日英対訳文対を用いたパターン翻訳器の自動作成法の検討

西村 拓哉　村上 仁一　徳久 雅人
鳥取大学大学院 工学研究科 情報エレクトロニクス専攻
{s062044, murakami, tokuhisa} @ ike.tottori-u.ac.jp

1　はじめに

現在，機械翻訳の分野において，対訳文対から自動的に翻訳規則を生成し翻訳を行う統計翻訳が注目され，研究が盛んに行われている [1]．統計翻訳では，イタリア語-英語など文法構造が類似する言語対において翻訳精度が高くなり，日本語-英語などの文法構造の異なる言語対においては翻訳精度が低くなる傾向がある [2]．別の翻訳手法にパターン翻訳がある [3]．パターン翻訳では文パターン辞書と単語辞書を用いて翻訳を行う．文パターンが有する大局的な文法情報を用いることで，翻訳文全体の構造を保持した翻訳精度の高い翻訳文を生成出来る利点がある．しかし，従来，文パターン辞書の作成は人手で行うため，開発にコストがかかる欠点がある [4]．

そこで本研究では，文パターン辞書を対訳文対から自動的に作成する手法を検討する．文パターン辞書の自動作成により，開発にかかるコストの削減が可能となる．また，パターン翻訳を統計翻訳の前処理に用いることで，日本語の文法構造が英語の文法構造に類似し翻訳精度が向上すると考えられる．

2　翻訳システム

2.1　パターン翻訳の基本概念

日英パターン翻訳では，まず日本語入力文 j が与えられたとき，日英パターン辞書と日英単語辞書 (句を含む) を参照する．次に，日本語入力文に適合した日本語文パターンに対応する英語文パターンの変数部を，単語に置き換えて翻訳を行う．尚，従来のパターン翻訳では日英文パターン辞書を人手で作成するので，開発に時間がかかる．しかし，日英文パターンが適合した場合には翻訳精度の高い翻訳文が得られる．図 1 に日英パターン翻訳の手順を示す．

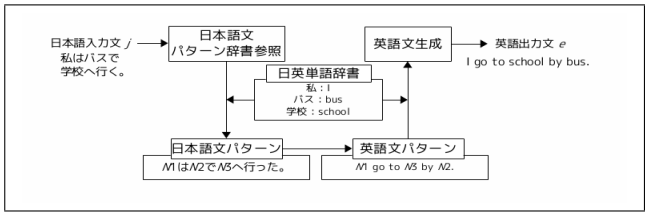

図 1　日英パターン翻訳の手順

2.2　統計翻訳の基本概念

日英統計翻訳は，日本語入力文 j が与えられたとき，全ての組合せから確率が最大値となる英語文 $\hat{e}$ を探索して翻訳を行う．

$$\hat{e} = \arg\max_{\mathbf{e}} P(e|j)$$
$$\approx \arg\max_{\mathbf{e}} P(j|e)P(e)$$

$P(j|e)$ は翻訳モデル，$P(e)$ は言語モデルと呼ぶ．図 2 に統計翻訳の手順を示す．

2.2.1　翻訳モデル

翻訳モデルは英語から日本語の単語列へ確率的に翻訳を行うためのモデルである．翻訳モデルはフレーズテーブルで管理されている．フレーズテーブルの例を表 1 に示す．

左から，日本語フレーズ，英語フレーズ，日英の単語翻訳確率の積，フレーズの英日方向の翻訳確率，英日の単語翻訳確率の積である．

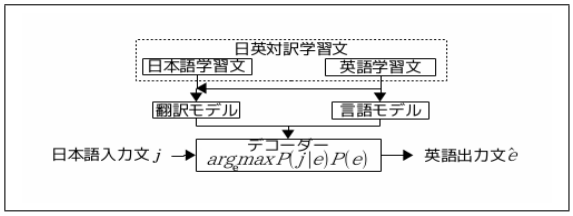

図 2　日英統計翻訳の手順

表 1　フレーズテーブルの例

ここ から	||| from here	||| 0.7778 0.1407 0.9545 0.2571
すぐ に	||| at once	||| 0.5 0.0111 0.1964 0.0030

2.2.2　言語モデル

言語モデルは単語列の生じる確率を与えるモデルである．日英統計翻訳では翻訳モデルにより生成された翻訳候補から英語文として自然な文を選び出す．統計翻訳では一般に，N-gram モデルを用いる．

3　翻訳システムの自動作成

3.1　本研究の翻訳システムの概要

パターン翻訳では，文パターン辞書と単語辞書を用いて翻訳を行う．文パターンが有する大局的な文法情報を用いることで翻訳文全体の構造を保持した翻訳精度の高い翻訳文を生成出来る利点がある．しかし，従来，文パターン辞書の作成は人手で行うため，開発にコストがかかるという欠点がある．

そこで本研究では，文パターン辞書を対訳文対から自動的に作成する手法を検討する．文パターン辞書の自動作成により，開発にかかるコスト削減が可能となる．また，パターン翻訳を統計翻訳の前処理に用いることで，文法構造が類似し翻訳精度が向上すると考えられる．図 3 に提案手法の枠組みを示す．

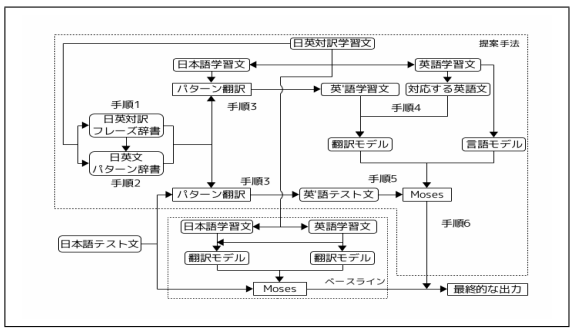

図 3　提案手法の枠組み

3.2　本研究の翻訳システムの手順

本研究の翻訳システムの手順を以下に示す．

手順 1　日英対訳フレーズ辞書の作成

まず日英対訳学習文から Moses[6] に付属している “train-model.perl” を用いてフレーズテーブルを生成する．次に，各フレーズ対の両方向へのフレーズ翻訳確率と単語翻訳確率について，それぞれ積を計算する．そして，閾値以上の確率を持つフレーズ対を抽出し，日英対訳フレーズ辞書を作成する．表 2 に日英対訳フレーズ辞書の例を示す．

表 2 日英対訳フレーズ辞書の例

この 本 ||| This book ||| 0.5176 0.1253
交通 事故 ||| traffic accident ||| 0.5803 0.2063

左から，日本語フレーズ，英語フレーズ，双方向へのフレーズ翻訳確率の積，双方向への単語翻訳確率の積となっている．

手順 2　日英文パターン辞書の作成

手順 1 で作成した日英対訳フレーズ辞書を用いて日英対訳学習文から日英文パターン辞書を自動的に作成する．日英対訳フレーズ辞書のフレーズ対が日英対訳学習文中で適合した場合に変数化を行い，日英文パターン辞書を作成する．尚，パターン翻訳を行うとき，変数部分に対して候補が莫大になることを防ぐために，変数部分を変換可能なフレーズに含まれる単語数について制限を設ける．図 4 に日英文パターン辞書の作成手順を示す．

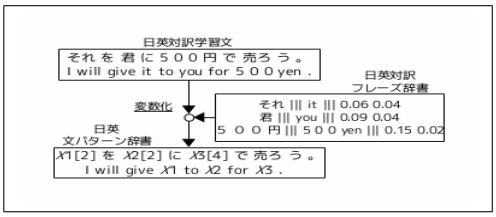

図 4 日英文パターン辞書の作成手順

図 4 の例では，日本語文 “それ を 君 に ５ ０ ０ 円 で 売ろ う ．” と英語文 “I will give it to you for ５００yen .” に対して，日英対訳フレーズ辞書を参照し，適合するフレーズに対して変数化を行う．3 つのフレーズ “それ”，“君”，“５ ０ ０ 円” が変数化され，日文パターン “*X*1[2] を *X*2[2] に *X*3[4] で 売ろ う ．” が生成される．そして同時に英文パターン “I will give *X*1 to *X*2 for *X*3 .” も生成される．日文パターンの変数部分 *X* に付与されている大括弧内の数値は変数化するときに用いたフレーズが含む単語数であり，単語数が “1” である場合には “2” とする．数値を “2” とする理由としては，入力文に対してカバー率を上げるためである．

手順 3　パターン翻訳

手順 1 で作成した日英対訳フレーズ辞書と手順 2 で作成した日英文パターン辞書を用いて，日英パターン翻訳を行う．図 5 に日英パターン翻訳の流れを示す．

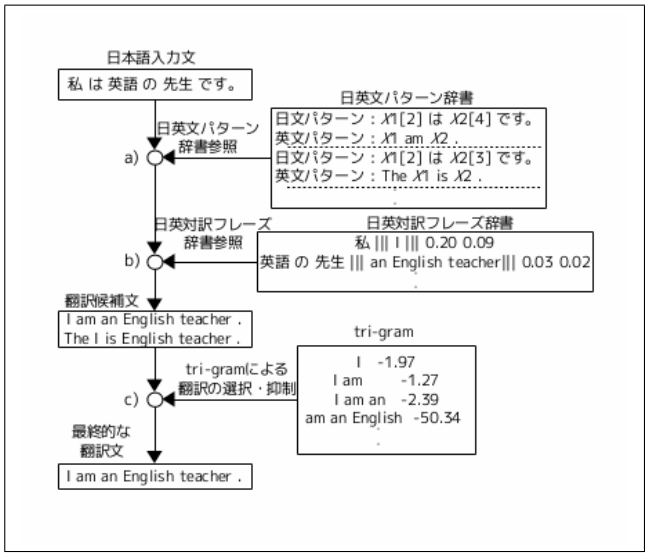

図 5 日英パターン翻訳の流れ

a)　図 5 の例において，日本語入力文 “私 は 英語 の 先生 です 。” に対して，日英文パターン辞書を参照し，適合するパターンを選択する．

b)　日英対訳フレーズ辞書を参照し，適合した日文パターンの変数部分にあたる日本語フレーズ “私”，“英語の先生” のそれぞれについて，英語フレーズを “I”，“an English teacher” を得る．

c)　英文パターンと変数部分の日本語フレーズに対応する英語フレーズを用いて翻訳候補を生成する．複数の候補が出現する場合には，tri-gram を用いて候補を 1 文に絞り込む．図 5 では，2 つの翻訳候補から最終的な翻訳文 “I am an English teacher.” を得る．また，最終的な翻訳文の tri-gram のスコアが低い場合には，適合すべきでないパターンに適合しているなど，翻訳文に誤りを含む可能性が高いと考えられる．したがって，翻訳文の tri-gram のスコアを用いて誤りを含む出力の抑制を行う．本研究では，tri-gram のスコアに対して “-1000” を閾値とし，“-1000” 以下のスコアを持つ最終的な翻訳文は出力しないものとする．

本研究では，パターン翻訳を用いて翻訳した日本語テスト文と日本語学習文をそれぞれ，英’ 語テスト文，英’ 語学習文と呼ぶ．

手順 4　統計翻訳の翻訳モデルと言語モデルの学習

英’ 語学習文と対応する英語学習文を用いて翻訳モデルを学習し，英語学習文を用いて言語モデルを学習する．

手順 5　統計翻訳における英語文生成

英’ 語テスト文に対して，手順 4 で学習した翻訳モデルと言語モデルを用いて英’ 英統計翻訳を行う．尚，本研究ではデコーダーに Moses を用いる．

3.3　ベースラインシステム

Moses を用いた日英統計翻訳システムをベースラインシステムとし，ベースラインと呼ぶ．翻訳モデルの学習には日英対訳学習文を，言語モデルの学習には英語学習文を用いる．

4　実験環境

実験データには，辞書の例文から抽出した日英対訳文対の単文，重文複文を用いる [7]．単文の翻訳実験には学習文に単文 100,000 文対，テスト文に単文 10,000 文を用いる．重文複文の実験には学習文に重文複文 100,000 文対，テスト文に重文複文 10,000 文を用いる．また前処理として，日本語文に対しては MeCab[8] を用いて形態素解析を行い，形態素と句読点の間にスペースを入れる．英語文に対してはコンマ，ピリオド，数字の前後にスペースを入れる．表 3 に単文の例を，表 4 に重文複文の例を示す．

表 3　単文の例

日本語文	彼 は 科学 上 の 偉大 な 発見 を し た 。
英語文	He made a great scientific discovery .

表 4　重文複文の例

日本語文	彼 は 父 の 後 を 継い で 医者 に なっ た 。
英語文	He followed in his father ’s footsteps and became a doctor .

4.1　言語モデルの学習

統計翻訳に用いる言語モデルには *N*-gram モデルを用いる．本研究では，SRILM[9] の “ngram-count” を用いて 5-gram の言語モデルを学習する．尚，スムージングに “-ndiscount” を用いる．

4.2　デコーダのパラメータ

本研究では，フレーズテーブル作成時のヒューリスティックに “grow-diag-final-and” を用いる．また，“distortion-limit” を “-1” とし，その他のパラメータについては Moses のデフォルトの値を用いる．

5 翻訳実験

5.1 パターン翻訳

5.1.1 実験方法

日本語テスト文と日本語学習文を用いてパターン翻訳器の作成とパターン翻訳を行う．まず，日英対訳学習文を用いて，日英対訳フレーズ辞書を作成する．今回の実験では，日英対訳フレーズ辞書作成時に用する閾値を “0.001” とする．次に作成した日英対訳フレーズ辞書と日本語学習文から日英文パターン辞書を作成する．そして日英対訳フレーズ辞書と日英文パターン辞書を用いて，日本語テスト文と日本語学習文に対して，パターン翻訳を行う．表 5 に，作成した日英対訳フレーズ辞書のフレーズ対数，日英文パターン辞書の文パターン数を示す．

表 5 各辞書のエントリ数

	フレーズ対数	パターン数
単文	40,787	95,853
重文複文	38,581	96,803

5.1.2 パターン翻訳結果

表 6 に，各実験条件における日本語テスト文でのパターン翻訳の出力文数を示す．

表 6 パターン翻訳の出力文数

	テスト文	学習文
単文	1143	64584
重文複文	349	50402

5.2 統計翻訳

翻訳モデルの学習には表 6 の学習文の出力結果 (英’ 語学習文) と，英’ 語学習文に対応する英語学習文を用いる．言語モデルの学習には英語学習文を用いる．学習した翻訳モデルと言語モデルを用いて英’ 語テスト文に対して英’ 英統計翻訳を行う．

5.3 ベースライン

ベースラインの翻訳モデルの学習には日英対訳学習文を用いる．言語モデルの学習には英語学習文を用いる．学習した翻訳モデルと言語モデルを用いて日英統計翻訳を行う．

6 実験結果

日本語テスト文 10,000 文での結果と，表 6 の日本語テスト文のパターン翻訳で得られた出力文のみでの実験結果について示す．

6.1 自動評価結果

自動評価には BLEU[12], NIST[13], METEOR[14] を用いる．表 7 に日本語テスト文 10,000 文での結果，表 8 にパターンに適合した日本語テスト文での結果を示す．尚，表 8 中の “パターン” はパターン翻訳のみでの翻訳結果を示している．

表 7 10,000 文での自動評価結果

単文			
	BLEU	NIST	METEOR
ベースライン	0.1130	4.5211	0.3160
提案手法	0.1101	4.5131	0.3175

重文複文			
	BLEU	NIST	METEOR
ベースライン	0.0947	4.0980	0.3021
提案手法	0.0977	4.1406	0.3049

単文の場合には提案手法の評価値がベースラインの評価値よりも低くなっている．一方，重文複文の場合では提案手法の評価値がベースラインの評価値よりも高く，提案手法の有効性が見られる．

6.2 人手による対比較評価

表 8 のパターンに適合した文からランダムに抽出した 100 文に対して人手による対比較評価を行う．

表 8 パターン翻訳の出力文での自動評価結果

単文 (1143 文)			
	BLEU	NIST	METEOR
ベースライン	0.2218	5.2390	0.4363
提案手法	0.1821	4.8417	0.4426
パターン	0.1630	4.5314	0.4230

重文複文 (349 文)			
	BLEU	NIST	METEOR
ベースライン	0.2814	5.1235	0.4562
提案手法	0.3618	5.8849	0.5438
パターン	0.3384	5.7004	0.5326

6.2.1 判断基準

人手による 4 つの判断基準に基づいて評価を行う．評価基準と評価例を以下に示す．

評価 1(>) 提案手法の結果がベースラインの結果よりも優れている
入力文 彼女 は 5 人 の 子供 を 育て た .
正解文 She has brought up ve children .
ベースライン She is ve children .
提案手法 She brought up ve children .

評価 2(<) 提案手法の結果がベースラインの結果よりも劣っている
入力文 農園 は 道路 に 接し て いる .
正解文 The farm abuts on the road .
ベースライン farm adjoins the road .
提案手法 The farm is roads are .

評価 3(≈) どちらも同程度に意味を理解できる，または，どちらも同程度に意味を理解できない
入力文 彼 に よろしく お 伝え ください .
正解文 Give him my good wishes .
ベースライン Please give my best regards to him .
提案手法 Please send him my best wishes .

評価 4(=) 提案手法とベースラインの翻訳結果が同一である
入力文 この ビール は 気 が 抜け て いる .
正解文 This beer is flat .
ベースライン This beer tastes flat .
提案手法 This beer tastes flat .

6.2.2 評価結果

表 9 に人手による対比較評価の結果を示す．

表 9 対比較評価の結果

	単文	重文複文
評価 1(>)	30 / 100	33 / 100
評価 2(<)	9 / 100	3 / 100
評価 3(≈)	50 / 100	31 / 100
評価 4(=)	11 / 100	33 / 100

自動評価結果とは異なり，対比較評価では提案手法がベースラインよりも翻訳精度が高くなっている．

7 考察

7.1 翻訳精度が高い出力の解析

重文複文の実験において，自動評価において提案手法がベースラインよりも翻訳精度が高くなっており，パターン翻訳のみでもベースラインより翻訳精度が高い結果となっている．提案手法がベースラインよりも優れている翻訳文について，表 10 に単文での翻訳例を，表 11 に重文複文での翻訳例を示す．

単文，重文複文のどちらの例においても，入力文に対して適切なパターンが適合し，翻訳精度の高い文が得られている．単文では，提案手法はベースラインよりも自動評価値が低い．しかし人手による対比較評価では提案手法がベースラインよりも優れている文が多い．この要因として，提案手法がベースラインよりも優れていた 30 文のうち，ベースラインの出力に動詞が存在しない文が 16 文あり，人手による対比較評価において提案手法の出力がベースラインを上回ったと考えられる．

表 10 単文での提案手法が優れている翻訳例

入力文	お前 に 限る 。
ベースライン	I accepted .
提案手法	There is nothing like you .
パターン	There is nothing like You .
日文パターン	*X*1[2] に 限る 。
英文パターン	There is nothing like *X*1 .
参照文	No one but you can do it .

表 11 重文複文での提案手法が優れている翻訳例

入力文	子供 に は 好き な よう に さ せ て いる 。
ベースライン	A child as you please .
提案手法	I allow the children do as they like .
パターン	I allow the children do as they like .
日文パターン	*X*1[2] に は *X*2[2] な よう に さ せ て いる 。
英文パターン	I allow the *X*1 do as they *X*2 .
参照文	I let the children do as they like .

7.2 誤り解析

単文の実験において，自動評価ではベースラインが提案手法よりも優れている結果となったが，反対に人手による対比較評価では提案手法がベースラインよりも優れている結果となった．そこで，対比較評価の結果に対して，提案手法がベースラインよりも劣っていた 9 文について解析を行った．解析の結果，適合した文パターンが不適切であった場合，日英対訳フレーズ辞書に登録されているフレーズ対に問題がある場合，提案手法の統計翻訳部が悪影響を及ぼしている場合，日本語単語の意味の違いによって誤った翻訳をしている場合，日本語単語の意味の違いによって誤った翻訳をしている場合，の 4 種類が見られた．表 12 に各場合における解析結果を示す．

表 12 誤り解析の結果

場合	件数
適合した文パターンが不適切	4
フレーズ対に問題あり	2
統計翻訳が悪影響	2
単語の意味の違いによる誤翻訳	1

表 12 の解析結果から，最も数の多い，適合した文パターンが不適切であった場合について考察する．表 13 に適合した文パターンが不適切であった場合の翻訳例を示す．

表 13 適合したパターンが不適切であった場合の翻訳例

入力文	土壌 が 肥え て いる 。
提案手法	An early summer rain is fertile soil .
パターン	An early summer rain is fertile soil .
日文パターン	*X*1[2] が *X*2[2] て *X*3[2] 。
英文パターン	An early summer rain *X*3 *X*2 *X*1 .
日本語原文	五月雨 が しとしと 降っ て いる 。
英語原文	An early summer rain is falling gently .
参照文	The soil is rich .

表 13 を見ると，適合した文パターンに問題があることがわかる．そこで適合した文パターンについて解析を行った結果，パターン作成時に用いる日英対訳フレーズ辞書のフレーズ対に問題があった．日本語原文と英語原文を見ると，日本語フレーズ “五月雨” と英語フレーズ “An early summer rain” が対応すると考えられる．しかし日文パターンでは日本語フレーズ “五月雨” が変数化されているが，英文パターンには英語フレーズ “An early summer rain” がそのまま残っている．フレーズ対を確認すると，日本語フレーズ “五月雨” が英語フレーズ “gently” と対応していた．また日本語フレーズ “しとしと 降っ” と “いる” がそれぞれ，英語フレーズ “falling”，“is” と対応していた．この結果，作成される文パターンの品質が低下し，翻訳精度が低下したと考えられる．

重文複文についても同様に誤り解析を行った結果，適合した文パターンが不適切であった場合が 1 件，日英対訳フレーズ辞書に登録されているフレーズ対に問題がある場合が 1 件，tri-gram での文選択が誤っている場合が 1 件であった．

したがって，日英対訳フレーズ辞書に含まれる不適切なフレーズ対が日英文パターン辞書の品質や提案手法の翻訳精度に影響する問題は，日英対訳フレーズ辞書作成時に用いる閾値の調整や品詞タグの利用により，適切なフレーズ対を抽出することで改善できると考えられる．

8 おわりに

本研究では，日英対訳文対から日英パターン翻訳器を自動的に作成する手法を提案した．実験の結果，単文の実験では，自動評価において提案手法はベースラインよりも評価値は低いが，人手による対比較評価において，ベースラインよりも良い結果となった．また，重文複文での実験結果では自動評価と人手による対比較評価のどちらにおいても提案手法はベースラインよりも良い結果となった．しかし，日英対訳フレーズ辞書作成時における不適切なフレーズ対によって，日英文パターン辞書の作成とパターン翻訳の翻訳精度が低下する原因となっている．この問題に対して，日英フレーズ辞書の作成時に用いる閾値の調整や品詞タグの利用により，翻訳精度が向上すると考えられる．

参考文献

[1] Richard Zens, Franz Josef Och, Hermann Ney “Phrase-based Statistical Machine Translation”, KI 2002, pp35-56, 2002.

[2] Holger Schwenk, Marta R.Costa-jussà, and José A.R.Fonollosa, “Continuous space language models for the IWSLT 2006 Task”, International Workshop on Spoken Language Translation 2006, pp166-173, 2006.

[3] 池原 悟, 佐良木 昌, 宮崎 正弘, 池田 尚志, 新田 義彦, 白井 諭, 柴田 勝征, “等価的類推思考の原理による機械翻訳方式”, 電子情報通信学会技術研究報告, pp7-12, 2002.

[4] 池原 悟, 宮崎 正弘, 白井 諭, 横尾 昭男, 中岩 浩巳, 小倉 健太郎, 大山 芳史, 林 良彦, “日本語語彙大系”, 岩波書店, 1997.

[5] Hideki Isozaki, Katsuhito Sudoh, Hajime Tsukada, Kevin Duh, “Head Finalization : A simple Reordering Rule for SOV Languages”, Association for Computational Linguistics 2010, pp.244-251, 2010.

[6] Philipp Koehn, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, Chris Dyer, Ondrej Bojar, Alexandra Constantin, Evan Herbst, “Moses: Open Source Toolkit for Statistical Machine Translation”, Association for Computational Linguistics 2007, pp177-180, 2007.

[7] 西山 七絵, 村上 仁一, 徳久 雅人, 池原 悟, “単文文型パターン辞書の構築”, 言語処理学会第 11 回年次大会, pp372-375, 2005.

[8] MeCab : Yet Another Part-of-Speech and Morphological Analyzer
http://mecab.sourceforge.net/

[9] The SRI Language Modeling Toolkit
http://www.speech.sri.com/projects/srilm/

[10] Jin’ichi Murakami, Masato Tokuhisa, Satoru Ikehara, “Statistical Machine Translation using Large J/E Parallel Corpus and Long Phrase Tables”, International Workshop on Spoken Language Translation 2007, pp151-155, 2007.

[11] Franz Josef och, “Minimum Error Rate Training in Statistical Machine Translation”, Association for Computational Linguistics 2003, pp160-167, 2003.

[12] Kishore Papineni, Salim Roukos, Todd Ward, Wei- Jing Zhu, “BLEU : a Method for Automatic Evaluation of Machine Translation”, Association for Computational Linguistics 2002, pp311-318, 2002.

[13] George Doddington, “Automatic evaluation of machine translation quality using n-gram co-occurrence statistics”, Human Language Technology ’02, pp128-132, 2002.

[14] Satanjeev Banerjee, Alon Lavie, “METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments”, Association for Computational Linguistics 2005, pp65-72, 2005.