

## 翻訳情報の提示によるクロスリンガル情報検索結果からの文書選択

鈴木 雅実    井ノ上 直己    橋本 和夫  
KDD 研究所

### 1 はじめに

言語横断 (クロスリンガル) 情報検索においては、検索結果の一覧表示の段階で、実際に閲覧すべき文書を選択するための翻訳情報が求められる。筆者らは、このような文書の取捨選択を容易にするための閲覧支援情報として、各文書中の主要キーワードを利用者の言語に翻訳して提供する枠組を提案した [4]。本発表では、原語でキーワードを示した場合、異なる対訳生成方法による翻訳キーワードをそれぞれ提示した場合に加え、キーワードではなく種々の方法で要約・翻訳された文を示した場合の効果の比較について、被験者を募って実験を行なった結果を報告する。

### 2 クロスリンガル情報検索への翻訳情報の提示

言語横断 (クロスリンガル) 情報検索においては、情報検索結果としてどのような文書集合を導くかという、検索性能自体の問題のほかに、その検索結果をどのように利用者に提示すれば実際の情報獲得に役立つかという視点が、モノリンガルな情報検索にも増して重要である。すなわち、検索要求から所望の文書集合のみを検索結果として導くことが困難である以上、利用者自身が検索結果を眺めること (browsing) によって、閲覧すべき対象の取捨選択が効率良く行なわれなければならない。モノリンガルな情報検索結果の一覧表示においてしばしば見られる、コンテンツの一部のテキスト等の表示も、クロスリンガル情報検索においては、検索対象言語に精通した利用者でない限り、原文のままでは文書の取捨選択に直接役立てることは相当に難しいか、多くの時間を要すると考えられる。

そこで、文書の選択閲覧を支援する情報は利用者の言語に翻訳されて提供されることが望ましいものと予想されるが、どのような翻訳情報を提示することが、検索結果から迅速に閲覧対象を判断するために有効かという点についての議論は少なかったように思われる。報告されている翻訳情報の提供方法については、HTML 文書のタイトルや主なヘッダの逐語訳の提供 (代表例は TITAN [1])、主要なキーワードのリスト (一例は CLINKS [4]) などがある。検索結果一覧表示段階での翻訳された要約テキストの提供事例はまだ見られないようである。<sup>1</sup>

いずれも、情報量や翻訳の質が問題となるが、これらの提示内容/方法の比較については未検証であると言える。文献 [6] は、WWW コンテンツのうち 1 件が数行程度のイエローページ<sup>2</sup>を対象として、名詞部分の辞書引き結果 (可能性のある訳語を列記したもの) を提示した “gisting” を参照して、被験者が各イエローページ項目を予め指定された 6 種類のジャンルに分類するというタスクを実行した結果を示し、利用者の “decision making” に十分役立つと結論づけている。しかし、検索対象がより長いテキストの場合について、文書選択/閲覧を支援する情報の種類/内容により、どの程度の効果があるかについての報告例は不明である。

そこで、次章の実験報告では、まず、筆者らの多言語検索サーバ CLINKS で導入した、出現頻度を反映した主要キーワードのリストを翻訳して提示する枠組の有効性の評価結果を示した上で (3.1)、さらに翻訳情報を要約テキストとして提示した場合の効果の測定結果を述べる (3.2)。

<sup>1</sup> 検索結果から選択されたコンテンツについて、翻訳要求により全文の機械翻訳結果が得られるものもある (商用化事例は、AltaVista [2]) が、品質についての保証はない。

<sup>2</sup> 一種のリンク見出しのようなもので、企業名とその所在地、サービス内容を簡潔に記した内容

### 3 比較評価実験

ここで述べる評価実験は、クロスリンガル情報検索の利用者が、検索結果の一覧表示の段階で、種々の形式で与えられる翻訳情報を参照して、探索課題に該当する文書をどの程度正しく選択できるかを比較しようとするものである。コンテンツ中の主要なキーワードを参照して該当文書を推定する、3.1の評価実験1については文献[5]で報告しているが、要約テキストの提示による3.2の評価実験2においても共通の部分が多いため、以下に再掲する。

#### 3.1 評価実験1：翻訳されたキーワード提示の効果

この実験では、被験者として日本人の大学生を主体とする19～32歳の男女64名の協力を得た。評価実験用に用意したテキスト・コーパスは、中国の経済分野の新聞記事(原文は中国語)で、これを英語および日本語に翻訳したものである<sup>3</sup>。記事数は224件で、これを用いて8組の探索課題と、それに対する仮想的な検索結果を作成した。各探索課題は、同コーパスの日本語版の記事タイトルから選んだ。すなわち、被験者に対して、「中国の小売業の現状」のような探索テーマを与えるとともに、該当記事1件を含む10件の記事検索結果を提示し<sup>4</sup>、その中からテーマに合致する記事を推定するタスクを実行してもらった。また、各課題毎の検索結果(記事番号と主要キーワードのリスト)は、計算機画面のブラウザ上にスクロール無しで表示できるように調整し、被験者が該当する文書番号をタイプするための入力欄を用意した。なお、この実験は被験者毎に個別に実行された。

##### 主要キーワードの提示条件と実験結果

以下の記述において、「正解率」とは、各探索課題において被験者が選択した文書番号(最大3個まで)が、実際の探索テーマが示す記事と一致した度合を指すものとする。キーワードの表示内容の提示条件A、B、Cの各場合に、被験者全体の正解率の平均値として得られた値、さらに、被験者が各探索課題への回答に費やした平均の時間についての比較を表1に示す。

表1 キーワード提示方法と正解率

キーワード表示条件	平均正解率	1課題当たりの所要時間
A: 出現頻度に基づく原語(英語)キーワード	67.2%	145 sec.
B: 翻訳(日本語)キーワード(注1)	69.5%	104 sec.
C: 翻訳(日本語)キーワード(注2)	75.8%	105 sec.

(注1) 英語キーワード群に対して、各々の英日対訳辞書中の最初の訳語を当てたもの。

(注2) 英語キーワード群に対して、各々の英日対訳辞書中の訳語候補の中から、日本語コーパス(インターネット上から収集したWWWコンテンツ)内の共起分布を参照して、尤度の高い組み合わせを選択したもの。

##### 主要キーワードの表示内容による文書選択への支援効果の相違

被験者にとっての外国語である英語で表示された場合(A)に比して、被験者の母国語である日本語で表示された場合(B、C)の方が、検索結果一覧の段階で目的に合致するコンテンツ(記事)を選択することが容易であり、判断に要した時間も大幅に短縮された。また、同様に日本語で表示された場合でも、英語のキーワードを基に辞書引きを行ない最初の訳語を用いる方法(B)と、翻訳生成側の日本語コーパス内の訳語候補同士の共起分布を考慮して訳語を生成する方法(C)とでは、後者の方が、コンテンツの選択をより容易にした。また、以上のような支援効果の程度は、コンテンツ選択の難度が高い場合により顕著であった。なお、この実験では、キーワードの表示個数を、3個、6個、9個、12個のいずれかに設定し、表示する条件の細分化を行なったが、正解率との間には

<sup>3</sup>英語版の1記事当たりの平均語数は380語程度。

<sup>4</sup>各探索課題に対する10件の検索結果集合については、実際にキーワード検索を実行するのではなく、224件の記事データに対し、コンテンツ中のキーワード分布の類似した記事をクラスタリングする処理を行ない、類似度がある程度近い同一クラスタに分類された記事10件づつを用いた[3]。

明確な相関関係は見られなかった。さらに、被験者の英語能力を測定する 10 分程度の筆記テストも同時に実施したが、そのスコアと提示条件 A での正解率の間にも、はっきりとした傾向は認められなかった。一方、インターネット利用経験の豊富な被験者は、正解率が若干高い傾向が見られた。

### 3.2 評価実験 2：翻訳された要約テキスト提示の効果

この評価実験 2 では、評価実験 1 で用いたコンテンツの主要キーワードの（翻訳）表示の代わりに、要約された翻訳テキストを提示して、探索テーマに該当する文書を被験者に推定させた。従って、対象コンテンツおよび探索テーマは評価実験 1 と同一の設定である。要約翻訳文の作成については、テキスト要約技術および機械翻訳技術の現状の到達レベルを考慮して、次のような 4 種類の提示条件を設けることとした。

#### 翻訳された要約テキストの提示条件

被験者が該当文書の選択のために参照する、各文書の要約テキストの提示条件は次の通りである。

- 条件 S1：手で翻訳された記事の最初の 75 文字 (2 行分に相当)
- 条件 S2：手で翻訳された記事を要約ソフト M で要約したもの
- 条件 S3：原文記事を要約ソフト P で要約したものに対応する人手翻訳の日本語
- 条件 S4：原文記事を要約 + 機械翻訳ソフト Q で翻訳した日本語

S1 から S4 までの条件で被験者に翻訳された要約テキストを提示し（提示する順序に応じて 8 種類のデータセットを用意）、被験者 60 名に対して 3.1 と同様な実験を行なった結果を、各条件毎のテキスト量とともに表 2 に示す。ただし、この要約テキスト表示の場合は、視認し易い状態で画面スクロール無しで表示することは不可能であることと、画面スクロール有りの条件で被験者のブラウザ操作を一定に制御することは困難であることから、計算機画面上での実験の代わりに、提示内容を A4 用紙に印刷した上で（S2～S4 では複数ページ）一斉実験を行なった。なお、本実験の被験者として、2 つの理工系学部の学生の協力を得た（M 大学 33 名、および T 大学 27 名）が、両集団の実験結果はほぼ同様であり、再現性についても問題がなかったものと考えられる。

表 2 要約テキスト提示方法の比較

要約種別	平均テキスト量	(比率)	平均正解率
S1	71 字	(9.4%)	78.3%
S2	88 字	(10.6%)	77.5%
S3	175 字	(23.3%)	72.5%
S4	227 字	(30.2%)	50.8%
元記事	752 字	(100.0%)	

#### 翻訳された要約テキストの表示内容による支援効果の相違

表 2 に示した実験結果から、S1, S2, S3 の条件が比較的高い平均正解率を示しており、あまり差が見られないのに対して、S4 の正解率の落ち込みが大きいことが分かる。この最大の原因は、S4 のみが機械翻訳の出力を用いていることと考えられる。情報量としては最も多い S4 の提示条件であるが、今回のタスク実行の助けとしては十分ではなかったことを示唆している。また、S1 のような単純な方法が、重要文を抽出する方式が採られていると思われる S2 とほぼ同様な結果となっていることは、重要箇所がテキストの冒頭部に現われやすい新聞記事を対象としたことによるものと推定される。

なお、課題 1 件当たりの平均所要時間は約 194 秒であり、日本語によるキーワード表示（評価実験 1 の条件 B および C）の場合の、1.8 倍ほどとなったことは（ただし、実験方法が異なるので参考値）、要約テキストの情報量が多く判断に時間を要したものと推定できる。さらに、被験者へのアンケート結果からは、コンテンツの要約情報の提示方法としては、今回の S1, S2 の条件程度のテキスト量が妥当とする意見が多いことが分かった。

### 3.3 評価実験 1・2 の比較

#### 条件毎の平均正解率の比較

主要なキーワードを利用者の言語（日本語）で提示した場合の文書選択の正解率と、翻訳された要約を提示した場合の正解率について比較を試みる。要約文の機械翻訳として生成されたテキストの表示（S4）を別にすれば、生成方法による若干の差異は認められるものの、両者はほぼ同レベルにあることが、表1および表2から導かれる。このことは、主要キーワードのみの表示でも、検索結果を「ふるいにかける」ような目的には、十分利用できる場面があることを示唆している。また、逆に言えば、（翻訳）要約の提示においては、単純な意味での情報量が多いにもかかわらず、（翻訳）キーワード表示と同等か、やや良好な程度のタスク達成率しか供し得ないことに加えて、翻訳の品質が鍵となることも示している。

#### 個別課題間の正解率の分布の比較

評価実験1のキーワード提示による該当文書推定の正解率は、8件の探索課題毎にかなりのばらつきを見せており、正解率の高い4課題（平均87%程度）と低い4課題（平均57%程度）に二分された。これに対し、評価実験2の要約提示では、際立った分布の違いは見られなかった（ほぼ60%～90%の間に分布）。しかし1題だけ、例外的に正解率の低いもの（課題3＝正解率25%）が認められた。この原因については、探索課題として示されたテーマ中に現われる語が、要約テキスト中に明示的に出現していないことによるものと考えている<sup>5</sup>。以上から、要約テキスト提示の方が、平均的に該当文書を探索する手助けとして有用であるが、より多くの時間を費やすのに対して、翻訳キーワードの表示は、効率良く検索結果を取捨選択するのに適していることが分かる。

## 4 まとめと今後の課題

本稿で報告した評価実験では、前述したように、検索結果を「ふるいにかける」ための助けとなる情報として、とりわけクロスリンガル情報検索においては、コンテンツの主要な部分の翻訳情報の提示が有効であることが示された。また、その形式としては、適切に翻訳された主要なキーワードの提示が、ほぼ適切に抽出された重要文からなる要約テキストに十分匹敵する場合もあることが明らかとなった。

また、キーワード提示、および要約表示には、各々の特性が存在し、おそらく的確な少量の要約テキストが、検索結果を最も効率良くブラウズする助けとなることが予想される（主要キーワードの表示のみではコンテンツのテーマや性格の推定は困難と思われる）。しかし、現状の技術レベルでは、要約・翻訳ともに直ちに精度を向上させることは難しいことから、コンテンツ中の重要なキーワードを抽出して適切な訳語に変換する方式は、膨大な外国語情報源の探索を行なう上で、重要な当面の研究課題であろう。

## 参考文献

- [1] TITAN, <http://titan.mcnet.ne.jp/>
- [2] AltaVista, <http://www.altavista.com/>
- [3] 青木 圭子, 松本 一則, 橋本 和夫, “大量文書向けのクラスタリング手法の評価”, 情報処理学会第56回全国大会, 1998.
- [4] 鈴木 雅実, 井ノ上 直己, 橋本 和夫: “多言語情報検索における利用者支援について - 主要キーワードの対訳付与に関する検討 -”, IPSJ-NL122-11, 1997.
- [5] 鈴木, 井ノ上, 橋本: “クロスリンガル情報検索結果の選択閲覧を支援する翻訳情報提示方法の比較”; 情報処理学会第57回全国大会; Vol.3, pp.169-170, 1998.
- [6] P. Resnik: “Evaluating Multilingual Gisting of Web Pages”, *AAAI Spring Symposium on Cross-Language Text and Speech Retrieval Electronic Working Notes*, 1997.

<sup>5</sup> 課題3のテーマは「インフラ建設における海外企業の協力期待」であり、要約テキスト中には道路・鉄道・港湾・電力等、いわゆる社会基盤としてのインフラに属する語が出現するものの、直接的な「インフラ」の語は含まれない。なお、実験1でも課題3の正解率は低く、約50%であった。このことは、「浅い」レベルのタスク遂行に留まる（あるいは単に語彙を知らない）被験者が多いことによるかもしれない。