

## マルチモーダル知識コンテンツを利用した英語コミュニケーションの学習支援

木 寺 敦 則<sup>†</sup> 桐 山 伸 也<sup>††</sup>  
堀 内 裕 晃<sup>††</sup> 竹 林 洋 一<sup>††</sup>

英語でのコミュニケーションに役立つ実用的な知識を、リアリティのある視聴覚情報と奥の深い言語情報からなるマルチモーダルナレッジとして提供することにより、「読む」、「書く」、「聞く」、「話す」の技能をバランスよく修得できる新たな英語学習システムを提案する。本研究は、英語教育・自然言語処理・ヒューマンインタフェースの研究者による文工融合の共同プロジェクトとして推進されているものであり、異分野間のシナジー効果によって、生きた英語の深い知識を自然言語処理によって概念表現に変換し、魅力的なマルチモーダル知識コンテンツを生成することに成功した。

### Multimodal Knowledge Contents Service for Improvement of Knowledge and Communicative Skills in English

ATSUNORI KIDERA,<sup>†</sup> SHINYA KIRIYAMA,<sup>††</sup> HIROAKI HORIUCHI<sup>††</sup>  
and YOICHI TAKEBAYASHI<sup>††</sup>

In this paper, we discuss the use of Multimodal Knowledge, which consists of significant linguistic information and realistic audio-visual information, for improvement of communicative knowledge and skills in English. We aim to develop a new English-Learning environment which enables learners to develop skills of "reading", "writing", "listening", and "communication" comprehensively. This research has started as a collaborative project of English-Learning, Natural Language Processing, and Human-Interface, and by transforming significant knowledge of live English into conceptual framework for communication, we succeeded in developing attractive Multimodal Knowledge Contents.

#### 1. はじめに

グローバル化の急速な進展により、英語によるコミュニケーションスキルが重要視されてきた。英語でのコミュニケーションに役立つ実用的な知識を、リアリティのある視聴覚情報と奥の深い言語情報からなるマルチモーダルナレッジとして提供し、「読む」、「書く」、「聞く」、「話す」技能をバランスよく修得できる新たな英語学習システムを提案する [1]。

筆者らのうちの一人は英語教育を専門としており、日本人の英語コミュニケーションスキル向上に役立つ知識 [5][6] を、英語教育の現場で長年にわたって蓄積してきた実績を持つ。また、他の一人の著者は、自然言語処理・音声情報処理によって自然言語 / 音声データと概念表現の関係を解明し、よりインテリジェントな受け答えができる対話システム

の開発に従事してきた経緯を持つ [3][4]。さらに別の著者の一人は、マルチモーダルナレッジサービス [2] の開発に従事しており、あらゆるユーザのあらゆる状況に応じて適切なナレッジ・情報を提供するシステムの実現を目標としている。

すなわち本研究は、英語教育・自然言語処理・ヒューマンインタフェースの研究者による文工融合の共同プロジェクトとして推進されているものであり、異分野間のシナジー効果によって、生きた英語の深い知識を自然言語処理によって概念表現に変換し、魅力的なマルチモーダル知識コンテンツを生成することに成功した。

#### 2. マルチモーダル知識コンテンツを利用した英語学習

##### 2.1 映像・音声・言語知識の統合的な利用

英語でのコミュニケーション能力を上達させるためには、留学やビジネス体験など、英語圏の現場でネイティブとの Face-to-face の会話を通して学ぶことが最も有効とされているが、現地へ出向くコス

<sup>†</sup> 静岡大学大学院情報学研究科

Graduate School of Information, Shizuoka University

<sup>††</sup> 静岡大学情報学部

Faculty of Information, Shizuoka University

トが大きいなどの理由により、現実にはビデオ教材や英会話教室などが多く利用されている。しかしながら、実用的な英語を「読む」・「書く」・「聞く」・「話す」技能を総合的に磨くという観点からは、ビデオ教材等では不十分である。

筆者らの一人は、リアリティ溢れる画像・音声情報と深遠な言語情報からなる実用的で役立つマルチモーダルナレッジを蓄積し、ユビキタス環境でオンデマンドにユーザに提供することを念頭においた、マルチモーダル知識コンテンツ配信システムの開発に携わってきた [2]。このシステムは、人間のリアルな生活のサポートを指向したものであり、マルチモーダルナレッジを簡単に作成できるオーサリングツールにより、誰もが手軽に新たなコンテンツを創出できる枠組みを提供している。

本論文では、この枠組みを利用した英語学習支援システムを提案する。すなわち、ネイティブへのインタビューや講演を撮影したリアルなデータから、英語教育の専門家の手で創出される深い英語の知識を抽出・蓄積していくことでコンテンツを増やし、蓄積した英語の知識ベースから関連する知識を、言語情報のみならず話者の表情や反応・ジェスチャーといった非言語情報をも含んだマルチモーダルナレッジとして学習者に提供するシステムを提案する (図 1)。

## 2.2 マルチモーダル知識ベースの構築

効率的な英語知識の蓄積と、知識コンテンツ間での相互参照を実現するために、抽出した英語知識に構造を与え、マルチモーダル知識ベースを構築する。従来の知識ベースの研究に、一般的な知識を集めて大規模知識ベースの構築を目指した研究 [7]

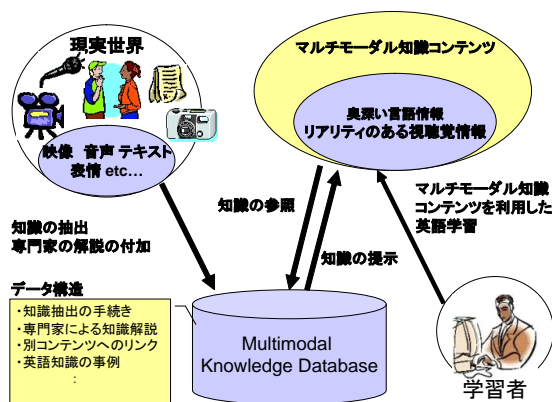


図 1 映像・音声・テキストの統合的利用

や、辞書に概念構造を持たせ、多角的・階層的な単語の検索を実現した研究 [8] などがあるが、それぞれの長所に注目して知識ベースを設計する。

取り扱う知識は、英語教育の専門家によって選別された実用的な知識だけを対象とした。また、文法や単語レベルの知識だけでなく、日本人が苦手とする対話技術や会話の駆け引きなど、コミュニケーションに関する幅広い知識を取り扱った。

知識ベースは幅広い知識を効率よく管理するためにオブジェクト指向データベースの形態をとり、抽出の手続き、知識解説、関連する知識コンテンツへのリンク、抽出した表現事例をひとまとめに管理する。

表現の用途や文法的な特徴の違いに基づいて英語知識を抽出し、知識の分類・体系化をおこなった。

例えば疑問表現のクラスに関しては、用途の違いによって「質問」「依頼」「勧誘」「許可」「確認」... などのサブクラスに分類し、更に質問の型に基づいて「5W1H型」「付加疑問文」「閉質問 (Yes / No で回答できる質問)」... などに細分化した。

「能動態 / 受動態」や「直接性 / 間接性」などのように、言語全般に共通する特徴についてはそれぞれのクラスが持つ属性として定義した。これらの特徴は話者の意図やニュアンスを強く反映するものであり、話し手の心情や発言の背景を理解するための大きな指標となる。

知識の生成、及び分類手法については 3 章で具体的に述べる。

## 2.3 マルチモーダル知識コンテンツの作成

英語教育の専門家が生成した英語知識を MPEG7 の枠組みに従って構造化し、映像・音声・テキストを用いて提示するマルチモーダル知識コンテンツを作成した。知識コンテンツの内部構造を図 2 に示す。また、作成したマルチモーダル知識コンテンツの画面を図 3 に示す。

コンテンツ作成の手順は次のようになる。まず、英会話の映像から意味のある部分だけを切り出し、それぞれに含まれる知識情報をメタデータとして登録する。次に、登録したメタデータと英語の知識ベースとの間でマッチングをとり、関連情報として提示する。同時に、コンテンツの中で使われている新たな英語の知識をテンプレートとして登録し、知識ベースの充実を図る。

### 3. コミュニケーションスキルの獲得に必要な英語知識

学習者にとっては、英語教育の専門家から指導を受けることが上達への最たる近道である。この見地から、我々は英語教育を専門とする一人の著者の監修によって、日本人に教示すべき英語知識を詳細な解説とともに具体化した。また、自然言語処理技術を用いてこれらの知識を蓄積し、新たな素材から抽出する手法を提案するとともに、抽出した知識を場面に応じて適切な表現に変換して学習者に提示する方策を検討した。

知識生成の流れを図4に示す。

#### 3.1 英語教育の専門家による知識の生成

コンテンツを作成するための素材として、英語圏のネイティブに対する対話形式のインタビューを収録した。インタビューで交わされた会話を英語教育

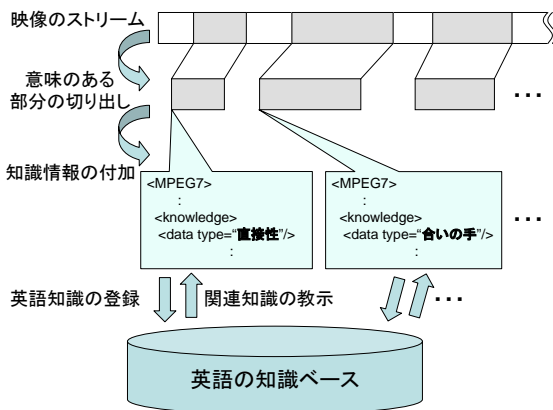


図2 コンテンツへのメタデータの登録



図3 マルチモーダル知識コンテンツの画面

を専門とする著者の一人が吟味した上で、日本人が不得手な表現や陥りやすい間違いなどを抽出し、解説を付与して英語知識を生成した。今回の収録素材から生成された代表的な英語知識を以下に挙げる。

- 「Joke Detection」  
面白いと感じるコンテンツが学べることを念頭におき、英語で発せられたジョークを検出して提示する。
- ネイティブが好んで用いる「自然な表現」  
“Why did you ~” という表現よりも “What brought you ~” の表現がネイティブが好む表現に近い。
- 話者の意図や感情が表出される「直接性/間接性」の違い  
“I shot the bird.” では「撃ち落した」ことになり、結果についての言及がある。  
“I shot at the bird.” では「狙って撃った」となり、弾が当たったかどうかの言及はない。
- コミュニケーションの潤滑油としての「合いの手」  
1. 相手の発話の最後を繰り返す部分  
2. 発言の中で話者が主張したい部分を別の単語を用いて言い換えている部分  
3. 話題を具体化する質問が使われている部分

#### 3.2 英語知識の蓄積・抽出手法

「Joke Detection」については、音声情報処理技術を用いて抽出をおこなった。具体的には、音声ストリームから基本周波数・パワーなどの韻律的特徴を表す情報を抽出し、笑い声の起きている部分や、声の高さが急激に変化している部分を検出した。

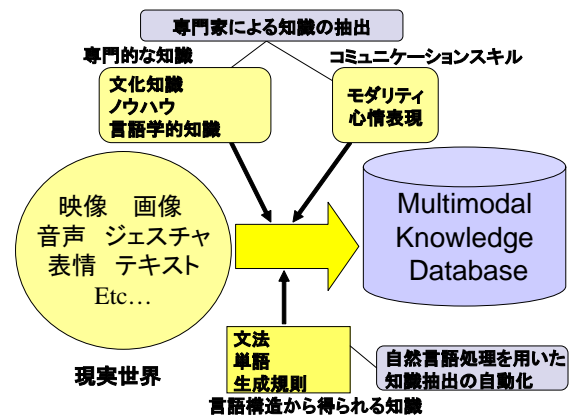


図4 英語知識の生成

その他のテキスト主体の英語知識を表現する方法として、具体的な単語の並びによる表現（以降、事例表現と呼ぶ）と、品詞や句などの構文情報の並びによる表現（以降、構文表現と呼ぶ）の2通りを用いることにした。3.1で挙げた知識の表現に際しては、「自然な表現」と「直接性/間接性」については事例表現を、「合いの手」については構文表現をそれぞれ用いた。

事例表現については、“What brought NPL”, “Why did NPL” や, “shot NPL”, “shot at NPL” (NPLは名詞句を表す) というレベルで表現を定義して蓄積する。構文解析ソフトを利用して抽出対象のテキストに構文解析を施し、表現にマッチする部分を抽出した。

構文表現については、「相手の発話の最後の句を繰り返している部分」のように抽象的な表現によって定義する。表現別に異なるアルゴリズムを蓄積していき、これを用いて抽出をおこなう。3.1節での「合いの手」の1.については、

- (1) 発話者が交代している発話境界を検出
- (2) 前出文の末尾の句と後続文の先頭の句を抽出
- (3) 2つの句が一致した場合、その部分を「合いの手」の例と判断

というアルゴリズムを作成することにより、今回収録した素材に現れる1.の知識を表出している箇所全てを正しく抽出することができた。

また「合いの手」の2.については、上記アルゴリズムの(2)で抽出した2句中の類義語の有無をソーラス辞書で調べることにより、2つの句が言い換えの関係にあるか否かを判定できる。3.についても、その質問が“how”や“what”などの疑問詞で始まるか否かを判定することで抽出できる。

### 3.3 学習者への知識の提示方法

抽出した知識に付属している解説が学習者に提示される。英語教育専門家の手による解説は段階的構造を持っており、学習者のレベルに合わせて提示する内容を変化させる制御が可能である。例えば、図3の例では“What brought you to Japan?”の方がネイティブに好まれる事実のみを提示しているが、より高いレベルの学習者に対しては、これがmake(使役動詞)+人+動詞の使役構文を拡張したbring(使役の意味を持つ動詞)+人+場所という英語特有の非人称主語構文の例であることまでが提示される。

同時に、関連知識を含むコンテンツと相互参照をおこなうことで、単一のコンテンツによる学習に

止まらず、数多くのコンテンツを参照しての英語学習が可能である(図5)。これにより、学習者に幅広い分野のコンテンツを提示し、さらに有意義な知識を効率よく提供することができる。

また、主に事例表現について、知識が蓄積される種類が増えるに従って、解説を各々別個に用意する手間が問題になると予想される。この問題も自然言語処理技術を駆使することで一部対処可能である。すなわち、

- (I) 蓄積した事例表現を「自然な表現」・「直接性」・「間接性」といったカテゴリに分類
- (II) 各カテゴリ内で共通に用いる解説文のテンプレートを用意(テンプレート中の事例表現によって変動する部分は概念表現で記述)
- (III) 実際の事例表現に応じて、テンプレートの概念表現部分を具体的な単語に置換して解説文を生成

という流れで、各カテゴリ内で共通の解説文が自動的に生成されるため、解説文作成の負荷が軽減される。

## 4. 評価実験

実験では、

- 1) 生成した英語の深い知識の学習者にとっての有用性
- 2) マルチモーダルコンテンツで学習することの有効性

を評価した。

### 4.1 実験方法

ネイティブに対する5分間のインタビュー映像を利用し、

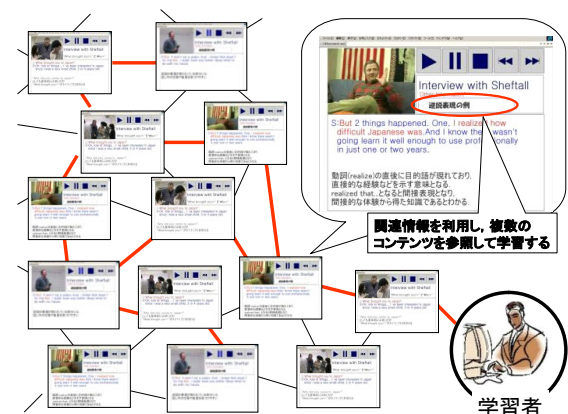


図5 マルチモーダル知識コンテンツによる英語学習

- A. 会話の内容のみをテキストで提示するコンテンツ
- B. 会話の内容と英語知識の解説をテキストで提示するコンテンツ
- C. 会話の内容のみをマルチモーダルに提示するコンテンツ
- D. 会話の内容と英語知識の解説をマルチモーダルに提示するコンテンツ

の4通りを作成した。

本学の学部生・大学院学生16名(男性13名,女性3名)に対して,アンケート及びコメント自由記述による主観評価を実施した。上記のAからDの4コンテンツを提示し,1)についてはAとB,及びCとD,2)についてはAとC,及びBとDをそれぞれ相対評価させた。コンテンツ提示の順番はランダムに設定し,各人が異なる順番でそれぞれのコンテンツを参照するよう指示した。

アンケートによる評価項目には次の4項目を設定した。

- a) 学習効果が高いのはどちらか
  - b) インタフェースが優れているのはどちらか
  - c) 学習意欲が湧くのはどちらか
  - d) 使い勝手が良いのはどちらか
- これらの評価項目に対し,被験者に
- (i) 知識の生成をおこなったコンテンツ,またはマルチモーダルコンテンツが優位である
  - (ii) 知識の生成をおこなわないコンテンツ,またはテキストコンテンツが優位である
  - (iii) 両者で同等である

の3段階による判定を求めた。

#### 4.2 実験結果

アンケートを1)の評価基準に従って集計した結果を表1に,2)の評価基準に従って集計した結果を表2に示す。表に示される数字は,それぞれの判定項目に対する回答者の人数である。

評価基準に対して厳正な結果を求めるため,1)については提示方式が同じ条件での知識解説付のコンテンツと会話内容だけのコンテンツの比較を,2)については提示内容が同じ条件でのテキストコンテンツとマルチモーダルコンテンツの比較を対象に集計をおこなった。評価基準に対して不適切なコンテンツの組み合わせについては評価の対象から除外した。

表1,表2で示されているように,マルチモーダルコンテンツ,及び知識解説を含むコンテンツともに,通常の学習コンテンツと比べて「優位である」

表1 生成した英語知識の有用性についての集計結果

評価項目	1-a	1-b	1-c
a) 学習効果が高いのはどちらか	24	1	7
b) インタフェースが優れているのはどちらか	26	3	3
c) 学習意欲が湧くのはどちらか	27	-	5
d) 使い勝手が良いのはどちらか	21	2	9

AとB,及びCとDの比較を集計の対象とした。

判定は以下の1-a~1-cの3段階によるものである。

1-a: 知識の生成をおこなったコンテンツが優位である

1-b: 知識の生成をおこなわないコンテンツが優位である

1-c: 両者で同等である

表2 マルチモーダルコンテンツの有効性についての集計結果

評価項目	2-a	2-b	2-c
a) 学習効果が高いのはどちらか	24	5	3
b) インタフェースが優れているのはどちらか	25	1	6
c) 学習意欲が湧くのはどちらか	29	3	-
d) 使い勝手が良いのはどちらか	25	5	2

AとC,及びBとDの比較を集計の対象とした。

判定は以下の2-a~2-cの3段階によるものである。

2-a: マルチモーダルコンテンツが優位である

2-b: テキストコンテンツが優位である

2-c: 両者で同等である

と回答した被験者が多かった。特に評価項目c)に関しては,評価基準1),2)ともに最も高い割合で「優位である」との回答が寄せられており,英語の深い知識をマルチモーダルに提示するコンテンツによって,学習者の意欲を大いに引き出せることが示された。

評価基準1)の評価項目d)に対する結果について,他の評価項目と比べて「優位である」と回答した被験者の割合が高くなかったのは,コンテンツ中の知識情報の密度がやや高く,一度に提示される情報量が多過ぎると感じた被験者が存在したことに起因すると考えられる。この点は今後対応策を熟慮する必要がある。

自由記述による主観評価においては,被験者から「テキストだけでなく,映像や音声を用いてマルチモーダルに提示するコンテンツは大変分かりやすかった」、「感情の表現方法など,実用的な解説をしてくれるのは大変役立つ」といったコメントが多数寄せられた。これは,生きた英語の知識をリアルティ溢れるマルチモーダルコンテンツで提示することの有効性の証である。

## 5. 考察

実験結果より,英語教育におけるマルチモーダル知識コンテンツの有効性が確認できた。

従来の英語学習であまり見られない,ジョークや

感情の表現といった新しい軸を取り入れることで、従来の画一的な学習教材と比べて柔軟で魅力的な知識コンテンツを学習者に提示することができる。

また、最初から英語教育用に学習コンテンツを構築するのではなく、身近な映像を利用して英語学習に応用できるメリットは大きい。映像に関する肖像権などの問題については、今後検討を進めていく。

将来的に、マルチモーダル知識コンテンツは情報提示技術の主流になると予想され、コンテンツが増えれば増えるほど英語学習に利用できる知識も豊富になり、それに伴って英語の知識ベースも充実していく。

今後、より一層効果的な英語学習システムを構築するために、下記の項目について重点的に検討を進める。

- 実用的な英語知識の抽出手法の決定  
対象とする知識を抽出の手続きとともに決定し、随時知識ベースに追加していくことで、取り扱う英語知識を充実させると同時に新しい英語学習の軸を追求する。
- 知識の構造化とマルチモーダル知識ベースの充実  
マルチモーダル知識コンテンツとの連携を密にし、関連知識との相互参照をより効果的におこなうために、英語知識を詳細に分類・体系化し、適切な知識表現のための知識構造を模索していく。
- マルチモーダル知識コンテンツによる効果的な知識表現手法の確立  
英語知識を適切に学習者に提示するための表現技法を迫るとともに、提示する知識の質や量の調節、ユーザの操作性の向上についても検討を進める。

## 6. おわりに

本稿では英語のコミュニケーションスキルの獲得のためにマルチモーダル知識コンテンツと英語教育専門家の知識の利用を実験し、その有効性を実証した。

今後は知識の抽出・蓄積・提示手法の効率化を進めると同時に、音声ストリームにおける韻律情報の分析に基づいて感情やニュアンスを抽出し、より高度な知識表現を扱う枠組を検討する。

## 参 考 文 献

- [1] 木寺敦則, 桐山伸也, 堀内裕晃, 竹林洋一: “英語のコミュニケーションスキル獲得のためのマルチモーダル知識コンテンツの利用,” FIT2003, (2003, 発表予定)
- [2] 竹林洋一, 鈴木優, 岐津俊樹, 浦田耕二, 網淳子, 宮澤隆幸, 金沢博史: “ユビキタス環境における音声対話システム MKIDS の開発,” 日本音響学会 2002 年春季研究発表会講演論文集, 2-5-13, pp.99-100, (2002)
- [3] 桐山伸也, 広瀬啓吉, 峯松信明: “話題知識を導入した文献検索音声対話システム,” 電子情報通信学会論文誌, Vol.J85-D- , No.5, pp.863-876, (2002)
- [4] 桐山伸也, 広瀬啓吉: “応答生成に着目した学術文献検索音声対話システムの構築とその評価,” 電子情報通信学会論文誌, Vol.J83-D- , No.11, pp.2318-2329, (2000)
- [5] 堀内裕晃: “譲渡不可能所有関係に関する一考察,” JELS (日本英語学会研究発表論文集), 15, pp.31-40, (1998)
- [6] 堀内裕晃: “語彙情報と譲渡不可能所有構文,” Ars Linguistica(Linguistic Society of Shizuoka), 5, pp.64-80, (1998)
- [7] Douglas B. Lenat, R. V. Guha: “Building Large Knowledge-Based Systems,” Addison-Wesley, Reading, MA, (1990)
- [8] George A. Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine J. Miller: “Introduction to WordNet: An on-line lexical database,” Journal of Lexicography, 3(4), pp235-244, (1990)