

介護場面の日本語学習に向けた例文データベースの構築支援技術 Support techniques for developing an example sentence database in the nursing-care situations

竹内 和広

Kazuhiro Takeuchi

大阪電気通信大学 情報通信工学部, 寝屋川市初町 18-8

Osaka Electro-Communication University, 18-8 Hatsucho, Neyagawa, Osaka

あらまし:近年, 医療福祉の分野に就業する外国人従事者が増えている. そのような従事者がサービスの受け手と会話をする場面では, 既存の教科書では経験のしたことのない言語表現に遭遇するため, 場面に対応した例文集の作成が有益である. 本稿では, そのような例文を収集し, データベース化する上での支援技術を紹介する. 具体的には, 登録する例文が教科書的な例文と相違する点を多面的に分析する技術と, その前提となる発話者の心理的・身体的な状態や状況の中間表現に関する研究について報告する.

Summary: Recently, it is said in Japan that more and more foreign workers are needed in the field of health and welfare. These workers on the job face various Japanese expressions that they didn't learn in the ordinary text books for Japanese language learners. In this paper, we discuss an efficient way to construct sentence database for the learners in specific purpose. In particular, we introduce two support techniques for constructing an intermediate semantic representation used in various methods to evaluate sentences from several points of view.

キーワード: 日本語学習, 学習用例文, 意味表現

Keywords: Japanese Language Learning, Example Sentences of Language Learning, Intermediate expression of meaning

1. はじめに

本研究では, コンピュータ支援により医療福祉分野の日本語を学習・運用する際に役立つ対話例集の作成・収集について検討する. 具体的には, 登録しようとする対話例文の適切性, 非文法性を計算文法, および, その他の意味処理機構に基づいて評価し, 外国人医療福祉従事者が日本語を学習する上で役立つ, かつ実際の医療福祉コミュニケーションを支援するシステム構築にも役立つような言語資源の構築を支援する技術を考える.

医療福祉分野に従事する外国人労働者は, 既に看護師や介護士の資格を取得し, 一般的な日本語能力や専門能力については一定の水準に達してから現場に入るとされている. しかし, このような施設や拠点で定常的に正規の日本語教員を十分な人数採用するこ

とは難しく, 日本語教育のボランティアや現場の関係者が, 上記のような外国人労働者(以下, 単に学習者と呼ぶ)への日常的な日本語教育の任を担うことが多い. そのため, 現在, 様々な機関や施設が, 学習者向け医療福祉分野日本語例文集を提供するようになって来たが, そのような例文に関して, 適切な日本語教師のアドバイスが得がたいことが問題となる.

そのような背景から, 介護場面において想定される, 未知の例文を日本語学習者に提示する上で必要となる日本語学習・運用向けの補足情報を提示する図1のような概要のシステムを構築している. このようなシステムは専門日本語教育の例文データベースを構築する上で有益であることが期待されるが, 既存の言語処理技術を単純に応用するだけでなく, 当該の専門分野のドメイン知識や辞書を整備することが必須である.

また、言語教育の専門知識も取り入れる必要がある。本稿では、そのようなシステムを構築する上で、言語教育支援システムの一般性を視座に、我々がシステム構築に際して使用、あるいは新規に研究開発した計算機技術について報告する。

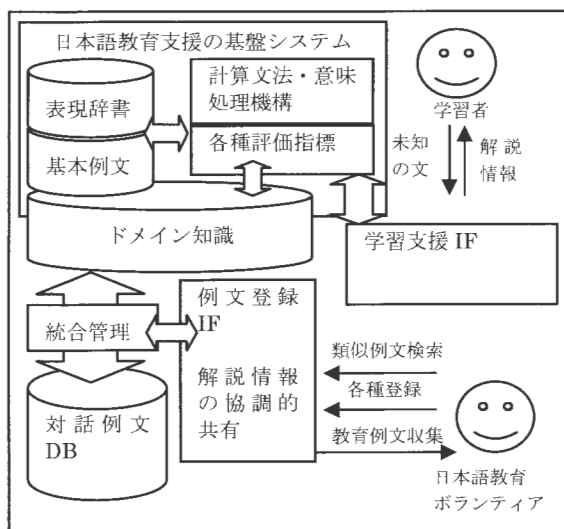


図1. 専門日本語学習支援システムの概要図

2. 計算機による言語学習支援

2.1 専門日本語教育支援の方向性

本研究では専門日本語教育をとりまく状況に関して、主に図1中にあるような2つの方向性の支援を考えている。

1つは現場で学習者の日本語教育に携わる関係者をサポートする方向性である。そのような関係者は必ずしも日本語教育を生業としているわけではないため、教材作成や日本語教育法の熟達に十分な時間をかけることはできない。この環境に対して、教材作成や日本語の教育上のポイントを低コストで実現でき、他の日本語教育ボランティアと情報を協調的に蓄積し共有する環境を提供する。

他方は、学習者の自習支援である。具体的には、未知の日本語をどのような位置づけで学べばよいかを、日本語教師のアドバイスを理想とした、近似的なアドバイスを、例文の解説情報として計算機により提供する方向性である。

さらに、コミュニケーションや協調の拠点として、計算機を使う方向性もある。そのような方向性においても、図1の基盤システムにおける処理機構は応用技術の開発において重要な役割を果たしうると考えているが、本稿の範囲を超える。

2.2 日本語文解説知識の構造化

日本語学習支援に対して有益なのは、現場に即した文に対しその解説情報を提供することである。しかし、言語の性質の本質として、日夜に生成される発話や文が多様であり、そのバリエーションは無限に存在する。従って、言語学習の際の無限に多様な文に対して、それぞれに個別な解説情報を詳細に提供することは不可能である。つまり、学習に即した解説情報として、当該の専門分野における日本語学習に適した限られた例文を用意し、その例文を用いて学習者に日本語の重要な性質、特徴、発話機能等を解説することが現実的である。

実際、言語教育において主流となる考え方の一つは、一般的な言語の性質をまず例文を用いて学習し、その上で専門的な語彙を積み上げていく方法である。具体的には、学習例文に基づく日本語学習においては、次のような知識階層から伝統的な教材が構成されているように見受けられる。

基礎知識: 初級の教科書に出現する例文、語彙(旧日本語能力試験で3級、4級レベルの出題範囲とされる文型、語彙)の列挙・解説。主に動詞の活用を基準として文型を規定する。表現の形式と目的が一致することを前提としている。文体は丁寧体「です・ます」を基本とする。

専門知識: 専門分野の特徴的な場面や行動、知識に関連した例文が示され、場面や専門性に応じた個別の語彙の列挙・解説。話者の態度や、聞き手への配慮を示す表現等も発展として解説が与えられる。

ここで、専門分野での言語運用では、語彙の列挙

中心の知識記述だけで、現場で接する日本語に対応することは容易ではなく、現場の日本語運用が、既に学習した基礎知識とどのような相関にあるかを解説情報として与えることが求められる。すなわち、協調的な例文データベースの構築・蓄積においては、現場で出現しそうな文を、解説情報を考慮せずに「むやみやたら」に登録しても教材として価値は低いと言える。

他方、計算機によって言語学習支援を行うためには、言語処理技術が必須となるが、言語処理技術においても言語の多様性に対応することは重要課題である。そのため、言語表現の多様性を評価する解析モジュールが提案されてきた。例えば、実際の言語運用で使われる日本語は、母語話者であっても誤用が存在し、そのような日本語もある程度は計算機によって解析が可能である必要がある。我々の研究グループも、既にそのような誤用例のデータベースを開発し、誤用や文体に即した不適切性を認識するための評価指標の比較[1]や学習時に気づきを促す上での効果的な情報提示方法、編集方法[2][3]を研究してきた。

また、日本語における誤用だけではなく、文体が多様性を生じさせる要因になる。例えば、丁寧語・敬語分析については、岩下ら[4]、白土ら[5]をはじめとする研究や学習支援システム構築の蓄積がある。それらの研究では、例えば、「食事」が「お食事」、「言う」が「おっしゃる」というように語彙レベルの言い替えが基本データとなる。

以上のような文の多様性を扱うためには、それを扱うための体系が必要なことを、誤用と敬語・丁寧表現の語彙知識を例に示した。このように、学習教材となる文は多面的な角度から言い換えを分析することが必要となる。

そのような背景から、我々は、計算機による学習者支援において、現場の学習者が知っている基礎知識を基本例文とその解説の形でデータベース化しておき、現場に即した文がどのような点で基本例文と類似し、また、どのような特徴を持つのかを多面的に評価し、

解説情報として学習者に提供する技術が基盤になると考えている。

2.3 解説情報基盤としての基本例文集

本研究では、日本語学習用に作成された一般的な基本例文集を用いて、学習者向けの日本語文の解説情報を提供することを考えている。この例文集は、学習者の基礎知識を考慮し、従来の文型積み上げ型の初級日本語教材と整合性を保ちつつ、実際の言語運用を重視して、次のような特徴を持つ。

- 1) 例文の提示:対話対
- 2) 例文の選択:日常的な行動, 状況が選択基準
(文法, 動詞の活用が主たる基準ではない)
- 3) 例文の情報:
語用論的バリエーションについて解説
基本的な対応関係の崩れについて例示

上記3)の語用論的バリエーションの例として、「評価を求める「どう」質問-返答」となる対話対を示す。

- Q:「食事はどうでしたか」
A(基本):「おいしかったです」
A(発展):「たくさん食べられました」

このように、評価を求める質問に対しては形容詞などの評価表現を含む文で返答することが基本的な対話対と考えられる。しかし、形式上は評価表現を用いずに、内容は「評価」として返答することも可能であるため、発話対として成立する。このような言語運用についても、既存の教科書を用いた学習による対話の練習や暗記にとどまらず、運用の場での学習支援を目指すために、バリエーションの提示や解説も視野に入れた。

2.4 基本例文からの言い替え例収集

すでに2.2節で述べたように、基本例文から多様な文の解説情報を提供するためには、基本例文から当

該の文がどのように言い替えられたかを評価することが必要となる。また、そのような評価指標は多面的であることが望ましい。

そのような評価指標の実現には、文の特徴を捉える上での前提となるドメイン知識を整備する必要がある。ドメイン知識の整備には、次の節で説明するが、知識を体系化することが必要であり、文の言い換え、誤用分析、敬語等をより適切に整理するために文を計算機で扱うための意味表現を設計することが必要となる。それは、MedSLT[6]等の特定分野の機械翻訳や異文化コラボレーションツールの内部意味表現として、目的に応じた適切な機械翻訳用の意味判断のために中間言語と呼ばれる意味表現モデルが採用されていることに対応する。

以上の方針に基づき、基本例文とドメイン依存の対話例との対応付けを基本とし、対話例の言い換え可能性について、以下の視座から言い換え例を収集している。

1) 発話主体の属性や人間関係による言い換え

敬語や丁寧語に関わる。

2) 基本例文からの発話意図による言い換え

基本例文:「ここに名前を書いてくださいーはい」

言い替え:「ここに名前を書きましょうーはい」

↓対象ドメインでの対における言い替え

対話例文:「今から着替えてくださいーはい」

言い換え:「今から着替えましょうーはい」

3) 感性表現に関わる言い換え

発話者の多様な身体・心情の状況の提示についても対応が困難である。例えば、医療福祉分野ではオノマトペを中心とした、患者の心理、身体状況に関する表出の気づきについて資料が不足している。

頭が痛い → 頭がズキズキする

目に違和感がある → 目がゴロゴロする

以上のような言い替えは、分野に独自の運用にも対応するため、語彙表現の収集だけではなく、ドメイン知識の効率的な収集に計算機技術を援用することが期待される。

3. 介護ドメインの知識表現

3.1 対話に関するドメイン知識

我々は本研究の開始段階において、医療福祉分野での外国人労働者の就労状態や動向を現場関係者や関連学会・団体から調査し、彼ら彼女らの多くが、介護・看護の場面で活躍を期待されており、さらに、そのような現場において、被介護者とのコミュニケーションについて難しさを感じていることを認識した。

その知見に基づいて、既存の日本人向けの介護看護の実践テキストを対象に、テキストマイニングの手法を利用して調査・分析・整理し、介護場面における対話を計算機で扱う上で必要な知識を体系的に整備し、対話例を整理・処理する上での基盤の構築を目指した。このような調査は、特定の分野の専門日本語教育での対話を計算機で扱う上では、対象となる専門分野独特の用語や表現の収集は必須となり、本稿はそこにテキストマイニング技法を援用した事例紹介となる。

本稿が支援として想定する、例文提示型の教育において、対話例がどのような要素から体系的に整理づけるかを検討した。具体的には、我々が設計した対話例の整理体系は、対話例及び、それらに出現しうる語彙・表現を保存する体系であると同時に、次の特徴を持つ。

特徴の一つは、「場面」という整理の体系を導入し、該当分野の特徴を適切に把握・制限し、計算文法や意味処理を定義することである。ただし、「場面」という抽象的なラベルは、内部的な意味表現としては曖昧性が高いため、学習者が関わる行動を基準に、そこに関連して発生する対話例を保存する。このような整理をする利点は単に対話例の検索キーワードを場面よりも詳細化できるだけではない。例えば、介護、看護者に対して気遣いのある発話するために、食事をする時

に「おいしいですよ」と発話することが、行動に関して「提案」という機能的な役割をもつことがある。そういった現象の説明にも、この整理体系は役立つ。それは、発話が敬語であるかどうかを自動判断し、学習者に意識させる支援に整理体系を発展的に利用できる可能性を持つ。

もう一つの特徴は、行動に関わる語彙を対話例とは別に予め豊富に用意しておくことである。行動には、それに参加する主体の評価がつきものである。そのため、行動に関係する対話においては評価的側面があることを常に留意する必要がある。例えば、被介護者が着替えを行った後に「チクチクします」といった発言をした場合は、着替えという行動に関連し、被介護者からマイナスの評価が出ていることを、介護者は受け取らなければならない。後述するように、専門家ではないサービスの受け手は、上の「チクチク」のような直感的で感覚的な表現[8,9,10]を用いがちであるが、そのようなドメイン依存性の高い表現の辞書の整備はなかなか進んでおらず、それらの表現が使われた対話例をすべて登録されることを期待することは難しい。また、不完全な文体をとることが多いため、計算機支援においても技術的な課題が多く、解決のためには不完全な文体に対応できる語彙知識をドメインに即した適切に整備する必要がある。この方向へのテキストマイニング技術の応用は、4節に述べる。

3.2 テキストマイニングを利用したドメイン知識整理

以上に説明したように、患者や被介護者が日常生活の中で、看護および介護に携わる学習者がいかなる形で接しているかを体系的に整理するために、テキストマイニングを利用する。

学習者が被介護者と関与する行動は、日本人のための介護・看護の教科書が日本語で多数出版されている。このような教科書から、学習者と被介護者が関係する日常生活上の行動に関連する語を収集し、それらの行動を体系付けする。その際の観点は、教科書における「洗髪の介助技術」「足浴の介助技術」といった

学習項目である。ただし、それらの学習項目に対して、対話例の実例を整理する上で適切と考えられる介護の日常生活における「場面」は、一対一では対応しない。そのため、以下のような階層を導入して関連語彙群を整理し、ドメイン知識の体系として利用する。

場面—協調活動—行動—関連語彙群

具体的な例で説明する。介護分野の人材養成教科書には、例えば、「着替えの介助」の技術解説として以下のような記述例[7]を考える。

「健側の手を持ちながら、腕を抜いてもらう。」

この文の係り受け解析を行い、その結果を利用した、「着替えの介助」という協調活動に関する行動および、関連語句を整理する。上の例ならば、解説文内に出現する行動に相当する「持つ」「抜く」という語と、それらの行動の対象となる「手」「腕」という語を抽出し、行動とその対象をグラフ構造で保存する。このような作業を当該の学習項目の解説文集合に適用すると、当該項目に関係する語の関連性がグラフ構造に表現ができる。これをドメインの知識表現として利用する。

グラフ構造を用いた知識表現は、当該の協調活動に関する語の関連性を表現することに役立つ。例えば、協調活動の「着替えの介助」において、「手」に関する行動は図2のように抽出することができる。図2は「着替えの介助」に関して、特定条件で抽出した部分グラフである。

このような知識表現は、例えば例文が、どの協調活動と強く関連するかを算出する上で役立つ。例えば、「手を持ち上げてください」という対話例の一部が与えられたときに、この知識表現から対話例に関連する協調活動と行動の対応を検討することができる。また、協調活動に「起床」、「食事」、「排泄」、「入浴」といった日常的な介護の場面に対応付けておき、当該の協調活

動が行われる場所、時間等の情報を付与すれば、文からの場面对応付けも可能となる。

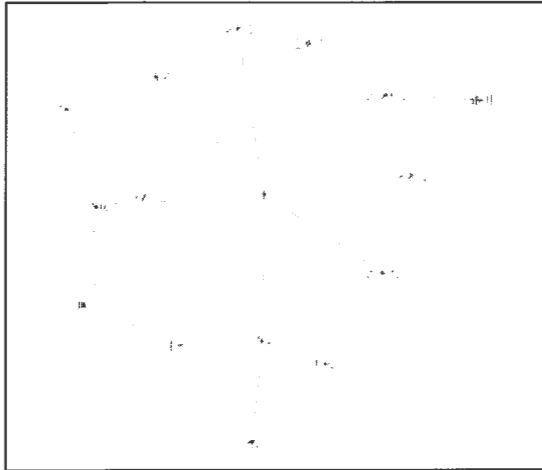


図 2. 協調活動に関係した手を対象とする行動群の例

上記では、行動とその対象の結びつけ手順のみを例示したが、通常、行動の対象語となる名詞の分類や関連性の定義が必要となるが、言い換えを踏まえて、次の評価対象となりうる名詞は、身体の部位、協調行動に関連した動作、「患者の身体、心情的な状態の対象となりうるもの」を関連資料から同様に半自動的な手段で整備している。

4. 多様な言い替えへの対応

4.1 評価表現のドメイン依存性

3 節では、医療福祉分野の日本人向けの教科書データからテキストマイニングすることにより、当該分野のドメイン知識を効率的に抽出整理する例を示した。それに対して、医療福祉サービスの受容者が自然に発する、それぞれの行動に関連付けるける直感的で感覚的な発話に対して、そこで使用される語彙をあらかじめ人手で網羅的に用意することには限界がある。そこで、インターネット上に大量に存在する直感的な記述をその使用文脈とともにあらかじめ収集しておき、当該の表現がシステムに入力された際に、その使用文

脈が一般的であるか否かを検討することを考えた。これは日本語教育ボランティアの教材作成時には、データ追加時の一般性を検討することに役立つ。また、学習者が自習支援の際には、当該の表現が特異であるか否かを検討する上で役立つ。ここでは、ドメイン依存性が高く、発話者の身体や心情の状況を表現する上で重要となる表現を自動収集する技術を紹介する。

ある表現が肯定的／否定的という二値の文脈においてどちらで出現しやすいかを検討した Turney らの研究[11]がある。Turney らの研究では、例えば、「良い」、「美しい」などの形容詞を代表とする評価語は肯定的な極性、「悪い」、「汚い」などの表現は否定的な極性を持つといったように、肯定的・否定的な文脈に特定の表現が出現する度合いをその表現の感性極性として捉える。

インターネット上には、商品についてユーザがレビューを行い、例えば 5 段階評価で 5 なら肯定的評価であり、1 ならば否定的評価であるといった、テキストデータとその評価値が組になったデータが多く存在する。これを教師データとして使用し、評価語に関して Turney の手法を適応すれば、良い評価に関連する語と、悪い評価に関連する語を数値化することが可能となる。

なお、このような評価語の極性値が評価対象となるジャンルごとに異なることは、既に知られている。我々は数多くのジャンルでの評価的意味を体系的に扱うために、意味的な中間表現として感性評価軸を導入した。

感性評価軸は、特定の評価対象について、対立する評価語を見出し語として提示し、当該の性質がどちらの語に似ているかを表現する図 3 のような形式のアンケート等でよく利用される意味表現方法である[12]。我々は、この評価軸をアンケート分析に関わる書籍を網羅的に調査し、できるだけ多くの評価的意味を表現する約 400 軸を用意した。

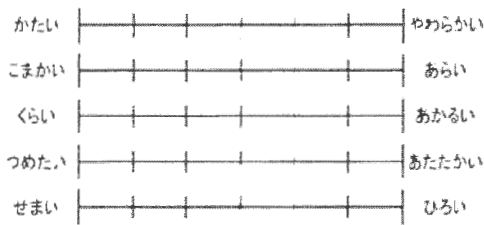


図3. 感性評価軸の例

この約400軸は直感的には肯定的評価を表す評価語と否定的評価を表す評価語の対であり、 i 番目の感性評価軸を形式上、以下のように示す。

$$P_i \leftrightarrow N_i$$

例として、Turneyの感性極性を応用した実験例を簡潔に示す。ここでは、食べ物の評価軸に関わる「スイーツ」ジャンルのテキストデータ、肌に対する評価軸を得ることが期待できる「化粧品」というテキストデータをインターネット上の評価サイトの書き込みから収集し、2つのコーパスとした。それぞれの分野の上位の特色となる感性評価軸の上位3軸は次のようになった。

- 化粧品ジャンル
細かい \leftrightarrow 粗い, かたい \leftrightarrow やわらかい, 引き締まった \leftrightarrow たるんだ
- スイーツジャンル
おいしい \leftrightarrow おいしくない, かたい \leftrightarrow やわらかい, 好き \leftrightarrow 嫌い

以上のように、2ジャンルごとに共通する軸があることを含め、ジャンルの特徴となる感性評価軸が存在する特性を端的に示している。

4.2 機械学習を用いたドメイン依存言い替え表現の収集

前節では、既を用意されている感性評価軸に対して、評価を行ったものであるが、用意した約400の感性評価軸には存在しない評価語が、意味的な中間表現である感性評価軸のうちの、どの使用文脈と近いかを類似度の算出により、評価軸の評価語と入れ替えが可能かを検討した。手順は以下の通りである。この手

順は Turney の感性極性算出法を機械学習手法の代表的な手法の一つである Support Vector Machines(SVM)を使って改良した研究[13,14]に基づいて、感性評価軸のドメイン適合性判定に応用したものである。

- 1) ジャンル j のコーパス $C^{(j)}$ を教師付学習コーパスとして使い、約400軸の感性評価軸の評価語それぞれの感性極性を算出する。軸 i の評価語 P_i, N_i の極性値をそれぞれ $Sp_i^{(j)}, Sn_i^{(j)}$ とする。
- 2) $Sp_i^{(j)}$ と $Sn_i^{(j)}$ のそれぞれに閾値で足りしをした上で、 $|Sp_i^{(j)} - Sn_i^{(j)}|$ の値上位30位以上の感性評価軸を、ジャンル j を代表する感性評価軸の集合 $S^{(j)}$ とする。
- 3) $S^{(j)}$ に含まれる評価語それぞれの出現周辺文脈をSVMによりモデル化する。
- 4) コーパス $C^{(j)}$ において語 w が出現する周辺文脈が $S^{(j)}$ に含まれる評価語のそれぞれのSVMモデルにおいて最も類似する評価語を選択し、その類似度が閾値以上なら当該の評価語と w を言い替え可能と見なす。

上記の手順で、各コーパスに出現するオノマトペ表現上位100語に関して実験を行った。LIBSVM[15]を用い、カーネルにはRBFを選択した。

実験で、オノマトペ表現を対象にしている理由は、言語的には、口語体で身体や心理的な状態を表現する上で、直感的、感覚的に用いられ、意味のドメイン依存性も高いことが知られているからである。また、応用的には上記のような理由から、オノマトペ表現の辞書記述は難しく、また、生成的であるため3節のような方法で予め意味内容を記述しておくことが困難な表現の代表だからである。

実験結果を表2,3に示す。表内では、言い替え可能であったオノマトペ表現を矢印で示してある。このように評価する対象の側面や属性の異なるジャンルごとに、辞書登録がなされていないオノマトペ表現が、意

味的な中間表現である感性評価軸の評価語に関して言い替え例を提案することが可能であることを示している。

現在、医療福祉ドメインでの評価対象と評価表現を評価する上での基盤資源となる、評価サイトのデータを蓄積中であり、オノマトペ表現だけでなく、特定の表現が感性評価軸上と相関をもつ表現かを評価できるように整備を進めている。

表2 化粧品ジャンルでのオノマトペ言い換え

かたい	↔	柔らかい	→	しっとり
重い	↔	軽い	→	すっきり
乾いた	→	さらさら	↔	湿った
多い	→	たっぷり	↔	少ない
清潔な	→	さっぱり	↔	汚い

表3 スイーツジャンルでのオノマトペ言い換え

かたい	↔	柔らかい	→	ふわふわ
重い	↔	軽い	→	さっぱり
乾いた	→	ばさばさ	↔	湿った
多い	→	ずっしり	↔	少ない
強い	↔	弱い	→	あっさり

5. まとめ

医療福祉ドメインに限らず、本研究の基盤部分は様々な専門分野の日本語教育に適応できる可能性がある。その際には、専門ドメインの知識を効率的に作成することが必要となる。そのような状況において期待できる例文の構築方法、関連知識の半自動抽出方法について論じた。ここで論じた基盤技術や知識表現を、工場などの技術研修向けの専門日本語学習支援等に、より一般的に適用できるよう、言語グリッドのサービスとして実装することを検討していきたい。

謝辞

本研究は科研費(基盤研究(C) 研究課題番号: 21500113)の支援を受けて実施した。

参考文献

- [1] 安藤秀明, 秋吉信吾, 竹内和広: 作文指導に向けた表現データベースの構築と表現評価指標の検討, 第15回公開シンポジウム「人文科学とデータベース」, pp.37-42, 2009.
- [2] So NODA, et al.: Awareness Promoting Visualization of Style Inconsistencies in Texts,

- IEEE 2nd International Symposium on Aware Computing, pp.291-297, 2011.
- [3] Kazuki SHIMAMURA, et al.: An effective visualization of style inconsistencies for interactive text editing, 16th International Conference on Artificial Life and Robotics, pp.154-157, 2011.
- [4] 岩下志乃, 岩切智希: 状況に応じた対話による敬語学習システム, 知能と情報, Vol.20, No.5, pp.709-719, 2008.
- [5] 白土保, 丸元聡子, 村田真樹, 神崎享子, 井佐原均, 日本語の敬語誤用判定システム, 計量国語学会, 計量国語学 28 卷 1 号, pp.1-20, 2011.
- [6] M. Rayner et al.: The MedSLT System, Coling 2008 Workshop on Speech Processing for Safety Critical Translation and Pervasive Applications, 2008.
- [7] 下正宗, 深見悦司: 『はじめての介護』, 成美堂出版, 2006.
- [8] 田守育啓, ローレンス・スコウラップ: 『オノマトペ - 形式と意味-』, くろしお出版, 1999.
- [9] Yumi NISHIMURA et al.: A Consideration on the Support for Japanese Onomatopoeia Learning in Japanese for Specific Purposes, 2nd Asian Conference on Education, pp.1164-1180, 2010.
- [10] 西村由美, 竹内和広: 目的別日本語教育におけるオノマトペ表現の重要性, 言語処理学会第17回年次大会, 2011.
- [11] Peter D. Turney: Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews, 40th Annual Meeting of ACL, pp.417-424, 2002.
- [12] 岩下豊彦: 『SD 法によるイメージの測定』, 川島書店, 1983.
- [13] 北中佑樹, 秋吉信吾, 竹内和広: 日本語教育に向けたオノマトペ表現データベースの構築, 日本知能情報ファジィ学会 第20回ソフトサイエンスワークショップ講演論文集, pp.49-52, 2010.
- [14] 北中佑樹, 竹内和広: 印象調査アンケートのためのオノマトペ表現による評価尺度の拡張, 電子情報通信学会技術研究報告「思考と言語」, 2010.
- [15] LIBSVM - A Library for Support Vector Machines.
<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>