テクニカルノート

末尾の名詞に着目したテレビニュース字幕の語義解析

井手一郎† 田中英彦†

増大する映像資源への自動的な索引付けを行う必要が高まっており，その際には付随する自然言語情報の利用が有効であると考えられる．本稿では，日本語ニュース番組への意味内容に立ち入った索引付けを実現するための基礎技術として，字幕の末尾の名詞に注目して意味属性を自動的に解析する手法の提案と，実際の番組へ適用した結果を報告する．既存の様々な概念辞書の分類体系は，必ずしもこのような目的に適っていないため，コーパス中から条件に適合する名詞を新たに抽出した．実際の番組に適用した結果，人物に関しては再現率94%，適合率78%，場所に関しては再現率82%，適合率60%の精度で語義解析ができた．

Semantic Analysis of Television News Captions Referring to Suffixes

ICHIRO IDE† and HIDEHIKO TANAKA†

Automatic indexing to image data is in strong demand. Utilizing accompanied natural language information is considered effective to accomplish the task. As a basis for semantic indexing, we propose an automatic semantic analysis method, which distinguishes semantic attributes of television news captions in Japanese, referring to suffixes. Classification in conventional concept based dictionaries are not fully applicable for such purpose, thus we extracted such suffixes from a corpus. The result was applied to actual news programs, and resulted in 94% recall, 78% precision for personal, and 82% recall, 60% precision for locational captions.

1. はじめに

テレビなどを通じて放送される映像資源の増加にともない，将来の検索や再利用を想定した自動的な索引付けへの要求が高まっている．自動的な索引付けを行うために，従来は主に画像認識的な手法が試みられていたが，近年，映像に付随する自然言語情報を利用する手法が脚光を浴びつつある．なかでも，字幕は元来キーワード的要素が濃厚な情報源であり，その効果的な利用が望まれる．

映像内容を反映した良い索引付けを行うためには，字幕から抽出するキーワードの語義を解析する必要がある．本稿では，字幕の末尾の名詞に着目した語義解析手法を提案し，実際の番組に適用した結果を示す．

なお，題材として，映像の検索や再利用の要求が最も高いと思われるテレビニュースに対象を絞って議論する．テレビニュースでは字幕が映像内容を端的に表し，また実際の番組を解析した結果，毎分平均4回の頻度で登場し，その利用は現実的である．

† 東京大学大学院工学系研究科
Graduate School of Engineering, The University of Tokyo

2. 字幕の特徴

2.1 文法的特徴

字幕には，一般の自然言語情報と異なる特徴があるため，必ずしも既存の処理手法をそのまま流用することはできない．主な特徴として以下のようなものがあげられる（文献6), 7) より一部引用）．

- 文章的な字幕
  - (1) 単文
  - (2) 主語の省略
  - (3) 体言止め
  - (4) 「へ」,「に」,「か」,「も」,「は」,「を」などの格助詞で終わる
- 非文章的な字幕
  - (5) 重要語句（主に名詞）の羅列
  - (6) 名詞のみからなる名詞句（地名や人名など

表1 テレビニュース中の字幕の種類
Table 1 Types of television news captions.

種類	割合	文法的特徴
(a) その他（タイトルなどを含む）	39%	(1)~(6)
(b) 場所・組織	28%	(5), (6)
(c) 映像中の人物	15%	(5), (6)
(d) 発言（要旨や翻訳を含む）	10%	(1), 口語文
(e) 日時	5%	(6)
(f) 放送上の技術的な情報	3%	(6)
(g) 映像内容の具体的な描写	2%	(2), (3), (5)

の固有名詞を含む）

このような文法的な特徴のため，既存の一般的な自然言語処理系でそのまま高度な処理を行うのは困難である．

2.2 内容的特徴

一方，字幕の内容的特徴については，**表1** に示すように分類できる．表中に示した「割合」は，実際の20分間のテレビニュース番組5本中に出現した587個の字幕を人手で分類したものに基づく数値である．また,「文法的特徴」は2.1節にあげたもののうちから該当する主なものを記した．

このうち，合わせて5割弱を占める（b）と（c）は，映像内容を忠実かつ簡潔に示すので，字幕をそのままキーワードとして利用可能である．しかし，映像内容を忠実に反映した索引付けを行う際は，キーワードの語義が明らかである必要があり，これらを用いるとしても,(1) 人物,(2) 場所,(3) その他，の判別を行う必要がある．したがって，次章以降では，このような判別を行うこと，すなわち字幕の語義解析を行う手法を紹介する．

3. 字幕の語義解析

従来より，一般的な文について文脈や格の解析を行って固有名詞の語義を解析したり[4]，言語情報のみでなく出現位置などの情報も用いて字幕の語義を解析する試み[7]が行われてきた．しかし，前者は文脈を利用するため，相互の直接的な関係が薄い字幕への適用は困難であり，後者は番組や放送局ごとのデザインや編集方針に依存する位置情報も用いるため，汎用性に欠ける場合があるという問題点がある．これらの問題点を考慮して，各字幕単独での解析を目指す．

たとえば,「橋本」だけでは，この名詞が人物を示すのか場所を示すのかは，人間でも分からないが,「橋本-首相」ならば人物であり,「橋本-駅」ならば場所であることが分かる．このように，一般に日本語においては，ある名詞句の末尾の名詞に着目することにより，その名詞句全体の語義を解析できることが多い☆．この性質を利用して，その多くが名詞句である字幕の語義を，末尾の名詞に着目して解析する．

4. 語義を決定する名詞の収集

まず，どのような名詞が末尾に存在すれば名詞句全体が（1）人物を示すのか,(2) 場所を示すのか，を知る必要がある．種々の概念分類辞書が存在するものの，語の用法を含めた語義による分類体系を有するものではないため，天下り的に上記の各分類に該当する分類項目を決定しにくい．そこで，RWCテキストデータベース[5]中の，RWC-DB-TEXT-95-2テキストコーパスを利用して，まずそのような名詞を収集した．

同コーパスは1994年1年間の毎日新聞の記事から27,000文を選択して，人手によって形態素解析を施したものである．文章の出自が新聞記事であることから，ニュース番組の字幕に登場するものと近い語彙が多く含まれていると考えられるとともに，人手により形態素解析されていることから，正しい解析結果であると見なせ，ここでの収集目的に適う情報源である．

しかし，これにより収集される語彙のみでは，一般的に字幕を解析するには不十分なため，収集した語を含む分類語彙表[2]の分類段落☆☆中の語もあわせて収集して，語彙の拡張をはかった．この拡張は，分類語彙表の分類体系そのものは本研究の目的に沿わないものの，分類段落中の各語の語義は類似していると仮定したうえでの処理である．

4.1 末尾に存在して人物を示す名詞

ここでは,「さん」や「俳優」のように，名詞句の末尾，あるいは単独で存在して，人物を示す名詞を収集する．

4.1.1 収集基準

以下の手順でRWC-DB-TEXT-95-2から名詞を収集した．

(1) 人のみに付きうる接尾辞「ら」,「たち」を探す
(2) その前の語が普通名詞か接尾辞ならば，その語を収集する

たとえば,「橋本（固有名詞）首相（名詞）ら（接尾）」からは「首相（名詞）」が収集される．なお，手順(2)において固有名詞は収集しないようにしている．

4.1.2 収集結果

収集時に利用した「ら」と「たち」を含めて，257語の名詞が収集された．このうち，190語は分類語彙表に記載のある語であり，残りの67語は記載のない

☆ 欧米の言語では，名詞句の先頭の名詞に着目する必要があることも多々あり，一概にこのようにはいえない．
☆☆ 分類項目中の細分類．10程度の語を含む．

ものであった．前者の語群が含まれる68分類項目中の141分類段落には，合わせて1,706語が属しており，最終的に後者と合わせた1,773語が収集された．

4.2 末尾に存在して場所を示す名詞

次に，「駅」や「台所」のように，名詞句の末尾，あるいは単独で存在して，場所を示す名詞を収集する．

4.2.1 収集基準

以下の手順でRWC-DB-TEXT-95-2から名詞を収集した．

(1) 地域を示す固有名詞☆を探す
(2) それに続く名詞を後方へたどる
(3) 名詞が途切れて，次に場所と関係することのある格助詞「から」，「で」，「へ」，「より」，「にて」が存在すれば，(2)の最後の名詞を収集する

たとえば，「横浜（地域）市（接尾）金沢（地域）区（接尾）で（格助詞）」からは「区（接尾）」が収集される．

4.2.2 収集結果

383語の名詞が収集された．しかし4.2.1項に示した基準では，これらに「アメリカ（地域）大統領（名詞）より（格助詞）」から収集される「大統領（名詞）」のように，人を示す名詞と重複するものが含まれているため，4.1.2項で収集した名詞を除去した．その結果340語が残り，そのうち281語は分類語彙表中に記載のある語であり，残りの59語は記載のないものであった．前者の語群に含まれる172分類項目中の279分類段落には，合わせて3,001語が属しており，最終的に後者と合わせた3,060語が収集された．

5. 実験：実際のニュース番組字幕への適用

前章での収集結果をもとに，以下の手順で実際のニュース番組中の字幕の解析を行った．解析を行ったのは，20分間の昼の定時ニュース5日分中の587個の字幕である．

5.1 実験手順

(1) 日本語形態素解析ツールJUMAN[3]を用いて形態素解析を行う．字幕がJUMANの辞書中に登録されている固有名詞（主に人名と地名）であれば，この段階で語義が決まる．
(2) 字幕の末尾の形態素が普通名詞ならば，4.1.2項および4.2.2項で得られた辞書と照合して，「人物を示す字幕である」あるいは「場所を示す字幕である」と判断する．

☆ 形態素解析の結果としてあらかじめ付与されている．

5.2 実験結果

以上のような手順で実験を行った結果，**表2**，**表3**に示すような結果が得られた．真の正解は第三者が与えたもので，解析結果をこれと比較して評価した．

表中，「本手法単独」とは5.1節の手順(2)単独により得られた結果とその評価であり，「JUMAN単独」とは手順(1)にあるように，JUMANによる解析の時点で語義が判明したもの（固有名詞）を示し，両者を合わせた総合的な結果とその評価を「総合」に示す．

5.3 考察

誤検出と検出洩れの原因としては，以下の複数のものが考えられる．

- **誤検出**
  - (1) **固有名詞の誤判定**（JUMAN）
    固有名詞には，人名にも地名にもなりうるものが存在．
    例）「土浦」（場所を人物と判断）
  - (2) **収集した名詞が不適切**（本手法）
    コーパスからの収集時に混入した不適切な名詞による影響や，分類語彙表の分類方針と本研究の分類目的との不一致による．「場所」の場合に顕著．
    後者の例）「会議」（場所）と「異論」（その他）が同一分類段落に属す
  - (3) **語義の多様性**（本手法）
    収集した名詞の中に，語尾に存在して「人物」，「場所」，「その他」の複数を示しうるものが存在．
    例）「写真-家」（人物）と「田舎-家」（場所）

これらのうち(1)と(3)は，多義性という，語が

表2 字幕解析結果（人物）
Table 2 Result of caption analysis (personal).

	本手法単独	JUMAN単独	総合
正答数	67個	15個	82個
回答数	81個	24個	105個
真の正解数	—	—	87個
適合率	82.7%	62.5%	**78.1%**
再現率	(77.0%)	(17.2%)	**94.3%**

表3 字幕解析結果（場所）
Table 3 Result of caption analysis (locational).

	本手法単独	JUMAN単独	総合
正答数	84個	49個	133個
回答数	174個	49個	223個
真の正解数	—	—	162個
適合率	48.2%	100.0%	**59.6%**
再現率	(51.9%)	(30.2%)	**82.1%**

持つ本質的な性質によるため改善は困難だが，(2)は収集基準の改良による改善が見込まれる．

- **検出洩れ**
  - (4) **固有名詞の語彙不足**（JUMAN）
    例)「卓」(人物),「オタワ」(場所)
  - (5) **収集した名詞の語彙不足**（本手法）
    例)「市議」(人物),「銀行」(場所)
  - (6) **解析の前処理の不足**
    字幕から余分な情報を除去するための前処理が不十分で解析に失敗．

これらのうち，(4)はJUMAN固有名詞辞書の改良を待たねばならないが，(5)はより大規模な解析済みテキストコーパスの利用により，(6)は解析手順の改良により改善が見込まれる．

6. おわりに

本稿では，テレビニュース中の字幕の語義解析手法と評価実験の結果を示した．その過程で，字幕の分析を行い，多くの字幕が比較的容易にキーワードとして利用可能であることを示した．また，既存の概念分類体系とは異なる方針で語義分類を行うため，語義解析上の手がかりとなる名詞の収集を行った．

その結果，本格的な実用化には一段の精度向上が望まれるものの，名詞句の末尾の名詞に着目した，人物と場所に関する語義分類手法の有効性が示された．

なお，本研究はニュース映像データベースの自動的索引付けに字幕を利用する[1)]ための基礎的な研究であり，現在，索引付けシステムでの試用を行っているところである．

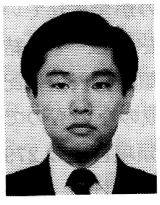
謝辞 JUMANは京都大学長尾研究室と奈良先端科学技術大学院大学松本研究室にて開発されたフリーソフトウェアであり，分類語彙表は国立国語研究所の成果物である．RWCテキストデータベースは技術研究組合新情報処理開発機構の成果物であり，同機構の許可の下に利用した．

本研究の遂行にあたって，角田達彦博士には全般にわたる適切な助言を，永松健司博士には自然言語処理に関する思慮深い助言を，平博司氏には字幕データ収集に協力していただいたことを深く感謝する．

参考文献

1) 井手一郎，山本晃司，田中英彦：ショットの分類に基づく映像データへの自動的索引付け，第56回情報処理学会全国大会論文集(2)，pp.263–264 (1998).
2) 国立国語研究所：国立国語研究所言語処理データ集5分類語彙表［フロッピーディスク版］，秀英出版 (1993).
3) 松本裕治，黒橋禎夫，山地　治，妙木　裕，長尾　眞：日本語形態素解析システムJUMAN，version 3.2 (1997). http://www-lab25.kuee.kyoto-u.ac.jp/nl-resource/juman.html よりダウンロード可能．
4) 那須川哲哉：文脈情報を利用したキーワード語義決定，1997年度人工知能学会全国大会（第11回）論文集，pp.348–349 (1997).
5) 技術研究組合新情報処理開発機構（RWCP）：RWCテキストコーパス (1996).
6) 若尾孝博，江原暉将，白井克彦：テレビニュース番組の字幕に見られる要約の手法，情報処理学会技術研究報告自然言語処理（NL），Vol.97, No.109, pp.83–89 (1997).
7) 渡辺靖彦，岡田至弘，長尾　眞：TVニュースで用いられるテロップの意味解析，情報処理学会技術研究報告自然言語処理（NL），Vol.96, No.89, pp.107–114 (1996).

(平成10年1月29日受付)
(平成10年6月 5 日採録)



井手　一郎（学生会員）
昭和47年生．平成6年東京大学工学部電子工学科卒業．平成8年同大学院工学系研究科情報工学専攻修士課程修了，修士（工学）．現在同研究科電気工学専攻博士課程在学中．自然言語処理，マルチメディア統合処理に興味を持っている．平成7年情報処理学会第51回全国大会奨励賞受賞．人工知能学会，電子情報通信学会各学生会員．

田中　英彦（正会員）
昭和18年生．昭和40年東京大学工学部電子工学科卒業．昭和45年同大学院博士課程修了．工学博士．同年同大学工学部講師．昭和46年同助教授．昭和62年より同教授，現在同大学院工学系研究科教授．この間昭和53～54年ニューヨーク市立大学客員教授．計算機アーキテクチャ，並列処理，自然言語処理，メディア処理，分散処理，CAD等の研究に興味を持っている．著書「非ノイマンコンピュータ」,「情報通信システム」，共著書「計算機アーキテクチャ」,「VLSIコンピュータI，II」,「ソフトウェア指向アーキテクチャ」．人工知能学会，日本ソフトウェア科学会，IEEE，ACM各会員．