

音声対訳コーパスからの 日本語待遇表現生成規則の自動獲得

木村 直樹† 松原 茂樹‡ 小川 泰弘† 稲垣 康善†

†名古屋大学大学院工学研究科

‡名古屋大学言語文化部 / CIAIR

nkimura@inagaki.nuie.nagoya-u.ac.jp

概要

音声翻訳システムが出力する日本語音声は自然な表現であるためには、生成される日本語に適切な待遇表現が用いられていることが望まれる。日本語にはさまざまな待遇表現が存在し、それらは、対話の状況や文法的制約といった種々の要因に基づいて使い分けられる。本論文では、特に述部の待遇表現に注目し、英日対話翻訳において日本語待遇表現の生成規則を自動獲得する手法を提案する。本手法では、上述の種々の要因が与えられたときに適切な待遇表現を決定するための規則を対訳コーパスから帰納学習により獲得する。このため、生成規則を作成するときに、発話を分析したり、要因の重要度を考慮するなどといった煩雑な作業は必要ない。ATR 音声言語データベースを用いて生成規則の学習および待遇表現の生成実験を行った。

Automatic Acquisition of Generation Rules of Polite Expressions from Paralell Corpus

Naoki KIMURA†, Shigeki MATSUBARA‡, Yasuhiro OGAWA†
and Yasuyoshi INAGAKI†

†Graduate School of Engineering, Nagoya University

‡Faculty of Language and Culture / CIAIR, Nagoya University

nkimura@inagaki.nuie.nagoya-u.ac.jp

ABSTRACT

This paper proposes a method for automatically acquiring generation rules of Japanese polite expressions from a spoken dialogue corpus. Japanese predicates have various types of polite expressions. These expressions should be chosen to suit discourse situations, grammatical constraints, and so on. This method makes the generation rules automatically by considering these factors. And our method enables to acquire the knowledge of discourse situations efficiently. An experiment on the dialogue sentences in ATR Speech and Language Database has shown the effectiveness of our method.

1 はじめに

人との対話では、待遇表現はコミュニケーションを円滑に行う上で重要な役割を果たす。待遇表現を用いることにより、発話に丁寧さを与えるとともに、聞き手や話題に登場する人物に対する敬意を表現する。日本語対話では多くの待遇表現が用いられる。特に、述部には待遇表現が多く現れる。どのような待遇表現が適切であるかは、対話の状況や文法上の制約など様々な要因に依存する。そのため、用い方を誤ると、相手

に対して失礼な表現となったり、発話として不自然になったりする。

音声翻訳を介した異言語間対話では、システムが話し手の発話を正しく翻訳することにより、話し手の意図が聞き手に伝わる。システムが出力した音声は自然な発話として聞き手に受け入れられるためには、システムが適切な待遇表現を生成することが望まれる。

これまでに、英日音声翻訳において適切な待遇表現を生成するための研究が行われている [2, 3, 6, 12]。このうち [2, 3] では、対話者の立場の上下関係や発話

の状況などをパラメータとし、それらの値に基づき適切な待遇表現を生成する。しかし、待遇表現の運用法をあらゆる文に対して網羅することは難しい。また [6, 12] では、用例に基づく翻訳システムにおいて対話者の社会的役割などを考慮した翻訳規則を追加することにより、適切な丁寧度の翻訳文を生成する。しかし、用例ごとに翻訳規則を作成する必要があるため、規則の作成に多くの時間と労力を必要とする。

そこで本論文では、英日対話翻訳において日本語述部の待遇表現生成規則を自動獲得する手法を提案する。本手法では、対訳コーパスから述部における待遇表現の生成事例を抽出し、帰納学習により規則を獲得する。

本手法では、学習に用いる事例は話者情報、英語発話及びその英日通訳発話から抽出したパラメータを用いる。翻訳のときにも学習のときにも値を自動的に獲得できるように、これらのパラメータを定めた。対訳コーパスから得られる事例から学習した生成規則を用いることにより、システムが通訳者と同様の基準で待遇表現を使い分け可能となることが期待できる。ATR 音声言語データベース [11] を用いて待遇表現の生成実験を行い、本手法を評価した。

本論文の構成を以下に示す。続く 2 章では、対訳コーパスから事例を抽出し、待遇表現生成規則を学習する手法について述べる。3 章では、英日対話翻訳における待遇表現生成について述べる。4 章では、待遇表現生成実験とその結果を述べ、考察を行う。

2 待遇表現生成規則の学習

本研究では、英語話者と英日通訳者の音声を収録した大規模対訳コーパスを用いる。まず、コーパスの各発話対応から、待遇表現の生成に関連する種々のパラメータの値を抽出し、生成事例を作成する (2.1 節)。次に、作成したすべての事例を用いて帰納学習を行い、待遇表現の生成規則を獲得する (2.2 節)。

2.1 待遇表現生成事例とその獲得

待遇表現生成事例は、待遇表現と、それを生成する要因となるパラメータの値のリストから構成される。これらはいずれも、通訳発話の中に待遇表現を含む対訳対応発話の組から自動的に獲得する。

なお、本研究では、派生文法 [5] に従って日本語形態素を記述する。派生文法では用言や助動詞が語形変化するのではなく、語幹に接辞が接続することにより動詞句が形成されると考える。例えば、「分かりました」は、「分 kar」,“(i)mas”,“(i)ta” から構成される¹。述部構成素を接続するための規則は極めて単

¹形態素中の“(i)”は連結母音と呼ばれる。連結母音には“(i)”のほかにも“(a)”がある。また、子音にも同じような振る舞いをするものがあり、それらは連結子音と呼ばれる。これらは接続する形

表 1: 待遇表現

形態素	品詞	形態素	品詞
(i)mas	派生接尾辞	kudasar	動詞
(r)are	派生接尾辞	mair	動詞
desita	繫辞	o	接頭辞
desu	繫辞	or	動詞
desyou	繫辞	sama	派生接尾辞
go	接頭辞	san'	派生接尾辞
gozar	動詞	御	接頭辞
irassyar	動詞	申 si	動詞
itadak	動詞	上 ge	動詞
itas	動詞	様	派生接尾辞

純であり、本研究のような構成的アプローチに基づく日本語生成の文法として適している [4]。

2.1.1 待遇表現

事例の作成でコーパスから抽出する待遇表現の種類を定めるために、ATR 音声言語データベースに収録された、旅行をドメインとする 334 個の対話 (全 4,304 ターン) を用いて、出現する待遇表現の種類を調査した。調査は、英日通訳者によるすべての発話を、派生文法に基づく形態素解析システム MAJO [8] を用いて形態素に分割して行い、述部に含まれる付属語²のみを対象とした。なお、本研究では、述部を文末の 1 文節³と定義する。

全 4,906 個 (1,508 種類) の述部から待遇を表す形態素を抽出した結果、19 形態素が該当した。その内容を表 1 に示す。本研究では、これらの 19 種類の形態素の組み合わせでもって一つの待遇表現と定める。この調査では、67 種類の組み合わせが見いだされた。主な組み合わせを表 2 に示す。本研究では、この 67 種類の組合せの中から最適な待遇表現を選択するための規則を自動的に生成する。

2.1.2 パラメータの種類

人間が対話において発声するときには、(1) 発話状況 (話し手の立場や聞き手との関係など)、(2) 発話内容 (発話の意図や聞き手に伝えたい事柄など)、及び、(3) 発話形式 (発話生成において採用する語彙や文法など) といった要因を総合的に判断して、用いるべき

形態素の末尾の音により現れたり現れなかったりするが、そこには単純な規則性が存在する。

² 付属語とは、派生文法における接尾辞、繫辞、補助動詞などのことを指す。

³ ただし、その文節の先頭の形態素が補助動詞である場合は直前の文節も含む。

訳文の発話形式

表 2: 待遇表現の組合せ

組合せ	使用例
(i)mas	分かりました
desu	結構です
gozar, (i)mas	ございますか
(なし)	よかった
o, (i)mas	お願いできますか
or, (i)mas	付いております
desyou	よろしいでしょうか
itadak, (i)mas	教えていただけますか
kudasar	教えてください
o, kudasar	お待ちください

待遇表現の種類を決定していると予想される。

上記の (1), (2), (3) は、対話翻訳おける発話生成では、話者の発話状況、原文の発話内容、訳文の発話形式にそれぞれ対応する。これら 3 つの要因それぞれに対して、値を自動的に獲得できることを条件に、事例の作成に用いるパラメータの種類を定めた。以下では、各要因ごとのパラメータの種類とその獲得方法およびその採用理由について述べる。また、表 3 に各パラメータのとり得る値を示す。

話者の発話状況

- 発話者の役割 対話者同士の立場の上下関係は待遇表現の使われ方に影響を及ぼす。一般的には、立場がより低い者の方が、より丁寧な表現を用いる。例えば、客と受付との対話における英語発話 “I’ll wait for your call.” の日本語訳は、客の発話であれば「電話を待っています」とした方が、また、受付の発話であれば「お電話をお待ちしております」とした方が発話として自然になる。コーパスに付与されている発話者情報をパラメータの値とする。

原文の発話内容

- 英語発話文の丁寧度 原文が丁寧であれば、翻訳文も丁寧な表現であることが望ましい。could, please, would が原文に出現するか否かでパラメータの値を表す。
- 発話に登場する人物 自分自身のことに言及するのであれば謙譲語を用い、相手のことであれば尊敬語を用いるといったように、発話中に登場する人物によって待遇表現は異なる。人物を表す語である I, you など原文に出現するか否かでパラメータの値を表す。

- 発話の話題となる事柄 発話の内容によって、文の丁寧度は変化する。ATR 音声言語データベース [11] の旅行受付発話の英日通訳を用いた調査によると、旅行者の人数を尋ねる発話では、「お知らせください (o/ 知 rase/kudasar/(i))」のように “kudasar” を用いた表現と「教えていただけますか (教 e/(i)te/itadak/e/(i)mas/(r)u/ka)」のように “itadak” を用いた表現がともに出現したのに対し、電話番号を尋ねる発話では、“kudasar” を用いた表現はほとんど現われなかった。注目している述部を含む日本語通訳文から獲得する。述部に最も近い名詞をパラメータの値とする。

- 日本語生成文の種類 待遇表現によっては文がある特定の種類であるときに限り使われるということがある。例えば、ATR 音声言語データベースの調査では、“kudasar” はその大半が依頼や命令を表す文で使われている。パラメータの値として、平叙、疑問、命令の 3 種類を考え、対象としている述部を含む日本語通訳文から獲得する。文末の形態素が終助辞 “ka” であれば“疑問”、命令の接尾辞であれば“命令”、それ以外であれば“平叙”とする。

- 日本語述部語幹 本論文では、語幹は述部を構成する形態素のうち最初に出現する動詞、形容詞もしくは名詞を指す。例えば、「お願いいたします」であれば“願 w”である。待遇表現 “o, itas” は、「お願いいたします (o/ 願 w/(i)/itas/(i)mas/(r)u)」のように、“願 w” とともに用いることができるが、「* お分かりいたします (o/ 分 kar/(i)/itas/(i)mas/(r)u)」のように“分 kar” とともに用いることはない。このように、語幹と待遇表現の間には接続しやすいものと接続しにくいものがある。注目している述部の語幹をパラメータ値とする。

- 日本語述部構成素 述部構成素の中には互いに共起しにくいものが存在する。例えば、希望の接尾辞として“(i)ta”があるが、「泊まりたいんです (泊 mar/(i)ta/i/n’/desu)」のように“desu” とともに用いることはできるが、「* 泊まりました (泊 mar/(i)mas/(i)ta/i)」のように“(i)mas” とともに用いない。本研究では英語文の解析により獲得できる“[可能]”、“[過去]”など 15 種類の文法機能をパラメータとして採用する。パラメータ値は、注目している述部にその文法機能を担う形態素が存在すれば“有”、存在しなければ“無”とする。

表 3: 各パラメータのとり得る値

パラメータ	値
発話者の役割	申込者, 担当者
英語発話文の丁寧度	could(有, 無), please(有, 無), would(有, 無)
発話に登場する人物	I(有, 無), my(有, 無), me(有, 無), mine(有, 無), we(有, 無), our(有, 無), us(有, 無), ours(有, 無), you(有, 無), your(有, 無), yours(有, 無)
発話の話題となる事柄	予約, フィリップス, お部屋, 承知, 電話番号, ホテル,[なし] など 997 種類
日本語生成文の種類	平叙, 疑問, 命令
日本語述部語幹	願 w, gozar, nar, 思 w, 分 kar, kasikomaru など 662 種類
日本語述部構成素	[否定](有, 無), [命令](有, 無), [様態](有, 無), [希望](有, 無), [過去](有, 無), [条件](有, 無), [継続](有, 無), [完了](有, 無), [可能](有, 無), [現在](有, 無), [使役](有, 無), [推量](有, 無), [授受](有, 無), [試行](有, 無), [過度](有, 無)

担当者:	Okey. Could, you tell me how many people and when you would like to stay, please ?
通訳者:	分かりました。何名様でしょうか。それと滞在期間もお知らせ願いたいんですけども。

図 1: 対訳コーパスにおける対応発声

2.1.3 待遇表現事例の作成例

図 1 に示す対訳対応発話の組の中の述部「分かりました」を例に、各パラメータの値の抽出方法を説明する。

まず、形態素解析により、それを“分 kar”, “(i)mas”, “(i)ta” に分解する。このうち、待遇表現は“(i)mas” となる。

各パラメータの値は次の通りである⁴。発話者の役割は対応する英語発話の話者の役割である“担当者”とする。英語発話文の丁寧度は対応する英語発声に could, please, would がともに出現するため、それぞれ“有”となる。発話に登場する人物は、英語発声に出現する you, me は“有”, それ以外は“無”とする。発話の話題となる事柄は、この事例では日本語発話文が述部のみからなり、それ以前に名詞が存在しないため、“[なし]”とする。日本語生成文の種類は、日本語通訳文「分かりました」の末尾の形態素が完了の接尾

⁴ 図 1 のように、対訳コーパスによっては、両言語間で文単位ではなく、対話ターンを単位とした対応のみが付与されていることもある。この場合は、上述のパラメータのうち、(2) 原文の発話内容のパラメータに限っては、ターン内のすべての英語発話を対象に値を抽出する。

図 2: 学習事例の例

パラメータ	値
発話者の役割	担当者
英語発話文の丁寧度	
could	有
please	有
would	有
発話に登場する人物	
me	有
you	有
その他	無
発話の話題となる事柄	[なし]
日本語生成文の種類	平叙
日本語述部語幹	分 kar
日本語述部構成素	
[過去]	有
その他	無
待遇表現	(i)mas

辞“(i)ta”であるため、“平叙”とする。日本語述部語幹は“分 kar”とする。派生文法における完了の接尾辞は英語では過去形に対応するため、日本語述部構成素の 15 種類のパラメータのうち、[過去]のみ値を“有”とし、それ以外は値を“無”とする。これらをまとめると表 2 のようになり、これを学習事例とする。

2.2 生成規則の帰納学習

前節で作成した学習事例のように、(パラメータ, パラメータ値, クラス) の形式で表される事例集合からの帰納学習によりクラス分類のための規則を獲得

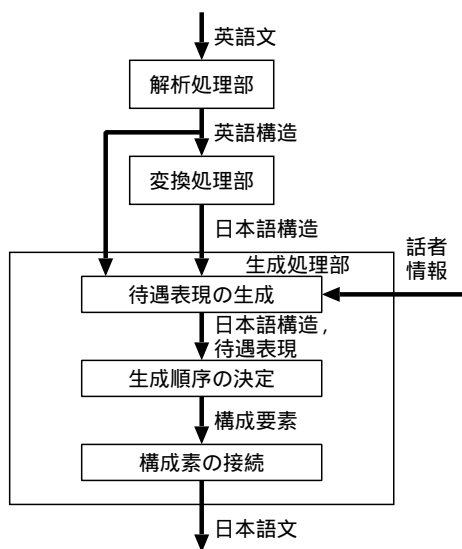


図 3: 英日対話翻訳システムの構成

するアルゴリズムは、これまでにいくつか提案されている [1, 9, 10]。本研究では、そのうちの一つである C4.5[10] を用いて待遇表現生成規則を作成する。

C4.5 は事例集合のエントロピーを最も減少させるテストを選び事例集合の分割を繰り返すことにより、決定木を学習する。分割が行われた集合は中間ノードに対応し、分割が行われなくなった集合は葉ノードに対応する。各葉ノードは、そのノードへの経路をもつ学習事例が属するクラスのうち、最も多くの事例が属するクラスを表す。また、中間ノードにはパラメータが、枝にはその値が割り当てられている。

待遇表現の運用に影響を与える複数の要因の重要度を人手により順序づけし、その結果を反映する待遇表現生成規則を作成することは難しい。しかし、C4.5 を用いることにより重要なパラメータから順に規則に反映させることができるため、要因間の重要度の違いを考慮した規則を自動的に作成できる。

3 対話翻訳における日本語述部生成

決定木を用いた待遇表現の生成は、根ノードを出発点とし、各ノードのパラメータの値に基づいて次にたどる枝を決定していく。葉ノードに到達したとき、そのノードに割り当てられたクラスが、訳文で用いる待遇表現であるとして判定される。

この待遇表現生成プロセスを対話翻訳システムに導入するには、図 3 に示すシステム構成を実現すればよい⁵。待遇表現生成モジュールには、話者情報 (発話

⁵ ここでは構成的アプローチに基づく一般的な機械翻訳システムへの導入を想定している。

状況)、英語発話の解析結果 (原文内容)、及び、日本語構造 (訳文内容) が入力される。日本語生成部の処理の流れは以下の通りである。

1. 待遇表現の生成：入力された情報から、前章で示したパラメータ抽出手法にしたがって事例を獲得し、それに待遇表現生成規則 (決定木) を適用する。
2. 生成順序の決定：変換処理部及び待遇表現生成部で生成されたすべての述部構成素を用いてその生成順を決定する。
3. 派生文法における連結の原則に従って各述部構成素を連結し、日本語述部を生成する。

例えば、“I understood.” を「分かりました」と翻訳する場合、日本語構造から、動詞が“分 kar”であり、時制が過去であるという情報が得られる。待遇表現生成処理の結果、“(i)mas” が選択されれば、述部が“分 kar”, “(i)mas”, “(i)ta” の順序で構成されることがわかる。それらを接続することにより「分かりました」を生成する。

4 待遇表現生成実験

4.1 実験の概要

本論文で提案した手法の利用可能性を評価するため、ATR 音声言語データベースを用いた待遇表現生成実験を行った。

まず、コーパス内の、旅行代理店やホテルのフロントなどにおける客と店員との対話 326 個のうち、英語話者発話と英日通訳者発話の対応組 4,144 ターン (5,535 文) から各日本語述部ごとに学習事例を作成した。事例数は 4,751 個である。また、これらの事例を用いて帰納学習を行った結果、ノード数 33,081 個の規模の決定木が得られた。決定木の一部を図 4 に示す。

実験のテストデータとして、上述のデータと異なる 8 対話 (160 ターン, 192 文) を用いた。そこから取り出した 155 個の日本語述部に対してテスト事例を作成した。実験では、テスト事例を用いて決定木をたどり、生成された待遇表現を評価した。

なお、対話翻訳システムの生成処理は、現在、待遇表現生成モジュールのみが実装されており、生成処理における語順決定や構成素接続などは引き続き実装を進めている段階である。システムが生成した述部全体を対象とするのであれば、人間による主観的な評価も可能であるが、付属語の組み合わせである待遇表現に対してそのような基準で評価することは難しい。そこで、本実験では、生成された待遇表現がコーパスに出現している待遇表現と同一であるか否かによって判定することとした。すなわち、正解 (コーパスの待遇表

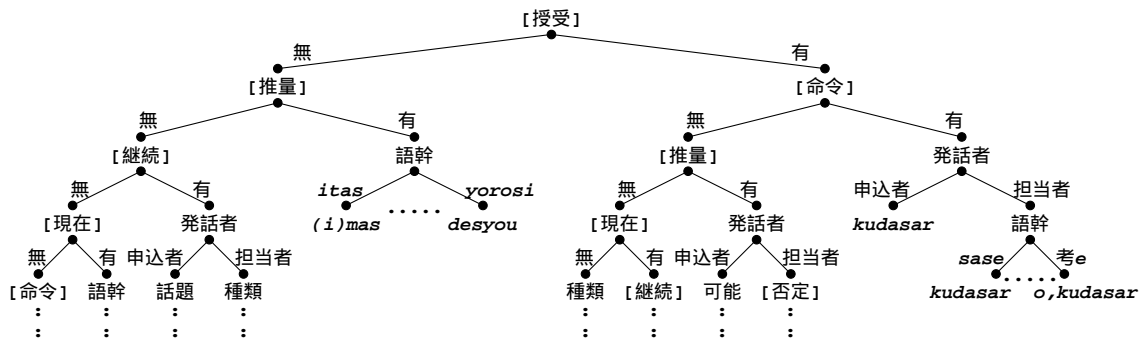


図 4: 作成した決定木の一部

表 4: 待遇表現生成実験の結果 (155 事例)

	事例数	割合 (%)
正解	119	76.8
不正解	14	9.1
失敗	22	14.2

現と完全に一致する), 不正解 (コーパスの待遇表現と部分的に一致, または, 一致しない), 失敗 (学習データに現れないパラメータ値をもつテスト事例のため, 決定木を葉ノードまでたどれない) の 3 段階で評価した。

4.2 実験結果と考察

実験結果を表 4 に示す。全テスト事例の 76.8% に相当する 119 事例に対する生成実験において, 生成された待遇表現がコーパスの待遇表現と一致した。このことから, 決定木学習を用いて待遇表現生成規則を自動獲得する本手法の利用可能性を確認した。9.1% に相当する 14 事例に対する生成実験では, コーパスとは異なる待遇表現を生成した。残りの 22 事例は待遇表現の生成に失敗した。学習事例に出現しないパラメータ値については, 決定木をたどる途中においてたどるべき枝が存在しないため失敗する。

待遇表現生成実験で作成された決定木の形状を分析すると, 上位ノードにおいて日本語述部構成素の存在, 発話者の役割, 述部語幹の種類といったパラメータが反映されていることがわかった (図 4)。このことから, 待遇表現の運用はこれらの要因と強い関係があることが予想される。

不正解 (正解と不一致) であると判断された 14 事例の内訳を表 5 に示す。このうち A は, 生成された待遇表現がコーパスの待遇表現を含んでおり, G はその逆となっている。待遇表現が “o, (i)mas” から “o,

表 5: 生成された待遇表現が正解と異なる事例

種類	生成された待遇表現	コーパスの待遇表現	頻度
A	o, itas, (i)mas	o, (i)mas	4
B	gozar, (i)mas	desu	3
C	desu	(i)mas	2
D	(i)mas	desu	1
E	desu	(r)are, (i)mas	1
F	desu	go, gozar, (i)mas	1
G	o	o, desu	1
H	様, desu	様, gozar, (i)mas	1

itas, (i)mas” になることは, 例えば, 「お待ちしています」が「お待ちいたしています」になることに相当する。どちらの述部を用いた場合でも聞き手にとって大きな違いはなく, 評価方法によっては正解とみなせる可能性がある。

C4.5 は各学習データの分布に基づき, 葉ノードに複数のクラス候補を割り振ることができる。待遇表現の候補を複数用意することにより, 待遇表現生成後の処理において, 構成要素間の接続可能性を考慮するなどして, 学習データ不足などに起因する不適切な待遇表現の生成を回避できる可能性もある。実際, 正解と一致しなかった 14 事例のうち 6 事例に対して生成された待遇表現は第 2 位の候補であった。

待遇表現を決定できなかった 22 事例に対しては, 次のような解決方法が考えられる。あるノードにおいて決定木の探索に行き詰まったときは, そのノードのすべての子ノードから探索を再開する。到達した複数の葉ノードにおける学習データの分布に基づき, 最大尤度の待遇表現を出力する。上記の 22 事例についてこの方法により待遇表現を生成したところ, 11 事例

に対してコーパスと同じ待遇表現を生成した。

5 おわりに

本論文では、英日対話翻訳において日本語述部の待遇表現生成規則を自動獲得する手法を提案した。本手法では、対訳コーパスから述部における待遇表現の生成事例を抽出し、帰納学習により待遇表現生成規則を獲得する。そのため、通訳者が発話状況、原文の発話内容、訳文の発話形式といった多くの要因に基づき待遇表現を使い分けることと同じように、システムが待遇表現を使い分けることが可能となる。本手法では、これらの要因として発話者の役割、英語発話文の丁寧度、発話に登場する人物、発話の話題となる事柄、日本語生成文の種類、日本語述部語幹、日本語述部構成素などのパラメータを用いた。これらのパラメータに対する値を自動的に獲得する方法を提示することにより、本手法の汎用性を示した。ATR 音声言語データベースを対象とした待遇表現生成実験の結果、本手法の利用可能性を確認した。

実験の結果、学習によって得られた決定木から、述部構成素の存在、発話者の役割、述部語幹の種類が待遇表現の運用基準と強い関連性を有していることがわかった。使用するパラメータを限定すれば、本手法を日本語音声対話システムの応答生成に利用できることが期待できる。これは今後の検討課題としたい。

参考文献

- [1] Breiman, L., Friedman, J., Olshen, R. and Stone, C.: *Classification and Regression Trees*, Wadsworth (1984)
- [2] 石原佳典, 奥村明俊, 赤峯享: 多言語機械翻訳における「ていねい」表現の生成, 人工知能学会第5回全国大会, pp.577-580 (1991).
- [3] 石原佳典, 亀井真一郎, 赤峯享: 対話文の英日機械翻訳における日本語待遇表現の生成, 情報処理学会第46回全国大会講演論文集(3), pp.155-156 (1993).
- [4] 木村直樹, 松原茂樹, 小川泰弘, 外山勝彦, 稲垣康善: 英日機械翻訳における日本語述部生成 - 派生文法に基づく方法 -, 情報処理学会第60回全国大会講演論文集(2), pp. 105-106 (2000).
- [5] 清瀬義三郎則府: 日本語文法新論 - 派生文法序説 -, 桜楓社 (1989).
- [6] Mima, H., Furuse, O. and Iida, H.: A Situation-based Approach to Spoken Dialog Translation Between Different Social Roles, *Proceedings of TMI-97*, pp.176-183 (1997).
- [7] 南不二男, 林大, 林四郎, 芳賀綏: 敬語の体系, 敬語講座1 敬語の体系, 明治書院 (1974).
- [8] 小川泰弘, ムフタル・マフスット, 外山勝彦, 稲垣康善: 派生文法に基づく日本語形態素解析, 情報処理学会論文誌, Vol.40, No.3, pp.1080-1090 (1999).
- [9] Quinlan, J. R.: Induction of Decision Trees, *Machine Learning*, Vol.1, pp.81-106 (1986).
- [10] Quinlan, J. R.: *C4.5 Programs for Machine Learning*, Morgan Kaufmann (1993).
- [11] 浦谷則好, 竹沢寿幸, 松尾秀彦, 森田千帆: 音声言語データベースの構成, テクニカルレポート TR-IT-0056, ATR 音声翻訳通信研究所 (1994).
- [12] 山田節夫, 隅田英一郎, 柏岡秀紀: 対話者の社会的役割を利用した訳し分け手法, 自然言語処理, Vol.8, No.1, pp.175-190 (2001).