

隣接情報に基づく対訳語の自動抽出手法

越前谷 博† 荒木 健治†† 桃内 佳雄†

† 北海学園大学工学部電子情報工学科

†† 北海道大学大学院情報科学研究科

echi@eli.hokkai-s-u.ac.jp araki@media.eng.hokudai.ac.jp

momouchi@eli.hokkai-s-u.ac.jp

出現頻度に基づく類似度を用いて対訳語を自動抽出する手法では、データスパースネスが問題となる。例えば、同じ言語の2つの単語列が常に同時に出現する場合、それらの訳語を決定することができない。このような問題に対し、我々は、抽出対象となる対訳語と隣接関係にある単語列の情報を自動獲得し、その情報に基づき対訳語を抽出する手法を提案する。本手法では、原言語文における抽出対象の単語列に隣接する単語列が目的言語文中のどの単語列に対応するのかの情報と目的言語文中のそれらの単語列とどのような隣接関係にある単語列が訳語となるのかの情報を獲得する。これらの情報を利用することで、言語間の対応関係を一意に決定することができる。原言語が異なる5つの対訳コーパスを用いた性能評価実験の結果、本手法に基づくシステムの抽出率は60.1%となった。この値は、Dice係数に基づくシステムの抽出率に対し、8.0ポイント高く、本手法の有効性を示している。

Method for Automatic Extraction of Bilingual Word Pairs based on Adjacent Information

Hiroshi Echizen-ya† Kenji Araki†† and Yoshio Momouchi†

†Dept. of Electronics and Information, Hokkai-Gakuen University

††Graduate School of Information Science and Technology, Hokkaido University

echi@eli.hokkai-s-u.ac.jp araki@media.eng.hokudai.ac.jp

momouchi@eli.hokkai-s-u.ac.jp

This paper describes a new method that extracts bilingual word pairs from parallel corpora using various languages by learning of adjacent information. Our system extracts the word strings that adjoin bilingual word pairs. Moreover, it decides whether bilingual word pairs exist in the right side of the selected word strings or the left side. By using such adjacent information, our system can extract correct bilingual word pairs. Evaluation experiments indicated that extraction rates was 60.1% in parallel corpora using five different languages, and this rate was more than 8.0% higher than the extraction rate of system using only Dice coefficient.

1 はじめに

対訳コーパスは言語学習用教材としても利用価値が高い。そのため、言語学習用教材に用例集を掲載している多くの書籍やサイトが存在する。言語学習用教材として対訳文を利用する場合、語彙レベルでの情報は不可欠と考えられる。文対応の対訳文のみが存在していても、ユーザにとってそれほど有益ではない。やはり、より詳細な言語知識として語彙レベルの情報が必要である。学習用教材の対訳コーパスから対訳語を抽出する場合、対訳コーパスのサイズや解析的な知識に強く依存することなく、様々な言語の対訳語を自動抽出できることが要求される。我々は、言語学習用教材である対訳コーパスの有効利用という観点から、様々な言語からなる複数の対訳コーパスを対象に対訳語を自動抽出する手法を提案する。

解析的な知識に強く依存することなく対訳コーパスから対訳語を抽出する手法として、Dice 係数などの出現頻度に基づく類似度を用いる手法 [1, 2] がこれまでに提案されている。しかし、問題点として、データスパースネスの問題が指摘されている。例えば (There is your book on the table. ; テーブル/に/あなた/の/本/が/あり/ます. ¹) と (There is your cat on the chair. ; 椅子/に/あなた/の/猫/が/います.) の2つの対訳文から式 (1) の Dice 係数 [4, 5] を用いて “book” の対訳語を決定する。

$$Dice(X, Y) = \frac{2 \times f_{XY}}{f_X + f_Y} \quad (1)$$

f_X と f_Y は単語 X, Y が独立に出現する頻度であり、 f_{XY} は単語 X, Y が対訳文中に同時に出現する頻度である。この式 (1) を用いて、“book” と “本” の類似度と “book” と “テーブル” の類似度を求めると、その値は共に、 $1.0 (= \frac{2 \times 1}{1+1})$ となり、“book” の対訳語を一意に決定することができない [2]。

このような問題に対し、我々は、対訳コーパスのみから対訳語に隣接する単語列の情報を自動

¹ “/” は単語の境界を意味する。本稿では、日本語文の単語分割は日本語形態素解析ツール「茶筌」[3] を用いて行う。

獲得し、その情報に基づき対訳語を一意に抽出する手法を提案する。本手法では、原言語の抽出対象である単語列に隣接する単語列が目的言語文中のどの単語列に対応するかの情報と、目的言語文中のそれらの単語列の左右のどちらに隣接している単語列が訳語となるのかの情報を獲得する。これらの情報を用いることにより、対訳語の出現頻度に依存することなく、対訳語を一意に決定できる。例えば、先述した対訳文の場合、“book” と “本” に隣接する単語列 “your” と “あなた/の” が対応関係にあること、さらに “book” と “本” がそれぞれ “your” と “あなた/の” の右側に隣接するという情報を獲得することで、“book” と “本” の対応関係を決定することができる。なお、このような隣接関係にある単語列の情報は、解析的な知識を用いて獲得するのではなく、学習の観点 [6, 7] により自動的に獲得する。隣接情報を対訳コーパスを用いて自動獲得することにより、使用する対訳コーパスを置き換えるだけでシステムの変更を必要とせず、様々な言語の対訳語を抽出することができる。性能評価実験では、原言語が英語、フランス語、ドイツ語、上海語、そして、アイヌ語であり、目的言語が日本語である5種類の対訳コーパスを対象に、本手法に基づくシステムを用いて対訳語の自動抽出を行った。その結果、60.1%の抽出率が得られた。それに対し、隣接情報を用いない、Dice 係数のみに基づくシステムにおいては、抽出率は 52.1%であった。したがって、抽出率は 8.0 ポイント向上し、本手法の有効性が確認された。

2 処理概要

本手法に基づくシステムの処理概要を図 1 に示す。ここで対訳コーパス中の文が膠着語の場合には、あらかじめ単語分割しておく必要がある。本システムでは、対訳コーパスを他の言語の対訳コーパスに置き換えるだけで、システムの変更は必要としない。また、言語リソースの乏しい言語も含め、様々な言語間の対訳語を自動抽出することができる。

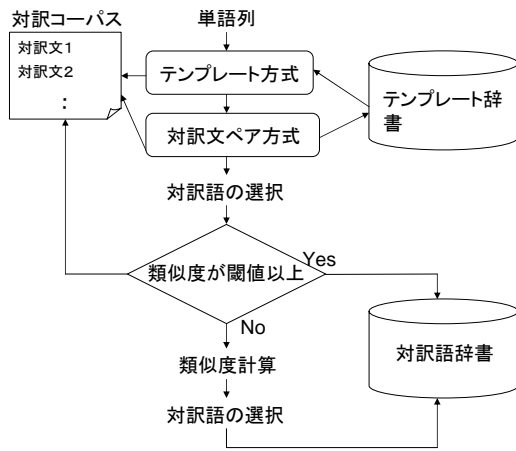


図 1: 処理概要

単語列を入力してから、対訳語を抽出するまでの処理の流れを述べる。始めに、1語以上で構成される原言語の単語列を入力する。テンプレート方式では、それまでに獲得されたテンプレートを用いて入力された原言語の単語列に基づき対訳語を抽出する。本稿における、テンプレートとは対訳語を抽出するための隣接情報を有する対訳知識である。抽出された対訳語には類似度が付与される。対訳文ペア方式では、入力された原言語の単語列が出現する対訳文とそれ以外の対訳文を用いて、対訳語と新たなテンプレートを得る。得られた対訳語とテンプレートには類似度が付与され、テンプレートはテンプレート辞書に登録される。ここで、複数の対訳語が得られた場合には、類似度の高い対訳語を一意に決定する。選択された対訳語の類似度が閾値以上の場合には、その対訳語を対訳語辞書に登録する。

また、類似度が閾値未満の場合には、正しい対訳語は抽出されていないとして、入力された単語列が出現する全ての対訳文を対象に対訳語の候補を抽出する。抽出された全ての対訳語の候補に対して類似度を付与し、類似度が最も高い対訳語を一意に選択する。そして、選択された対訳語を対訳語辞書に登録する。

2.1 対訳文ペア方式

対訳文ペア方式による対訳語の抽出およびテンプレートの獲得処理について述べる。本手法では、テンプレートを得るために、原言語文間と目的言語文間の両方において表層レベルで一致する単語列（以後、共通部分と呼ぶ）を利用する。共通部分を得るには単語列が最低2回以上出現することが条件となる。したがって、比較的出現頻度の高い単語列がテンプレートに選ばれることになる。そして、出現頻度の高い単語列間の対応関係の決定に有効と考えられる Dice 係数を用い、類似度を付与する。その結果、テンプレートの信頼性を比較的高い精度で求めることができる。

2.1.1 対訳語の抽出処理

始めに、対訳文ペア方式による対訳語の抽出処理について述べる。以下に、その処理の詳細を述べる。

- (1) 入力された原言語の単語列が原言語文に出現する対訳文を対訳コーパスから選択する。
- (2) 選択された対訳文の原言語文において、入力された単語列に隣接する単語列を共通部分に持ち、かつ、目的言語（日本語）文間で共通部分が存在する対訳文を対訳コーパスから選択する。
- (3) 目的言語文間において、以下の処理を行う。
 - i) 共通部分が1つの場合にはその共通部分を対象に、また、共通部分が2つ以上の場合には最も文頭に位置する共通部分を対象として、共通部分の左に隣接する語から文の先頭までの単語列を抽出する。
 - ii) 共通部分が1つの場合にはその共通部分を対象に、また、共通部分が2つ以上の場合には最も文末に位置する共通部分を対象として、共通部分の右に隣接する語から文の末尾までの単語列を抽出する。

- iii) 共通部分が 2 つ以上存在する場合、共通部分同士全ての組み合わせを用いて、共通部分に挟まれた単語列を抽出する。

(4) 入力された原言語の単語列と目的言語文から抽出された単語列を用いて全ての組み合わせを生成する。そして、以下の処理を行う。

- i) 目的言語文から抽出された単語列の品詞を調べる。その結果、抽出された単語列の先頭が名詞、動詞、形容詞、副詞、接続詞のいずれにも該当しない場合、もしくは末尾の語の品詞が名詞、動詞、助動詞、形容詞、副詞、接続詞のいずれにも該当しない場合には、その組み合わせを破棄する。
- ii) 組み合わせられた単語列間で語数の比が 1:3 もしくは 3:1 より大きい場合には、その組み合わせを破棄する。

(5) 処理 (4) の結果、残った組み合わせにおいて、入力された原言語の単語列と目的言語文から抽出された単語列の類似度を以下の式 (2) を用いて求める。

$$sim(W_S, W_T) = \frac{2 \times f_{ST}}{f_S + f_T} \quad (2)$$

ここで、 f_S は単語列 W_S が原言語文に独立に出現する頻度、 f_T は単語列 W_T が目的言語文に独立に出現する頻度、そして、 f_{ST} は単語列 W_S, W_T が対訳文中に同時に出現する頻度である。

なお、処理 (4) の ii) において、単語列間の語数の比を制限しているのは、語数の比が大きく異なる対訳語は対応関係の点から不適切である可能性が高いという考えからである。

図 2 に英日の対訳文を用いた場合における、対訳文ペア方式による対訳語抽出の具体例を示す。図 2 では入力された原言語の単語列 “house” が出現する対訳文 1 と “house” に隣接する単語列

入力された単語列: house

対訳文 1

(Please keep it until you leave this house.

;この家/を/出る/まで/持って/いて/下さい。)

対訳文 2

(I'd like to send this letter. ;この手紙/を送りたいの/です。)



名詞対訳語: (house;家)

図 2: 対訳文ペア方式による対訳語の抽出例

“this” を共通部分に持ち、かつ目的言語文間においても共通部分の存在する対訳文 2 を選択する。そして、目的言語文間において、2 つの共通部分 “この” と “を” に挟まれた単語列 “家” を対訳文 1 の目的言語文から抽出する。次いで、入力された “house” と抽出された “家” の組み合わせを対訳語とし、式 (2) を用いて類似度を求める。本処理の結果、複数の対訳語が抽出された場合には、3.3 の対訳語に対する順位づけの結果に基づき対訳語を一意に決定する。

2.1.2 テンプレートの獲得

ここでは、3.1.1 の処理で得られた対訳語とそれに隣接する共通部分を利用して、テンプレートを獲得する処理について述べる。以下に処理の詳細を述べる。なお、以下の処理過程は、3.1.1 で述べた処理を継続している。

- (1) 原言語文と目的言語文において、対訳語が存在していた部分を変数に置き換える。
- (2) 原言語文と目的言語文のそれぞれにおいて、個々の共通部分と変数の組を抽出し、それらを組み合わせでテンプレートを得る。
- (3) テンプレートから変数を除いた単語列を対象に、原言語文の単語列と目的言語の単語列の類似度を式 (2) を用いて計算する。そして、その類似度を付与し、テンプレート辞書に登録する。

図 3 に英日の対訳文を用いた際の対訳文ペア方式によるテンプレートの獲得例を示す。図 3 では、テンプレートとして (this @ ; この /@) と (this

対訳文 1

(Please keep it until you leave *this* house.
;この家/を/出る/まで/持つ/て/い/て/下/さい.)

対訳文 2

(I'd like to send *this* letter. ;この手紙/を/送/りたい/の/です.)



(Please keep it until you leave *this* @.
;この@/を/出る/まで/持つ/て/い/て/下/さい.)



テンプレート:

(this @;この/@) (this @;@/を)

図 3: 対訳文ペア方式によるテンプレートの獲得例

@;@/を)が獲得された。抽出された“house”と“家”を変数“@”に置き換え、共通部分と変数の組を抽出する。原言語文においては、共通部分が1つであり、その右側に変数が存在するため“this @”が抽出される。目的言語文においては、共通部分が2つ存在するため、“この/@”と“@/を”が抽出される。そして、抽出された共通部分と変数の組同士を組み合わせることでテンプレートとして (this @;この/@) と (this @;@/を) が得られる。

ここで、テンプレート (this @;@/を) 中の単語列“this”と“を”は誤った対応関係である。しかし、そのようなテンプレートは正しい対訳語に比べ、類似度が低下する。例えば (this @;@/を) の場合、“を”の出現頻度は非常に高くなる。しかし、“を”は“this”と対応関係にはないため、それらが同時に出現する可能性は高くない。したがって、式(2)の f_T のみが f_{ST} や f_S に比べ、非常に高くなり、類似度は低下すると考えられる。このように出現頻度の高い単語列であっても原言語と目的言語間で対応関係が成立しない場合は、類似度は低下する。なお、本稿では、テンプレートの原言語の単語列を原言語部分、目的言語の単語列を目的言語部分と呼ぶ。

2.2 テンプレート方式

ここでは、3.1.2で述べた処理より獲得されたテンプレートを用いて対訳語を自動抽出する処理について述べる。以下に処理の詳細を述べる。

- (1) 入力された原言語の単語列が出現する対訳文を選択する。
- (2) 選択された対訳文に対し、入力された単語列に隣接する単語列と一致し、かつ、目的言語文中の単語列と一致するテンプレートをテンプレート辞書より選択する。ただし、テンプレート中の原言語部分の変数が右側に存在する場合は、入力された単語列が原言語文中の共通部分の右側に存在し、テンプレート中の原言語部分の変数が左側に存在する場合には、入力された単語列が原言語文中の共通部分の左側に存在していなければならない。
- (3) 対訳文の目的言語文に対し、以下の処理のいずれかを行う。
 - i) 選択されたテンプレートの目的言語部分において、右側に変数が存在する場合、対訳文の目的言語文に対し、共通部分の右側に隣接する名詞、動詞、形容詞、副詞、接続詞、助詞を含まない名詞句、助詞を含まない動詞句を抽出する。
 - ii) 選択されたテンプレートの目的言語部分において、左側に変数が存在する場合、対訳文の目的言語文に対し、共通部分の左側に隣接する名詞、動詞、形容詞、副詞、接続詞、助詞を含まない名詞句、助詞を含まない動詞句を抽出する。
- (4) 入力された原言語の単語列と目的言語文から抽出された単語列の組み合わせにおいて、言語間の語数の比が 1:3 より大きい場合には、その組み合わせを破棄する。
- (5) 処理(4)の対象にならなかった組み合わせに対し、入力された原言語の単語列と目的言語文から抽出された単語列の類似度を式(2)より求める。
- (6) 他に選択されたテンプレートが存在する場合には、処理(3)から処理(5)を繰り返す。

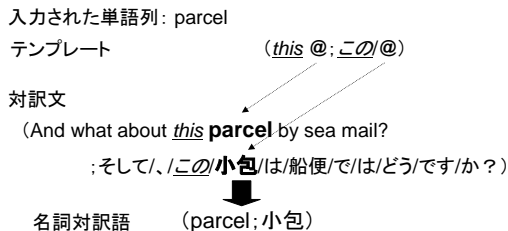


図 4: テンプレート方式による対訳語の抽出例

図 4 に英日の対訳文を用いた際のテンプレート方式による対訳語の抽出例を示す。図 4 では図 3 より獲得されたテンプレート (this @; この/@) より対訳語として (parcel; 小包) が得られた。テンプレート (this @; この/@) において、変数 “@” を除いた部分 “this” と “この” が、対訳文との共通部分として存在する。そのため、対訳文の目的言語文から、“この” の右側に隣接する名詞 “小包” が抽出される。このテンプレート方式より、複数の対訳語が抽出された場合には、3.3 の対訳語に対する順位づけの結果に基づき対訳語を一意に決定する。

2.3 対訳語に対する順位づけ

3.1 の対訳文ペア方式と 3.2 のテンプレート方式より、複数の対訳語が得られた場合には、以下の処理より、順位づけを行う。その結果、上位 1 位となった対訳語のみが対訳語辞書への登録対象となる。

- (1) 複数の対訳語の中で、類似度が最も高い対訳語を選択する。
- (2) 処理 (1) の結果、複数の候補が存在する場合、抽出時に隣接関係にあった単語列の類似度が最も高い対訳語を選択する。
- (3) 処理 (1), (2) の結果、複数の候補が存在する場合、対訳コーパス中で最初に出現する対訳語を選択する。

このような順位づけの結果、上位 1 位の対訳語の類似度が閾値以上の場合、その対訳語を対訳語辞書に登録する。

2.4 隣接情報を用いない場合の対訳語の抽出

3.3 より選択された対訳語の類似度が、閾値未満の場合には、検索対象を広げて改めて対訳語を抽出する。具体的には、入力された単語列が存在する対訳文の目的言語文中の全ての名詞、動詞、形容詞、副詞、接続詞、助詞を含まない名詞句、助詞を含まない動詞句を抽出する。入力された原言語の単語列と目的言語文から抽出された全ての単語列との組が対訳語の候補となる。全ての候補の類似度を、式 (2) より求め、類似度が最も高いものを選択する。その結果、類似度が同じである複数の対訳語が存在する場合は、対訳コーパス中で最初に出現する対訳語を対訳語辞書に登録する。

3 性能評価実験

3.1 実験方法

実験データとして用いた対訳コーパスは英語 - 日本語、フランス語 - 日本語、ドイツ語 - 日本語、上海語 - 日本語、アイヌ語 - 日本語対訳コーパスの 5 種類である。対訳コーパス中の対訳文は文献 [11] ~ [15] に掲載されている対訳文を使用した。対訳文の数はそれぞれ 393 文、399 文、377 文、392 文、233 文であり、総数は 1,794 文である。また、原言語の 1 文平均語数は 6.8、目的言語 (日本語) 文の 1 文平均語数は 8.8 である。

実験は図 1 の処理概要に基づくシステムに対し、表 1 の英語 - 日本語対訳コーパスを与える。そして、単語列を入力し、抽出された対訳語を対訳語辞書に登録する。その際に入力する単語列は文献 [11] の語句一覧に掲載されている、1 語以上で構成される名詞、動詞、形容詞、副詞、接続詞の英単語である。抽出対象の全ての英単語について対訳語が抽出された後、対訳語辞書を初期化、すなわち空の状態にする。そして、フランス語 - 日本語、ドイツ語 - 日本語、上海語 - 日本語、アイヌ語 - 日本語対訳コーパスの順で、同様の処理を繰り返す。その際、使用するシステムに対する

変更は行わない。

なお、図1の処理概要で使用する閾値として、今回の実験では0.5を使用した。式(2)では類似度の範囲は0.0~1.0となるため、中間の0.5を正誤の基準とした。

3.2 評価方法

本システムにより得られた対訳語の評価は、文献[11]~[15]に掲載されている語句一覧の訳に準拠し、人手で行った。そして、評価結果を抽出率で表す。抽出率は、全抽出対象の対訳語における、正しく抽出された対訳語の割合である。

3.3 実験結果

本手法の有効性を確認するために、Dice係数に基づくシステムとの比較実験を行った。表1に本システムとDice係数に基づくシステムの抽出率を示す。Dice係数に基づくシステムとは、3.4で述べた、隣接情報を用いない場合の対訳語の抽出処理に基づき構築されたシステムである。

表1: 本システムとDice係数に基づくシステムの抽出率

原言語	Dice係数	本手法	語数
英語	49.7%	58.0%	169
フランス語	47.9%	56.7%	240
ドイツ語	53.3%	61.0%	195
上海語	54.9%	62.9%	264
アイヌ語	54.0%	61.5%	213
合計	52.1%	60.1%	1,081

3.4 考察

表1より、本手法の抽出率はDice係数に基づくシステムの抽出率に比べ、全体で8.0ポイント高くなった。また、対訳コーパス毎では、どの対訳コーパスにおいても、7.5~8.8ポイント高く

なっている。この結果は、本手法が特定の言語に依存することなく、使用した全ての言語に対して有効であることを示している。

また、出現頻度に基づく類似度を用いて対訳語を抽出する場合、一般的に、出現頻度の高い対訳語は類似度の信頼性が高くなり、比較的高い精度で抽出できると考えられる。それに対し、低頻度の対訳語では、1で述べたように、類似度の値が等しい複数の対訳語が得られ、一意に対訳語を決定できずに抽出率が低下する。そこで、本実験において、出現頻度と抽出率の関係を調査した。その結果を表2に示す。

表2: 本システムとDice係数に基づくシステムの抽出率の内訳

原言語	Dice係数		本手法	
	1	2以上	1	2以上
英語	35.7%	77.2%	46.4%	80.7%
フランス語	37.5%	76.6%	49.4%	76.6%
ドイツ語	39.5%	75.0%	51.3%	76.3%
上海語	40.0%	79.8%	49.1%	85.9%
アイヌ語	45.0%	63.5%	56.9%	66.3%
合計	39.6%	73.8%	50.4%	76.8%

表2より、出現頻度1の対訳語の抽出率が10.8ポイント向上した。この結果は、対訳語の類似度だけでは正しい対訳語を決定できない状況において、本手法が非常に有効であることを示している。

4 むすび

本稿では、多言語対訳コーパスから対訳語を抽出する手法として、隣接情報に基づく対訳語の自動抽出手法について述べた。原言語が英語、フランス語、ドイツ語、上海語、アイヌ語、目的言語が日本語である5種類の対訳コーパスを用いた性能評価実験の結果、本手法の抽出率として60.1%が得られた。これは、Dice係数に基づくシ

システムに比べ 8.0 ポイント高い値であり，本手法の有効性を示している．今後は多義語の問題を解決すると共に，多言語翻訳システムに本手法を利用していく予定である．

謝辞

本研究の一部は，北海学園大学ハイテク・リサーチ・センター研究費によって行なわれている．

参考文献

- [1] F. Smadja, K. McKeown and V. Hatzivassiloglou, “Translating collocations for bilingual lexicons: A statistical approach, *Computational Linguistics*, Vol. 22, No. 1, pp. 1–38, 1996.
- [2] 北村美穂子, 松本裕治, “対訳コーパスを利用した翻訳規則の自動獲得,” *情処学論*, vol. 37, no. 6, pp. 1030–1040, 1996.
- [3] 松本裕治, 北内啓, 山下達雄, 平野善隆, 松田寛, 高岡一馬, 浅原正幸, “日本語形態素解析システム『茶筌』version 2.2.9 使用説明書”, 奈良先端科学技術大学大学松本研究室, 2002.
- [4] M. Kay and M. Röschen, “Text-Translation Alignment,” *Computational Linguistics*, vol. 19, no. 1, pp. 121–142, 1993.
- [5] C. D. Manning and H. Schütze, “Foundations of statistical natural language processing,” The MIT Press, 1999.
- [6] 越前谷博, 荒木健治, 桃内佳雄, 柝内香次, “実例に基づく帰納的学習による機械翻訳手法における遺伝的アルゴリズムの適用とその有効性,” *情処学論*, vol. 37, no. 8, pp.1565–1579, 1996.
- [7] 越前谷博, 荒木健治, 桃内佳雄, 柝内香次, “翻訳例に基づく再帰チェーンリンク型学習による機械翻訳手法,” *電子情報通信学会論文誌 D-II*, Vol. J85-D-II, No. 12, pp. 1840–1852, 2002.
- [8] 辻慶太, 芳鐘冬樹, 影浦峽, “対訳コーパスにおける低頻度語の性質 - 訳語対自動抽出に向けた基礎研究 -,” *電子情報通信学会信学技報*, NLC2000-16, pp. 47–54, 2000.
- [9] 佐藤健吾, 斎藤博昭, “サポートベクタマシンを用いた対訳表現の抽出,” *自然言語処理*, Vol. 10, No. 4, pp. 109–124, 2003.
- [10] K. McTait, “Automated Dictionary Extraction for “Knowledge-Free” Example-Based Translation,” In *Proc. 7th International Conference on Theoretical and Methodological Issues in Machine Translation*, pp. 111–118, 1997.
- [11] 早川嘉春, ジョン・スネリング, “エクスプレス：英(イギリス)語,” 白水社, 東京, 1998.
- [12] 筑紫文耀, “エクスプレス：フランス語,” 白水社, 東京, 2001.
- [13] 小塩節, “エクスプレス：ドイツ語,” 白水社, 東京, 2004.
- [14] 榎本英雄, “エクスプレス：上海語,” 白水社, 東京, 2004.
- [15] 中川裕, 中本ムツ子, “エクスプレス：アイヌ語,” 白水社, 東京, 2004.