：`Takiko Mizunoe, Yoshiko Okada`
翻字：水の江瀧子, 岡田嘉子

翻字に失敗したものを詳細に調べてみると，以下の3種に大別できる．

1. 英語を日本語として判別している (`Seagaia, Jean-Henri`など)
2. 形態素解析が誤っている (`Tairoike`など)
3. Snippet 中に語が出現していない

1に関しては，辞書に存在していない語を表層のみで日本語か英語かを判別するのは困難なので，処理時間は大きくなるが，日本語として判定した場合の対訳語抽出処理と並行して英語として判定した場合のの対訳語抽出処理を行なうことで対応できると考える．3に関しては，現在は検索結果の上位10件のの Snippet のみで処理を行なっているので，実際の HTML ファイルを取得するなど，情報源の拡張を行なうことで対応できると考える．2の実際の例は図3のような場合である．この場合，「太」や「路」といった形態素に対し，解析結果の読みの情報が誤っているため，クエリ全体とローマ字表記が一致せず，対訳語を抽出することができない．この問題の解決には形態素解析器の精度の向上が求められる．

[4]CPU:PentiumIII　1GHz　2機，Memory:2GB，OS:RedHat Linux ver7.2 の環境において Perl5 を使用した．

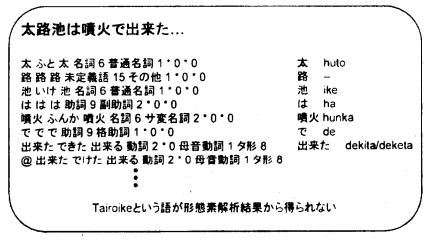

図 3: 形態素解析失敗例

4.3.2 英語の対訳語抽出

処理時間は1キーワードにつき平均で約7秒であった．表5より，ヒット率，精度ともに低い結果となり，英語と判別された文字列については，提案手法によって機械翻訳の辞書にない語を適切に補完できなかった．しかし，中には機械翻訳では正解がとれないような物も，Web を使うことにより取得できたものもあった．その例を以下に示す．

成功例

原文：`When was the Japan Biological Informatics Consortium established?`
抽出：`Japan Biological Informatics Consortium`
翻字：バイオ産業情報化コンソーシアム

翻字に失敗したものを詳細に調べてみると，以下の4種に大別できる．

1. パタンマッチにより抽出された語が不正確である
2. キーワードと対訳語は隣接しているが，パタンにマッチしていない
3. キーワードと対訳語が離れて存在している
4. Snippet 中にキーワードと対訳語のうち片方しか出現していない

4に関しては日本語の場合と同様に，情報源を拡張することで対処できると考える．3の場合は，提案手法では抽出することができないので他の方法を考える必要がある．1の失敗例には例えば以下のようなものがある．

失敗例 1

原文：`Where was Mike Volpi born?`
抽出：`Mike Volpi`
翻字：{グループのマイク・ボルピ, ボルピ}

このように，パタンマッチにより抽出された文字列が，正解を含むより長いものであったり正解の一

部であったりするということが見られた．これは節3.1.2で述べたように，日本語は表層ではどこまでが一語なのか判別しにくいことに起因していると考えられる．この対処には，形態素解析などを用いて隣接する対訳語を過不足なく抽出する方法が考えられる．

また，以下のような例もあった．

失敗例 2

原文：`What is the name of a magazine published by Larry Flynt?`
抽出：`Larry Flynt`
翻字：{ウディ・ハレルソン，ラリー・フリント}

これは，クエリ（原文から抽出した検索語）の側に不正解であるウディ・ハレルソンと，正解であるラリー・フリントの２語が存在し，それぞれがパタンマッチにより抽出されたためである．提案手法では訳語抽出について表層のみで正解かどうかを判断するのは難しいので，節3.1.3で述べたように得られた候補全てを利用して問題文を作成している．上の例のように，明らかな翻字間違いの場合の問題文に対して質問応答するのは処理時間が長くなったり，不適当な解を取ってきてしまうなど適切ではない．また，現在のシステムでは異なる問で同一の語が出てきた場合，その都度検索をしているなど無駄が多い．さらに，Web 上の情報源は常に同一であると限らないため，時間により，正しい対訳が得られたり得られなかったりということもある．

それら全てに対処する方法として，確信度付きの翻字辞書を整備する方法を考えている．抽出した語を翻字辞書の項目から探し，確信度が閾値以上であれば対訳語を提示し，新たな Web 検索は行なわない．閾値以下，あるいは辞書にない場合にのみ Web 検索を行ない，対訳語が抽出された場合に，確信度づけを行なう．これにより，確信度の高い語のみ翻字候補として使用されるため，処理時間の軽減などに寄与すると考える．

確信度づけの手法としては，辻らの提案するバリデーション [3] や，外池らの提案する手法 [5] など，サーチエンジンを使用する方法や，質問応答のスコアによる方法が考えられる．前者は翻字候補抽出と同時に行なえば，質問応答に質問文を渡す前に確信度の低い語を除去できるという利点がある．後者は上位に正しい解がくる質問文ほど，正しい翻字であるという仮定に基づいたものである．この場合は，一度全ての翻字候補について質問応答システムに渡す必要がある．今後，どちらの方法がより適当か検討していきたい．

4.4 原言語での質問文タイプ同定の効果

節 3.2の手法により，質問文から抽出した質問文タイプの精度を評価した．結果を表6に示す．ここで，MT+ は機械翻訳と WEB による翻字処理を行なった質問文を既存の日本語質問文解析により質問文タイプを決定したものであり，TYPE は原言語のパタンマッチにより質問文タイプを決定したものである．

結果を見ると，全体的な精度は開発用より評価用の方が落ちているものの，原言語でタイプを決定した方が精度がよいことがわかる．特に，`rate`，`money`，`vol`においては機械翻訳の結果ではほとんど正しく同定できないのに対し，原言語でのタイプ同定を用いると比較的高い精度で取れることがわかる．これは例えば，機械翻訳が`what percent`を「どのパーセント」と誤訳することなどが起因している．

その一方で，`ORGANIZATION`に関しては両方式とも精度がよくなかった．これはもともと特定しにくいのに加え，`party`のように翻訳（パーティー，政党など）に曖昧性のある語は質問タイプ同定に用いなかったことが原因と考えられる．また，開発用セットに出てこなかった`speed`，`area`，`weight`に関しては一問も正しく同定できなかった．

このような問題を解決するには，開発用の質問文をさらに集める，翻訳に曖昧性のある語で文の特徴となり得る語に関しては，周辺の語から対訳の曖昧性解消を行なうことなどが考えられる．

5 まとめ

本稿では，英日言語横断質問応答に対して，既存の日本語質問応答を利用する手法を考察し，その際に問題となる箇所に対して二つの解決手法を提案した．

第 1 に Web 上の情報を用いて，機械翻訳の未翻訳語を翻字する手法を提案した．質問文から固有名詞と思われる語を抽出し Snippet から対訳語を探すことで，ローマ字列の翻字に対して有効に働くことがわかった．その一方で英語の固有名詞の翻字に対してはさらなる検討が必要であることもわかった．

第 2 に原言語から質問文タイプを決定する手法を検討し，質問応答の精度向上に寄与することを示した．パタンマッチにより，機械翻訳と既存の質問文解析を使用する手法よりも，比較的高い精度で質問文タイプを決定できることがわかった．その一方で同定しづらい質問文タイプや，対応していない質問文タイプがあることがわかった．

この二つの提案手法を同時に用いた場合に，システム全体の精度が最も向上することがわかった．しかしながら，MRR で評価した時に日本語質問応答システムの約 1/3 程度の精度しかないので今後はそ

表 6: 質問文タイプ同定の精度

開発用 300 問					
タイプ	PER	LOC	ORG	dat	tim
質問数	38	43	19	40	3
MT+	84.2	76.7	26.3	90.0	0
TYPE	76.3	95.3	31.6	95.0	100
タイプ	int	per	rat	len	spe
質問数	0	4	38	3	0
MT+	0	0	0	33.3	0
TYPE	0	75.0	86.8	33.3	0
タイプ	are	wei	yea	mon	vol
質問数	0	0	6	31	23
MT+	0	0	33.3	9.7	30.4
TYPE	0	0	50.0	93.5	95.7
タイプ	num	non	計		
質問数	0	52	300		
MT+	0	98.1	56.7		
TYPE	0	100	86.7		
評価用 200 問					
タイプ	PER	LOC	ORG	dat	tim
質問数	29	30	20	24	15
MT+	100	76.7	15.0	91.7	6.7
TYPE	100	83.3	15.0	95.8	86.7
タイプ	int	per	rat	len	spe
質問数	0	1	10	3	2
MT+	0	0	0	0	0
TYPE	0	0	70.0	33.3	0
タイプ	are	wei	yea	mon	vol
質問数	1	1	3	20	18
MT+	0	0	33.3	0.5	55.6
TYPE	0	0	100	80.0	94.1
タイプ	num	non	計		
質問数	0	23	200		
MT+	0	91.3	55.6		
TYPE	0	95.7	79.5		

PER:PERSON　　LOC:LOCATION
ORG:ORGANIZATION
dat:date　　tim:time
int:interval　　per:period
rat:rate　　len:length
spe:speed　　are:area
wei:weight　　yea:year
mon:money　　vol:vol
num:num　　non:none

の差を埋めていく検討をしたいと考えている．具体的には，翻字手法のさらなる検討，確信度付きの翻訳辞書の整備，質問文タイプ同定の検討などを行なっていきたいと考えている．

参考文献

[1] 佐々木裕．統合学習による質問応答システムの新しい構成法 ～ＣＬＱＡに向けて．自然言語処理研究会報告 2004-NL-163，情報処理学会，2004.

[2] 関根聡．言語横断質問応答システム 言語処理学会 第 10 回年次大会 発表論文集，pp.321–324，2004

[3] 辻慶太，佐藤理史，影浦峡．対訳人名における翻字・サーチエンジンの有効性評価 言語処理学会 第 11 回年次大会 発表論文集，pp.352–355，2004

[4] Tatsunori Mori. Japanese Q/A System using A* Search and Its Improvement: Yokohama National University at QAC2. In *Working Notes of the Fourth NTCIR Workshop Meeting*, pp. 345–352, 6 2004.

[5] 外池晶嗣，佐藤理史，宇津呂武仁．４択クイズを連想問題として解く自然言語処理研究会報告 2004-NL-161，情報処理学会，2004.

[6] Cross Language Evaluation Forum. http://clef.iei.pi.cnr.it/

[7] Bernardo Magnini, Alessandro Vallin, Christelle Ayache, Gregor Erbach, Anselmo Peñas, Maarten de Rijke, Paulo Rocha, Kiril Simov, and Richard Sutcliffe. Overview of the CLEF 2004 Multilingual Question Answering Track. In *Working Notes for the CLEF 2004 Workshop*, 9 2004.

[8] Text Retrieval Conference. http://trec.nist.gov/

[9] 日本アイ・ビー・エム株式会社．翻訳の王様バイリンガル Version5，2002.

[10] 日本電子化辞書研究所．EDR 電子化辞書，2001.

[11] Ellen M. Voorhees. Overview of the TREC 2002 Question Answering Track. TREC 2002, 2002.

[12] NII-Test Collection for IR. http://research.nii.ac.jp/ntcir/