

意味タグ付与の応用による伝達行為タグ付与の試み

野澤元・黒田航・井佐原均
独立行政法人情報通信研究機構 (NICT)

1. はじめに

近年、会話の書き起こしテキストに対して発話行為・談話タグを付与する試みが行われ、専用の特殊化された手法が開発されてきた。これに対して、本研究では複層意味フレーム分析における意味タグ付与の手法を応用することで、発話行為や談話機能を含んだ、広い意味での伝達行為タグの付与を試みる。具体的には、複層意味フレーム分析における意味タグを用いて、話し手と聞き手の間に成立する伝達の場合を記述し、それによって様々な水準での伝達行為を表現するというものである。タグ付きテキストにおいて、複層意味フレーム分析との共在が可能であり、言語表現の意味と遂行される伝達行為との関係性を表現することができる。

2. 伝達行為の記述

発話が何らかの伝達行為を遂行するという事実は、Austin (1962) によって初めて明示的に指摘され、それ以来、様々な方法でこのような伝達行為の記述が試みられてきた。しかし、言語表現の意味記述や統語記述に比べると、この分野の発展は遅れているように思われる。これには次にあげるような、いくつかの理由があると思われる。

第一に、伝達行為を記述するための手法が、特定の種類の伝達行為や、特定の分野における発話のみを対象とするような、特殊化したものになっていることがある。発話行為や談話機能といった伝達行為に、どのぐらいの種類が存在し、それらがどのような発話の単位によって遂行され、また、互いにどのような関係性があるのかについては、経験的な問題であり、必ずしも全てが明らかにされたというわけではない。従って、現段階ではより柔軟で制限の少ない手法によって、様々な分野における伝達行為の記述が試みられる必要がある。言語処理研究においても、発話行為や談話機能を記述する手法が開発されてきたが(荒木雅弘他 1999)、それらの枠組みは、記述上の制約が非常に強い。例えば、ある発話が遂行する伝達行為を、相互排他的に単一の種類として認定する必要はない。そのような伝達行為間の相互制約的關係は、量的な記述に対する分析の結果として抽出されるべきものであり、記述の前提とされるべきではない。現在の伝達行為の研究は、まだその

ような制約を発見してゆく段階にあると思われる。

第二に、発話とそれが遂行する伝達行為の関係性について、割り切った視点が必要だということである。多くの場合、発話が遂行する伝達行為は、発話の内在的な性質と見なされ、例えば、発話とその伝達行為の関係性が、「発話の力」といった比喩的な用語によって表現されるなど (Austin 1962)、ある意味で神秘化される場合さえある。しかし、発話が遂行する伝達行為とは、ある発話が行われている時に、同時に話し手と聞き手の間に成立する別の行為である。つまり、それは生起した発話を取り囲む環境の一部であり、その記述はそのような環境の記述の一環として行われるべきものである。本論では、ある話し手によって発話が生成され、ある聞き手によって理解される状況を、伝達の場合と呼ぶ。従って、発話が遂行する伝達行為の記述は、伝達の場合の記述の一部ということになる。

書き起こされた発話を分析し、それが遂行する伝達行為を記述する際、その伝達行為を発話の内在的な性質として混同する理由の一つは、伝達行為の認定のために、しばしば発話の表現の意味がその手掛かりとして用いられることである。文のような比較的大きな単位の言語表現は、多くの場合、時空間的に異なったいくつかの状況を指示する形態素や語の複合体として存在している。最も典型的には、文が生成され、理解される伝達の場合を指示する部分と、伝達の場合とは時空間的に切り離された離隔状況を指示

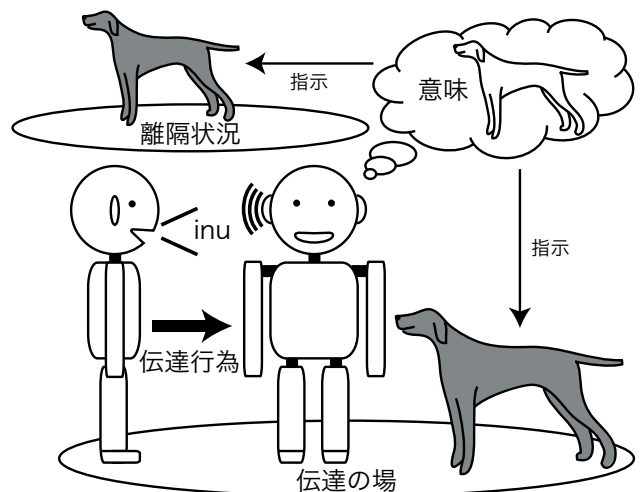


図 1

発話が表す伝達の場合と離隔状況の関係性

する部分を含んでいる。例えば、「遊びに行きたいね」の「遊びに行き」の部分は、未来における離隔状況を指示しているが、助動詞「たい」や終助詞「ね」は、伝達の場合における話し手の心情を指示している。ここで重要なことは、その発話が遂行する伝達行為の一つを、例えば、<提案>として認定した場合、それは発話の表現の意味を手掛かりとして、伝達の場合を推定しているに過ぎないということである。これはある発話が会話の中に位置付けられていて、先行発話との関係が考慮される場合であっても、書き起こされた発話のテキストに対する分析である限りは同じである。なぜなら、先行発話が遂行する伝達行為もまた、発話の表現の意味を手掛かりにして認定されるからである。

発話とそれが遂行する伝達行為の記述が蓄積されると、その記述に基づいて、両者の関連性について分析することが可能になる。しかし、そのような分析の結果として抽出される関連性は、仮定される聞き手の理解や解釈の機構のモデルに大きく依存することになる。発話が遂行する伝達行為の記述自体は、言語表現の意味と伝達行為の関連性を分析するためのデータとして位置づけられ、理解や解釈の理論とは分離されることが望ましい。

2. 複層意味フレーム分析

2.1. 複層意味フレーム分析の概要

複層意味フレーム分析 (Multilayered Semantic Frame Analysis: 以下、MSFA と略記) (黒田他 2005) は、Berkeley FrameNet (Baker *et al.* 1998) から着想を得て開発された意味分析の手法である。テキスト内の文が表現する事態水準の意味を分析し、分析によって抽出された意味は、テキストに対して意味タグを付与することで記述される。MSFA では特に事態水準の意味に着目しており、事態は意味フレーム (Fillmore 1982) というフレーム構造を持つと仮定している。意味タグは意味フレームの構造を反映し、事態のラベルとその事態における役割のラベルから構成される。分析の具体的な手順は、(1) 文の中で表現された意味フレームを特定し、(2) 意味フレームを想起させる要素である語や形態素を、支配項や喚起項として指定し、(3) その意味フレームの参与体を表す語や形態素を、フレーム要素として指定し、そして、(4) それぞれの参与体に意味フレーム内での役割を指定するという流れになる。

2.2. 会話に対する複層意味フレーム分析

MSFA は特定の分野における言語表現だけを分析の対象としているわけではない。しかし、研究の初

期における分析対象は、主として新聞記事のテキストであったため、現在の分析手法や表記法は、それに合わせた形式を取っている。そのため、いくつかの分野における言語表現の分析には適しておらず、特に会話を取り扱うことが出来ないという問題点があった。本研究では、MSFA のこれまでの表記法にいくつかの点で変更を加え、会話の分析に対応させた表記法を提案する。

2.3. 伝達の場合と離隔状況

第一に、これまでの MSFA には、文の中で表現された意味フレームが、伝達の場合と離隔状況のどちらを指示するのかについての区別がなかった。しかし、発話が遂行する伝達行為の記述をする場合、文の中で認定された意味フレームが、伝達の場合を指示するのか、離隔状況を指示するのかという区別は重要である。

1. 私はあなたに命令する。
2. 私はあなたに命令した。

例えば、(1) が表す一群の意味フレームには、<命令>フレームが含まれているが、それは伝達の場合における伝達行為を指示している。これに対して、(2) が表す<命令>フレームは、その参与体が(1)と同一でありながらも、過去に存在した離隔状況における伝達行為を指示している。このように、単に意味フレームを記述するだけでは、文が遂行する伝達行為としての<命令>フレームと、伝達の場合とは時空間的に異なる離隔状況における行為としての<命令>フレームを判別することが出来ない。

これまでの新聞記事のテキストに対する MSFA では、新聞記事の書き手である記者と、その読み手である読者の間に成立する伝達行為を指示する意味フレームとして、<報道>、<報告>、<語り>のフレームが特定されていた。もちろん、新聞記者も様々な意図を持ちながら記事を書くと思われるが、新聞記事における伝達行為は基本的に限定されており、また、それぞれの記事の中でも、話の流れによって伝達行為が大きく変化することはないと考えられる。しかし、会話においてはより多様な伝達行為が遂行され、それらは会話の推移に沿って頻繁に交替する。会話に対する MSFA では、そのような伝達行為の変遷を明示化するためにも、伝達の場合を指示する意味フレームと離隔状況を指示する意味フレームは区別される必要がある。

ここで注意しなければならないのは、会話における発話が遂行する伝達行為の記述は、あくまでも発話の環境である伝達の場合の記述であるということ

ある。分析者は発話に含まれる、形態素、語、語句、場合によっては文全体が表す意味を手掛かりにして、伝達の場の一部としての伝達行為がどのようなものであったのかを推定し、それを記述しているのである。その意味では、伝達の場における伝達行為の記述は、言語表現のいわゆる意味の記述とは異なる。ただし、後で述べるように、言語表現のいわゆる意味と、伝達の場の記述が一致することもある。

3. 君のせいで、僕は先生に怒られた。
4. ?君のせいで、僕は先生に怒られずに済んだ。

意味フレームが指示する状況を、伝達の場と離隔状況に区別した形での、(3)に対するMSFAを図2に示す。左端の列に形態素単位で分割した文が配置され、最上部の行には、伝達の場か離隔状況かという意味フレームの指示レベルが、また、その下にはフレーム名が表示されている。この分析では、<因果>、<非難>、<怒る>、<叱責>、<加害>、<語り>の六種類の意味フレームが認定されているが、そのうちで離隔状況を指示する意味フレームは、<因果>、<怒る>、<叱責>、<加害>の四種類であり、残りの<非難>と<語り>は伝達の場の記述である。

図2において名詞「せい」は、離隔状況における<因果>フレームの喚起項であると同時に、伝達の場における<非難>フレームの喚起項にもなっている。これは名詞「せい」が、ある状況における特定の事態間の因果関係を表すと同時に、その因果関係を理由とした非難の伝達行為に用いられることを示している。「せい」が単なる因果関係を表すだけでなく、非難の伝達行為を遂行するという事実は、(4)のような表現が不自然に聞こえることからわかる。「僕は先生に怒られずに済んだ」という表現が表す事態は、通常では感謝の理由となるが、「せい」は非難の伝達行為を遂行するため、両者が一つの文の

中で共在すると、概念的な不整合が生じるのだと考えられる。

(3)における名詞「せい」は、それぞれが伝達の場と離隔状況を指示する異なった意味フレームを喚起するが、(1)における動詞「命令する」の場合は、それが支配項となって喚起する意味である<命令>フレームが、伝達の場における伝達行為を指示している。つまり、(1)はその文自体の伝達の場を記述しているのであり、表現のいわゆる意味の記述と、伝達の場の記述が一致している。これは、明示的遂行発話に特徴的に見られる指示の構造であり、話し手は伝達の場を命令として記述することで定義し、命令行為を遂行しているのだと考えられる。

2.4. 主要な参与体の識別子化

これまでのMSFAの表記法は、単文の分析を前提としており、複数の文にまたがるテキストや、会話のテキストには対応していなかった。その原因は、ある言語表現において喚起される複数の意味フレームの間で、それぞれの参与体が同一指示であることを表現するために、参与体を表す形態素や語に対して役割を付与するという手法がとられていたからである。例えば、(3)における<非難>フレームの[非難する者]と、<叱責>フレームの[叱責される者]が同一指示であることは、両者が「僕」という同じ語にタグ付与されていることによって表現されている。つまり、語や形態素といった表現が、実質的に参与体の表示となっていたのである。

しかし、この手法では、分析の対象となる言語表現において、複数の形態素や語が同一指示を持っている場合、そのままでは、それらの同一性を表現することができない。本研究で提案するのは、表現の中に繰り返し現れる主要な参与体を識別子化し、フレーム要素である形態素や語に対してはその識別子をタグ付けするとともに、識別子に対して役割をタグ付けするという手法である。参与体を識別子化し

図 2

指示レベル	離隔状況	伝達の場	離隔状況	離隔状況	離隔状況	伝達の場	
フレーム名	因果	非難	怒る	叱責	加害	語り	
君	原因	非難される者				聞き手	
の	MARKER						
せい	EVOKER	EVOKER					
で	MARKER						
、							
僕	結果	非難する者		叱責される者	被害者	話し手	
は							
先生			怒る者	叱責する者	加害者		
に							
怒ら			GOVERNOR	EVOKER	害の内容		
れ			EXTENDER	EXTENDER	EVOKER		
た				EXTENDER			
。							

指示レベル	離隔状況	伝達の間	離隔状況	離隔状況	離隔状況	伝達の間
フレーム名	因果	非難	怒る	叱責	加害	語り
P1 (僕)		非難する者		叱責される者	被害者	話し手
P2 (君)	原因	非難される者				聞き手
P3 (先生)			怒る者	叱責する者	加害者	
君	P2	P2				P2
の	MARKER					
せい	EVOKER	EVOKER				
で	MARKER					
、						
僕	結果	P1		P1	P1	P1
は						
先生			P3	P3	P3	
に						
怒ら			GOVERNOR	EVOKER	害の内容	
れ			EXTENDER	EXTENDER	EVOKER	
た			EXTENDER			
。						

図 3

た形での、(3) に対する MSFA を図 3 に示す。

上部の指示レベルとフレーム名を表示する行の下に、三つの参与体の識別子とそれらが担う役割を表示するための行が配置されている。例えば、名詞「僕」は、<非難> フレームと<叱責> フレームおける同一の参与体 <P1> を指示しており、<P1> はそれぞれのフレームにおいて[非難する者]と[叱責される者]の役割を担っていることが表現されている。言語的な伝達の間において、最も基本的な伝達行為である<語り> フレームについては、参与体 <P1> と <P2> が、それぞれ[話し手]と[聞き手]の役割を担っていることがわかる。

会話の分析では、<語り> フレームの記述が話し手と聞き手の役割交替を表現する。離隔状況を指示する意味フレームを省略して、記述を伝達の間に限った、(5) の会話に対する MSFA を図 4 に示す。

- (5) A: 犯人は田中だ。
 B: いや、犯人は山田だ。

左から四番目の列は、「犯人は田中だ。」の範囲の表現が、参与体 <P1> と <P2> が、それぞれ[話し手]

と[聞き手]の役割を担う<語り> フレームの[語りの内容]であることを表している。また、右端の列は、「いや、犯人は山田だ。」の範囲の表現では、参与体 <P1> と <P2> が、逆に、それぞれが[聞き手]と[話し手]の役割を担っていることを表現しており、役割の交替が起きていることがわかる。

参考文献

荒木雅弘・他 . 1999. 「発話単位タグ標準化案の作成」, 『人工知能学会論文誌』, vol.14(2), pp.251-260.
 Austin, John Langshaw. 1962. *How to Do Things with Words*. 2nd ed. Oxford: Oxford University Press.
 Baker, Collin, et al. 1998. *The Berkeley FrameNet project*. in Proceedings of the COLING-ACL, Montreal, Canada.
 Fillmore, Charles J. 1982. "Frame Semantics", *Linguistics in the Morning Calm*. Seoul, South Korea: Hanshin Publishing Co., pp.111-137.
 黒田 航・他 . 2004. 「日本語のための意味タグ体系を定義する試み: FrameNet の視点から」, 『自然言語処理第 10 回大会発表論文集』, pp.148-51.

図 4

指示レベル	伝達の間	伝達の間	伝達の間	伝達の間	伝達の間	伝達の間	伝達の間
フレーム名	断定	主張	語り	否定	断定	反論	語り
P1 (A)	断定者	主張者	話し手	被否定者	被断定者	被反論者	聞き手
P2 (B)	被断定者	被主張者	聞き手	否定者	断定者	反論者	話し手
犯人	断定内容	主張内容	語りの内容	否定内容			
は	GOVERNOR						
田中							
だ							
。							
いや				GOVERNOR		反論内容	語りの内容
、							
犯人					断定内容		
は							
山田							
だ					GOVERNOR		
。							