

発話と発語内行為命名動詞を用いた 言語行為コーパスの自動構築と分析

大山 鉄郎^{1,a)} 稲垣 陽一^{2,b)} 関 洋平^{3,c)}

概要：近年，Siri に代表されるような，対話理解・生成システムが注目されている．これらの対話理解・生成システムと人がスムーズな対話を行うには，システム側が人の発話の意味を理解し，適切な返答をする必要がある．本研究では，対話理解・生成システムの実現に向けて，発話の意味の分析や判定に有用である，ブログをもとにした言語行為コーパスの自動構築を行なう．具体的には，発話の引用と発語内行為命名動詞（illocutionary act naming verb）というパターンの機械的抽出によって，ブログが話題とする多様な対象領域から発話と発語内行為命名動詞の対データを収集する．構築した言語行為コーパスの精度は，3名の被験者の人手判定を正解とした重み付け精度で 0.522 となり，自動構築した言語行為コーパスの有用性を示した．また，構築した言語行為コーパスの利用例として，入力を発話，出力を言語行為タイプとする分類器の構築を行った．言語行為タイプは，会話に頻出する言語行為や，その言語行為に対して対話理解システムの反応を変化させる必要があるかという点から，11 種類を定義した．構築した分類器の分類精度は，マクロ平均で精度 0.555，再現率 0.545，F 値で 0.546 となった．また，4名の被験者の人手判定を正解とした重み付け精度は 0.560 となり，自動構築した言語行為コーパスを利用することで，言語行為が適切に分類できることを示した．

1. はじめに

近年，Siri ^{*1}に代表されるような，対話理解・生成システムが注目されている．これらの対話理解・生成システムとユーザがスムーズな対話を行うには，システムがユーザの発話の意味を適切に理解する必要がある．ここでの発話の意味とは，文の内容としての意味ではなく，ユーザがどのような意図でその発話を行ったかという意味を指している．システムが発話の意味を正しく判定できるようになることで，ユーザの発話に対して，その意図を汲んだ反応をすることができる．

本研究では，発話意味を理解するための枠組みとして言語行為 [4] を用いる．言語行為は，語の働きを，事実の描写や記述といった側面からではなく，命令や約束など行為の遂行といった側面から捉えた概念である．

言語行為や発話意図に関する研究では，コーパスが利用されることが多い [6], [7], [8]．言語行為がアノテーション

されたコーパスは，言語行為に関する分析や，言語行為を判定するための分類器の作成など，発話に関する広い範囲の研究で有効な資源だといえる．しかし，コーパスの構築は，人手によるアノテーションを必要とするため作成コストが高く，大規模なコーパスの構築が難しいという問題点がある．

本研究では，対話理解・生成システムの実現に向けて，大量かつ多様な文を含むと考えられるブログ記事から，言語行為コーパスの自動構築を行なう．具体的には，言語行為を示唆する発語内行為命名動詞（illocutionary act naming verb） [5]（以下，命名動詞とする）に着目することで，発話と言語行為の対データをブログから抽出する．命名動詞は，発話と共起することで，その発話の言語行為を特定するための手がかりとなる動詞である．命名動詞の有用性について，以下に例を示す．

- (1) 頼む からうちのチームに入ってくれないか？
- (2) その意見は絶対に間違っていると 反対した
- (3) 「明日の天気が晴れだったらテニスをしよう」と 誘った

上記の例では，「頼む」，「反対する」，「誘う」という命名動詞が含まれている．これらの命名動詞は，発話の言語行為を特定する上で有用だと考えられる．命名動詞は，(1) のように発話自体に含まれたり，(2)，(3) のように発話の後に続く形に続いたり，様々な形で発話に出現する．こ

¹ 筑波大学大学院 図書館情報メディア研究科

² 株式会社きざしカンパニー

³ 筑波大学 図書館情報メディア系

a) ohyama@slis.tsukuba.ac.jp

b) inagaki@kizasi.jp

c) yohei@slis.tsukuba.ac.jp

*1 <http://www.apple.com/jp/ios/siri/>

のとき,(3)のような引用された発話は,引用符+命名動詞という形式的パターンを持っており,容易に任意のテキストから機械的に抽出することができる.この際,引用された発話に続く命名動詞は,その文の書き手が発話に対して行った言語行為のアノテーションと捉えることができる.

本稿では,(3)の形式で出現する発話に焦点をあて,発話と発語内行為命名動詞を用いた言語行為コーパスの自動構築手法について提案する.また,コーパスの利用例として,11種類の言語行為を分類する分類器の構築と評価を行う.

本稿の構成を以下に示す.まず,2節で関連研究について紹介し,3節でコーパスの自動構築手法について説明する.4節では,構築したコーパスの適用例として,分類器の構築・評価を行い,5節でまとめと今後の課題を示す.

2. 関連研究

発話意図や発語内行為を付与したコーパスに関する研究は,いくつか行われている [6], [7], [8].

翠ら [7] は,音声対話システムへの応用を目的に,発話行為タグと,意味内容タグの2つを定義した.発話行為タグは,本研究における言語行為を表し,意味内容タグは,発話がどのような内容に対して言及しているかを表す.これらのタグを京都観光案内対話コーパスの文節単位ごとに付与し,タグの出現傾向などからコーパスの分析を行っている.翠らの定義した発話行為タグセットは,人手による付与であること,複数のレイヤーから構成され柔軟な定義ができる点で本研究と異なるが,発話の意図をアノテーションするという点では本研究と類似している.

熊本ら [6] は,話し言葉対話システムの自然言語処理部として,発話意図タイプ(本研究における言語行為)がアノテーションされたコーパスから,発話意図タイプ決定ルールを自動生成するアルゴリズムを提案している.ルールの自動生成という点と,本研究のコーパスの自動構築の点において類似性が見られる.しかし,熊本らの手法では,あらかじめ発話意図タイプがアノテーションされたコーパスが必要であり,自動生成されるルールはそのコーパスに依存している.

平尾ら [8] は,スムーズな対話を目指して,雑談を対象に発話役割の推定を行っている.掲示板の対話を対象にデータ収集を行い,それらに対して6種類の発話役割を付与し,入力文と構築したコーパスに含まれる文の類似度から発話役割を推定する.

言語行為を対象とするコーパスに関連した研究では,コーパスの全体または部分に対して人手によるアノテーションを必要としており,構築にかかるコストが高くなる傾向がある.そのため,構築されるコーパスは,ドメインの限定や規模などに制約がかかることが多い.

本研究では,言語行為コーパスの自動構築手法を提案す

る.命名動詞を用いることで,ブログが話題とする様々な対象領域から発話を自動収集する.命名動詞をブログの書き手が付与したアノテーションとみなすことで,手作業では構築することが難しい,大規模かつ多様な文を含む言語行為コーパスを自動構築することができる.

3. 言語行為コーパスの自動構築について

本節では,言語行為がアノテーションされたコーパスの自動構築手法の提案と,その評価について示す.

3.1 収集する発話内容

本研究では,久保ら [5] を参考に,収集する発話の言語行為を決定した.久保らは,言語行為の発語内行為を型・群に分類して説明している.型は,言明型,行為拘束型,行為指示型,宣言型,感情表現型の5種類である.群は型の細分類で,各型に3~5種類の群が存在する.

各型の概要を以下に示す.

- 言明型
物事の状態を表す行為.話し手は,その命題で表されている事態が真であるという前提をもとに発話する.
例.明日は祝日だから学校はお休みです
- 行為拘束型
発話時点以降,命題内容が表している事態に,話し手自身をかかわらせようとする行為.命題内容が表す事態は発話時点から先のことで,遂行者は話し手でなければならない.
例.絶対に 大学に合格すると誓う
- 行為指示型
命題内容に含まれる事態や出来事が相手により実現されることを目指す行為.命題内容は,これから先のことであり,相手には命題内容を実現する能力があるという前提を必要とする.
例.今日の講義資料とノートを見せてくれませんか?
- 宣言型
その発話の遂行とともに新たな事態が発生する行為.その命題で表されている事態が自分の発話によって実現可能であることを前提としている.
例.運動会の開催を宣言します
- 感情表現型
話し手の発話時点における心的状態を言葉で表す行為.
例.今まで本当にありがとうございました

3.2 発語内行為命名動詞への着目

提案手法では,久保らが発語内行為の分析に用いた命名動詞を利用して,発話を収集する.久保らの研究では,命名動詞として機能する引用動詞を手掛かりとして,小説や新聞記事を対象に発語内行為を分析している.

命名動詞とは,発話が意図している言語行為を表現する

動詞のことである．以下に例を示す．

「レース用に使うバイクを提供してくれ」と頼んだ。

このような文がある場合，次のようなことが理解される．

- (1) この文は，発話の引用と，命名動詞から構成されている．
- (2) 発話は「レース用に使うバイクを提供してくれ」である．
- (3) 命名動詞は「頼む」である．
- (4) 命名動詞「頼む」は，発話が意図している言語行為が「行為指示型」であることを表現している．
- (5) 発話に対する言語行為の判定は，この文の書き手がやっている．したがって，命名動詞は書き手による言語行為アノテーションとみなすことができる．

このような対データ（発話と命名動詞）は，パターン抽出によって効率的に収集可能である．したがって，言語行為アノテーションの付与された発話コーパスを大規模に自動構築できると考えられる．本研究では，これらの命名動詞を利用して発話を収集することを考える．

3.3 発話内行為命名動詞と発話を用いたコーパスの自動構築

本研究では，きざしカンパニー^{*2}がクロールしたブログ記事（約 2 億 3000 万エントリ）を対象にしてデータ収集を行った．ブログ記事上で，発言を引用した部分であると考えられる引用括弧内の文字列と，命名動詞が連続して出現する部分を収集する．命名動詞は，久保ら [5] が定義したものを使用した．命名動詞の一致の判断には現在形，過去形を使用し，その命名動詞がサ変動詞である場合は，その名詞部分を用いた．発話の引用を表す文字列として“「」”と“””を用いた．このパターンによって，例えば，命名動詞「尋ねる」の発話収集は，“「」と尋ねる”，“「」と尋ねた”を，命名動詞「約束する」の発話収集は“「」と約束”と一致する文字列を収集することになる．発話へのアノテーションは，発話の収集に使用した動詞の形にかかわらず，現在形を使用した．収集に使用した命名動詞の例を表 1 に示す．

以下にブログ記事の一部を示し，命名動詞と発話収集までの流れを説明する．

ホテルに帰ってきてゆっくりしていたら、太郎がコンビニに行きたそうにしていたが、「夜は危ないから一人で出歩かない方がいい」と忠告した。太郎は少し悩んでいたが、やはり自分でも危ない

表 1 収集に使用した命名動詞の例

型名	含まれる命名動詞数	命名動詞
言明型	34	供述する，同意する，告白する
行為拘束型	24	約束する，遠慮する，脅す，断る
行為指示型	36	尋ねる，忠告する，頼む，相談する
宣言型	20	宣言する，却下する，認める
感情表現型	24	謝る，感謝する，挨拶する

と感じていたのか、明日の朝に行くことにしたよ
うだった。

記事中から，発話の引用部分を表す括弧と命名動詞「忠告する」を用いて，「夜は危ないから一人で出歩かない方がいい」と「忠告する」を収集する．

この場合，命名動詞は「忠告する」である．命名動詞「忠告する」は，発話が行為指示型であることを示している．すなわち，「夜は危ないから一人で出歩かない方がいい」という発話が意図している言語行為は行為指示型であることを表現している．したがって，ブログの書き手は，発話が意図する言語行為が行為指示型であると判断しているとみなすことができる．例では，引用した発話の話し手とブログの書き手は同一であると推測されるが，もし異なっていたとしても，書き手はその発話に対して適切だと判断した命名動詞を付与しているとみなせるため，同じ手法によって収集できる．

この手法によってブログが話題とする多様な対象領域から，発話の引用 + 命名動詞というパターンの機械的抽出によって，発話を収集することができる．

3.4 言語行為コーパスの評価

提案手法によって収集した対データの数を命名動詞別にまとめたものを表 2 に示す．

収集に使用した 138 種類の命名動詞の中で，収集した発話数が一番多かった命名動詞は「答える」の 71,715 件，一番少なかったものは「選任する」の 2 件であった．「選任する」や「ゆする」といった命名動詞で収集できた発話が少なかった原因としては，発話の引用 + 命名動詞というパターンで使用されることが少ない，ブログにおいて典型的でない動詞であったことが考えられる．

本研究では表 2 から，発話を 1,000 件以上収集することができた 66 種類の命名動詞を用いてコーパスを構築した．コーパスの構築に使用した 66 種類の命名動詞から合計で 438,747 件の発話を収集した．収集した数の平均は 6,647 件であった．命名動詞が属する型の内訳は，言明型 23 動詞，行為拘束型 8 動詞，行為指示型 18 動詞，宣言型 4 動詞，感情表現型 13 動詞だった．なお，構築したコーパスに含んでいない，発話の収集数が 1,000 件未満だった 72 種類の命名動詞からは，平均 297 件，合計で 21,373 件の発話

^{*2} <http://www.kizasi.co.jp/>

表 3 収集した発話例

型	命名動詞	発話
言明型	供述する	約5年間で2500万円以上の売り上げがあった
行為拘束型	脅す	民主党の候補者を当選させないと、それなりのことをさせてもらう
行為拘束型	誓う	二人で支えあい、助け合いながら、笑顔の絶えない幸せな家庭を築きたい
行為指示型	頼む	昨日の試合を録画してた生徒はビデオテープを貸してくれ
行為指示型	尋ねる	カラオケで歌ったり、話題に出すと話が盛り上がる日本人アーティストは誰だと思いますか？
感情表現型	感謝する	4万人近くの署名を集め、1月の閉店を回避できたのはみなさんの協力のおかげ

表 2 収集した発話数

No	型名	命名動詞	発話数
1	言明型	答える	71,715
2	言明型	説明する	41,429
3	宣言型	宣言する	21,791
4	行為指示型	尋ねる	19,276
5	言明型	返事する	17,011
<< 省略 >>			
62	感情表現型	呆れる	1,151
63	行為拘束型	遠慮する	1,139
64	行為拘束型	脅す	1,100
65	感情表現型	叱る	1,085
66	行為指示型	求める	1,069
<< 省略 >>			
134	行為指示型	せびる	17
135	行為拘束型	承る	13
136	行為指示型	たかる	10
137	行為指示型	ゆする	6
138	宣言型	選任する	2

を収集した。

収集した対データの一部を表3に示す。収集した発話自体には命名動詞に類する言葉は含まれていない。しかし、対データとして収集された命名動詞は発話に対して適切であることがわかる。

より定量的な評価のため、1,000件以上収集した命名動詞の発話から各型200件、計1,000件に対して、著者と他2名の合計3名の被験者で人手判定を行った。各型200件に含まれる命名動詞は同数に調整している。被験者にはガイドラインとして、型の定義、それぞれの型の命名動詞、その型に分類される例文(各型5文ずつ)とそれが分類された理由を提示した。人手判定では、各発話に対して最も適切な型を付与してもらった。

コーパスの評価として、人手判定による結果を正解とした自動判定の精度(Accuracy)を計算した。注意点として、人手判定が型を付与しているのに対して、構築したコーパスは発話と命名動詞の対データからなっている。そのため、アノテーションされているのは命名動詞である。精度を計算する際の自動判定の型は、久保らの定義に従い、対象の命名動詞が属する型とした。

精度は、判定可能な対象に対する精度と全てに対する重み付き精度の2種類で計算した。判定可能な対象に対する

精度では、被験者3名のうち、2名以上の判定が一致している発話を、判定可能な対象として評価データに使用した。評価データは916件となり、人手判定を行った1,000件の発話から84件が取り除かれた。全てに対する重み付き精度は、1,000件全ての発話を評価データとして、全員一致なら1,2人一致なら0.66,1人一致なら0.33,全員一致しなければ0のスコアを与えて精度を算出した。

結果を表4に示す。なお、被験者間の κ 値[1]は、被験者Aと被験者Bの κ 値が0.64,被験者Aと被験者Cの κ 値が0.50,被験者Bと被験者Cの κ 値が0.48となった。 κ 値の平均は0.54となり、中等度の一致(moderate)が得られた。

表 4 言語行為型自動判定の精度(Accuracy)

	精度
判定可能な対象に対する精度	0.553
重み付き精度	0.522

表4より、言語行為型自動判定の精度は、判定可能な対象に対する精度で0.553,重み付き精度で0.522となった。

3.5 言語行為コーパスの自動構築に関する考察

3.5.1 コーパスの自動構築手法について

提案手法では、66種類の命名動詞から平均6,647件、合計で438,747件の発話を収集し、大量かつ多様な文を収集することができた。

発話の引用に用いられる命名動詞の選択は、ブログの書き手のみの判断に拠っている。そのため、ブログの書き手の嗜好等によってある程度の揺れが発生し、久保らの定義した内容とは必ずしも一致しない可能性がある。しかし、被験者同士の κ 値は0.54と、中等度の一致(moderate)が得られている。このことから、命名動詞の選択には、ある程度の統一した指標が存在していることが示された。

自動判定の精度は、判定可能な対象に対する精度で0.553,重み付き精度で0.522と良好な値が得られた。これにより、発話の引用に続く形で用いられる命名動詞は、発話に対して適切な言語行為を表しており、自動構築した言語行為コーパスは適切な言語行為がアノテーションされているといえる。

3.5.2 収集した発話の曖昧性

構築した言語行為コーパスの誤りの原因の1つに、収集した発話の曖昧性が挙げられる。以下に例を示す。

- (1) 「あなたのために優勝します」と答えた
- (2) 「おい、絶対にこれだけはやれよ！」と叱った
- (3) 「あなたのおかげで遅刻せずに済んだ」と説明した

発話を収集するのに使用した命名動詞から考えると、(1)、(3)は言明型、(2)は感情表現型となる。しかし、発話だけを見ると、(1)は「誓う」などの行為拘束型、(2)は「命令する」などの行為指示型、(3)は「感謝する」などの感情表現型と捉えることもできる。これらのように、アノテーションされた命名動詞以外にも適切な命名動詞を持つ発話が存在する。人手判定では、最も適切な言語行為型を1つ選択してもらっているため、このような発話は精度の計算上、定義として誤っていなくても、誤りとされることがある。

構築した言語行為コーパスは、用途に応じて複数の命名動詞を組み合わせて1つのカテゴリとして使用することを想定している。そのように使用することによって、実用上の曖昧性が多少軽減することがある。例として、分類器への応用を考える。分類器を構築する際、まず、目的に応じた分類タイプを定義し、次に、それに沿った命名動詞を言語行為コーパスから選択する。このとき、定義した分類タイプから見て曖昧性の高い命名動詞は使用しない、お互いに混同する命名動詞を1つの分類タイプとして定義するなどの操作によって、コーパス自体の曖昧性の問題点は変わらないが、分類の曖昧性は下がることもある。このような理由から、本稿では、コーパス構築の際に曖昧性を考慮して命名動詞を絞るといった処理は行っていない。

3.5.3 収集できなかった発話の分析

提案手法は、発話の引用+命名動詞という文字列パターンによって発話を収集する。そのため、文字列パターンに一致しない発話は収集できない。収集できなかった原因を3つ示し、それぞれについて考察する。

- (1) 定義したパターンで使用される頻度が少ない命名動詞
- (2) 発話の引用の後に命名動詞が続かない
- (3) 発話に命名動詞が含まれている

(1)の定義したパターンで使用される頻度が少ない命名動詞は「せびる」や「承る」などが該当する。これらの命名動詞は、発話の引用を示す“「」”に続く形で出現する例がほとんど見られなかった。これは、他の形で使われることが多い動詞であるか、その命名動詞がブログにおいて典型的でないということが考えられる。これらの命名動詞の対データの収集は、現在の収集方法では改善が難しいと考えられる。しかし、本稿の目的は、精度の高い言語行為コーパスの構築であるため、出現数上位の動詞によって十分な量のコーパスが構築できるなら、重要な問題ではないと考えている。ただし、使用される頻度が低いことが、重

要度の低い言語行為であることを示している訳ではない。この原因によって収集できていない命名動詞については、それが重要な言語行為を表すものであるかどうかの検討が必要である。

(2)に該当する、発話の引用の後に命名動詞が続かない例を以下に示す。

- 「代表取締役として、これからも会社を支え続けて下さい」と取締役会で選任した

例の発話「代表取締役として、これからも会社を支え続けて下さい」は、「選任する」という命名動詞の対データとして収集することが適切である。しかし、発話の引用の後に命名動詞ではない語が続いてしまっているため、パターンに一致せず対データを収集できていない。この例は、パターンには一致していないが、書き手は発話に対して「選任する」という命名動詞をアノテーションしていると判断できる。そのため、収集するべき有用な対データである。現在は、発話の引用+命名動詞という文字列パターンの一致によって対データを収集しているが、この例を収集するためのパターンとして、係り受け構造に着目し、発話の引用が命名動詞に係っている場合に対データを収集するというものが考えられる。

次に、(3)の発話に命名動詞が含まれている例について考察する。例を以下に示す。

- 「賛成多数により田中氏を代表取締役に選任する」

この例も先ほどと同じく、「選任する」という命名動詞と対になる発話として収集することが適切である。しかし、命名動詞自体が発話に含まれてしまっており、発話の引用の後に命名動詞が続いていないため収集できない。この例では、発話自体に適切な命名動詞を含んでいるため、その命名動詞を抽出することで対データを収集できるように見える。しかし、発話に含まれた命名動詞は、その発話の言語行為を必ずしも決定しない。例を以下に示す。

- 「長い間会社を支え続けていた彼を、代表取締役として選任したと伝えて欲しい」と頼んだ

上記の例では、発話中に命名動詞「選任した」を含んでいるが、発話の対データとして収集する命名動詞としては「頼む」が適切である。発話に含まれる命名動詞は、発話を構成する一部分であり、発話の引用に続く命名動詞の使い方とは異なり、発話全体の言語行為を指しているとは限らない。このように、発話中に言語行為が含まれることは、発話の言語行為を決定する要因にはならないといえる。

発話中に言語行為が含まれる誤りでは、誤りである命名動詞より後に、適切な言語行為を表す語(例では「伝えて欲しい」)があらわれることが多く、ある程度のパターン化は可能だと考えられる。しかし、本稿では、精度の高いコーパスの自動構築を行うため、発話の引用+命名動詞のみで十分な対データが収集できた命名動詞のみをコーパスの構築に使用した。

表 5 言語行為タイプの概要

型	言語行為タイプ	概要	含まれる命名動詞
言明型	事実	物事の実事を述べる行為	供述する, 報告する
言明型	打ち明け	相手に対して今まで隠していた事実を明らかにする行為	打ち明ける, 告白する
言明型	同意返答	相手の発話に対して同意する行為	同意する
言明型	反対返答	相手の発話に対して反対する行為	反対する
行為拘束型	脅迫	発話によって相手の態度を変えようとする行為	脅す, 威嚇する
行為拘束型	約束	誓いなどで話し手自身の行動を制限する行為	約束する, 誓う
行為拘束型	断り	相手の依頼や勧誘に対しての断り行為	断る, 遠慮する
行為指示型	依頼	相手に対して頼みや誘いを行い, 何らかの行動を求める行為	頼む, せがむ
行為指示型	情報要求	相手に情報や意見などを求める行為	尋ねる, 相談する
行為指示型	励まし	相手を励ます行為	励ます, 激励する
感情表現型	感情	感情を表現している行為	感謝する, 挨拶する

4. 対話理解・生成システムの実現に向けた言語行為の判定

ユーザと会話するインターフェースの実現には、まず、ユーザの入力に対して、その発話が何を意図しているのかを理解する必要がある。本研究では、発話に対して言語行為の種類を判定することで、発話の意図を理解する。言語行為を正しく判定することで、会話型インターフェースは、ユーザの入力に適切な反応をすることができる。

本節では、前節で作成したコーパスを用いて、入力を発話、出力を言語行為とした分類器を構築し、評価を行うことで、言語行為の分類と構築したコーパスの有用性を明らかにする。

4.1 言語行為タイプについて

分類器の出力となる言語行為は、久保らが定義する群(型の細分類)を参考にして、会話に頻出する言語行為であるか、その言語行為に対して会話型インターフェースの反応を変化させる必要があるかという点から、会話型インターフェースに有効と考えられる言語行為(以下、言語行為タイプとする)を11種類定義した。

定義した言語行為タイプには、「事実」と「打ち明け」、「情報要求」と「依頼」というような、定義は類似しているが、会話を行う上で有用であると考えられる言語行為が含まれている。これによって、型分類や従来の発話行為に基づく分類 [6], [8] では同一と見なされていた言語行為を区別できる。

各言語行為タイプに、その言語行為タイプを代表するような命名動詞を2つずつ割り当てた。ただし、「同意返答」と「反対返答」については、適切な命名動詞が1つしかなかったため、1つとなっている。

本稿で構築したコーパスは、発話に対して命名動詞がアノテーションされている。そのため、分類器の構築に使用する際は、必要な分類タイプ(本節では言語行為タイプ)を定義した後に、それに沿うような命名動詞を割り当てて使

用することを想定している。これによって、命名動詞という小さな言語行為の差異を、用途に応じて柔軟に使用することができると考えられる。言語行為タイプに割り当てた命名動詞は、プログにおいて典型的であること、人によって用途が揺れていないことなどを考慮して決定した。

表5に、定義した言語行為タイプとそれに含まれる命名動詞を示す。

4.2 分類器の構築

入力された文を、定義した11種類の言語行為タイプに分類する分類器の構築を行った。分類器にはSVMを使用し、カーネルは、線形カーネルを用いた。実装にはLIBLINEAR[2]を使用した。素性として、文の名詞、動詞、形容詞、副詞、感動詞、bigramを用い、形態素解析にはMeCab [3]を使用した。訓練データとして、前節で構築した言語行為コーパスから、11種類の言語行為タイプそれぞれ1,000文を用いて、10分割の交差検定を行った。各言語行為タイプ1,000文の選択は乱数を利用して決定したが、言語行為タイプ内のそれぞれの命名動詞の発話数は同じになるように調整している。

4.3 分類器の評価

表6に言語行為タイプの分類精度を示す。言語行為タイプの分類は、マクロ平均で精度0.555、再現率0.545、F値で0.546であった。「依頼」、「情報要求」の分類精度は高い傾向があり、「打ち明け」、「同意返答」、「反対返答」は低い傾向が見られた。

また、分類された結果と人手判定による結果の精度(Accuracy)を計算した。人手判定は、各言語行為タイプ100文、計1,100件に対して、著者と他3名の合計4名の被験者で人手判定を行った。被験者にはガイドラインとして、言語行為タイプの定義、それぞれの言語行為タイプに含まれる命名動詞、その言語行為タイプに分類される例文(各言語行為タイプ2文ずつ)と、分類された理由を提示した。

精度は、判定可能な対象に対する精度と全てに対する重

み付き精度の 2 種類で計算した。判定可能な対象に対する精度では、被験者 4 名のうち、3 名以上の判定が一致している発話を、判定可能な対象として評価データに使用した。評価データは 752 件となり、人手判定を行った 1,100 件の発話から 348 件が取り除かれた。全てに対する重み付き精度は、全員一致なら 1、3 人一致なら 0.75、2 人一致なら 0.50、1 人一致なら 0.25、全員一致しなければ 0 のスコアを与えて精度を算出した。結果を表 7 に示す。なお、被験者間の κ 値 [1] の平均は、0.57 となり、中等度の一致 (moderate) が得られた。

より詳細な分析を行うため、判定可能な対象に対する精度の算出に用いた評価データを用いて、Confusion Matrix を作成した。Confusion Matrix は、各サンプルの正解に対して、何に分類されたかを表にまとめたものである。結果を表 8 に示す。表 8 では、行が正解の言語行為タイプを、列がその正解に対して分類された言語行為タイプを表している。一番右の列は、その行が指す言語行為タイプの判定可能とされた対象の総数を表している。「事実」、「依頼」、「打ち明け」は高く、「反対返答」、「断り」、「同意返答」は低いという傾向が見られる。

表 6 言語行為タイプの分類精度

	精度	再現率	F 値
事実	0.569	0.534	0.551
打ち明け	0.407	0.471	0.437
同意返答	0.418	0.586	0.488
反対返答	0.532	0.487	0.509
脅迫	0.654	0.484	0.556
約束	0.551	0.602	0.575
断り	0.520	0.500	0.510
依頼	0.642	0.631	0.636
情報要求	0.618	0.606	0.612
励まし	0.570	0.571	0.571
感情	0.625	0.518	0.566
平均	0.555	0.545	0.546

表 7 言語行為タイプの分類精度 (Accuracy)

	精度
判定可能な対象に対する精度	0.690
重み付き精度	0.560

4.4 分類器の考察

言語行為タイプの F 値は 11 分類で平均 0.546 となり、ある程度の分類精度を示した。特に「依頼」と「情報要求」は、類似した言語行為タイプでありながら、比較的高い分類精度を示した。「依頼」と「情報要求」は定義的には類似しているが、プログで発話する際には、文末表現に特徴が見られるなど、定義の類似と実際の発話の類似が必ずしも相関しないことがわかった。

人手判定を正解とした場合の分類精度は、判定可能な対象に対する精度で 0.690、重み付き精度で 0.560 となり、言語行為の分類において有用な値となった。これらの値から、自動構築した言語行為コーパスは、定義した分類タイプによく合致する命名動詞を選択することで、目的に応じた分類器が構築できることが示された。

表 8 の Confusion Matrix による分析では、「事実」と「打ち明け」を混同している傾向が見られた。特に「打ち明け」に関しては、全体の 4 割近くが「事実」に誤分類されていた。「事実」に誤分類する主な原因として、「打ち明け」は今まで隠していた事実を明らかにする行為であり、名詞や動詞などの現在使用している分類器の素性では、2 つのタイプの異なりを捉えられていなかったことが考えられる。一方、「打ち明け」の判定可能な対象の総数は 87 と比較的高い値となっている。人手判定では、発話の内容が隠していた事実かどうかを経験的に判断して言語行為タイプを判断するため、「打ち明け」は人にとって易しく、分類器にとっては難しい分類タイプだといえる。また事実は、打ち明けだけでなく、同意返答、反対返答、感情と様々なタイプに分類されやすいことがわかった。これは、他の言語行為タイプより多様な表現を含むためだと考えられる。

4.5 誤り分析

誤りとして、単独の発話だけでは判定できない例を以下に示す。

- (1) 私も子どもたちを一人で登校させるのは不安です
- (2) 必要としている人がいるならば、提供すべきだと思う
 - (1) は「反対返答」、(2) は「同意返答」の言語行為タイプが付与されている。しかし、これらの発話は、発話単独では他の言語行為タイプと混同してしまい、判定が不可能なことがある。(2) の発話が属する「同意返答」は、定義上、発話に加えて先行発話を前提とする言語行為である。そのため、現在のように単独の発話だけを入力とする場合、「反対返答」や「事実」と区別することが難しいことがある。(1) の例では、発話がどのような文脈で行われたかという情報が必要になる。もし、先行発話が「近辺で事件があり、犯人は捕まっていますが通常通り学校はあります」であれば、「反対返答」という言語行為タイプは適切であるといえる。しかし、対象が「犯人が捕まるまで、集団登校にして保護者を同伴させましょう」であれば、言語行為タイプは「同意返答」のほうが適切であると考えられる。同様に (2) の例も先行発話の内容次第では、「反対返答」となる可能性がある。

定義上は、発話単独では分類できないはずの「反対返答」と「同意返答」であるが、表 6、表 8 を見ると、ある程度の精度で分類できていることがわかる。特に表 8 では、2 つの言語行為タイプの混同は見られない。これは、実際の会話においては、反対する返答と同意する返答は、定型文が

表 8 判定可能な対象の Confusion Matrix

		分類結果											合計
		事実	打ち明け	同意返答	反対返答	脅迫	約束	断り	依頼	情報要求	励まし	感情	
正解	事実	42	15	16	16	6	4	5	6	4	5	10	129
	打ち明け	32	44	1	1	1	0	1	0	4	0	3	87
	同意返答	1	1	33	0	0	0	3	1	0	3	1	43
	反対返答	0	0	1	24	2	0	4	0	1	1	0	33
	脅迫	0	0	1	1	46	0	1	1	0	0	1	51
	約束	1	2	1	0	0	63	0	0	0	0	2	69
	断り	0	0	0	4	0	0	30	0	0	0	0	34
	依頼	0	2	1	2	3	5	2	82	2	14	2	115
	情報要求	0	3	11	2	2	0	2	0	66	1	2	89
	励まし	0	0	0	0	1	1	2	0	0	46	1	51
感情	1	2	0	0	0	0	2	1	0	2	43	51	
	合計	77	69	65	50	61	73	52	91	77	72	65	752

あるなど、名詞や動詞の形態素，n-gram に特徴があらわれているためである。

5. おわりに

本研究では、対話理解・生成システムの実現を目指し、言語行為コーパスの自動構築とその評価を行った。

構築したコーパスは、言語行為を特定する手がかりとなる命名動詞と発話の引用を用いることで、ブログが話題とする多様な対象領域から、発話と命名動詞の対データを収集することができた。

構築した言語行為コーパスの自動判定は、人手判定を正解として、判定可能な対象に対する精度で 0.553，重み付き精度で 0.522 と良好な値が得られた。これにより、発話の引用に続く形で用いられる命名動詞は、発話に対して適切な言語行為を表しており、自動構築した言語行為コーパスは適切な言語行為がアノテーションされているといえる。なお、このときの被験者間の κ 値の平均は 0.54 と中等度の一致 (moderate) が得られた。

構築した言語行為コーパスの有用性を評価するため、言語行為コーパスから会話に有効だと考えられる言語行為タイプ 11 種類を定義して、入力を発話、出力を言語行為タイプとする分類器を構築した。言語行為タイプの分類は、マクロ平均で精度 0.555，再現率 0.545，F 値で 0.546 という結果が得られ、一定の精度の分類ができることを示した。人手判定を正解とした場合の分類精度は、判定可能な対象に対する精度が 0.690，重み付き精度が 0.560 と良好な値を示した。これらの値から、自動構築した言語行為コーパスを利用した分類器は有用であるといえる。

本稿で行ったような、用途に合わせて言語行為タイプを定義し、定義に沿った命名動詞を取捨選択して分類器を構築する手法は、様々な用途に用いられる会話型インターフェースのより柔軟な構築に繋がる可能性がある。

本稿では、発話と発話内行為命名動詞を用いることで、

言語行為コーパスの自動構築を行えることを示した。また、構築したコーパスから言語行為タイプの分類器を作成することで、その有用性についての評価・考察を行った。

今後の課題としては、コーパスの自動構築の際に用いるパターン抽出方法の高度化と、命名動詞についての再検討を行っていくことで、より良い言語行為コーパスの自動構築手法の検討が挙げられる。

謝辞 本研究の一部は、科学研究費補助金基盤研究 C (課題番号 24500291) ならびに筑波大学図書館情報メディア系プロジェクト研究の助成を受けて遂行された。

参考文献

- [1] Jacob Cohen. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, Vol. 20, No. 1, pp. 37-46, 1960.
- [2] Rong-en Fan, Xiang-ru Wang, and Chih-jen Lin. LIBLINEAR: A Library for Large Linear Classification. *Journal of Machine Learning Research*, Vol. 9, pp. 1871-1874, 2008.
- [3] Taku Kudo, Kaoru Yamamoto, and Matsumoto Yuji. Applying Conditional Random Fields to Japanese Morphological Analysis. In *Processings of the 2004 Conference on Empirical Methods in Natural Language (EMNLP2004)*, pp. 230-237, Barcelona, Spain, 2004.
- [4] ジョン・R・サール著, 山田友幸監訳. 表現と意味: 言語行為論研究. 誠信書房, 2006.
- [5] 久保進編著, 阿部桂子, 越智希美子, 鈴木光代, 向井留実子共著. 発話内行為の意味ネットワーク: 言語行為論からの辞書的対話事例分析. 晃洋書房, 2002.
- [6] 熊本忠彦, 伊藤昭. コーパスに基づく発話意図タイプ決定ルールの自動生成. *情報処理学会論文誌*, Vol. 40, No. 6, pp. 2699-2707, 1999.
- [7] 翠輝久, 大竹清敬, 堀智織, 柏岡秀紀, 中村哲. 京都観光案内対話コーパスにおける対話行為タグの設計と分析. *情報処理学会研究報告: 音声言語情報処理*, Vol. 75, No. 8, pp. 39-44, 2009.
- [8] 平尾卓也, 松本和幸, 北研二, 任福継. コーパスに基づく雑談を目的とした発話役割同定. *言語処理学会第 18 回年次大会 (NLP2012)*, pp. 109-110, 2012.