

対話時の音声情報と言語情報に着目した発話意図の符号化

Encoding intention in utterances by using speech information and linguistic information

後藤 義貴[†]
Goto Yoshitaka

中平 勝子[†]
Katsuko T. Nakahira

北島 宗雄[†]
Muneo Kitajima

1 はじめに

コミュニケーションの場において対話関係は流動的に変化していき、その結果様々な雰囲気が存在する。対話関係の変化やその場に現れる特徴量を観測し、提示することにより、その場に介在しない第三者でもそのコミュニケーションの場を理解することが可能になると考える。

先行研究ではそのような特徴量を把握するために運用できるフレームワークが提案された [1]。このフレームワークではコミュニケーションの場で直接観測できる発話や身体動作、それらによって醸し出される、観察のすることができない雰囲気がコミュニケーション場を特徴付けるとされている。3要素の変化コミュニケーション変化して行くと考えられている [2]。本稿では発話に着目し、発話中の音声情報と言語情報から発話意図を符号化する手法の提案する。

対象とするコミュニケーション場はゼミやディスカッションの様な、お互いの立場を明確にして説明や意見交換を行う場とする。このような場の発話意図を符号化することで、その場のコミュニケーションの流れを目で見えるようにすることができ、コミュニケーションの理解に寄与できると考える。

2 コミュニケーション場の符号化

本稿で提案するコミュニケーション場の符号化手順は以下の通りである。手順3では文の先頭の形態素を文頭句、文の最後の形態素を文末句として、それぞれの形態素から分かる言語情報と音声情報で符号化を行う。また、文頭句では言語情報を、文末句では言語情報と音声情報両方を使用する。

1. 場で発言された内容の書き起こし及び発話時間の記録
2. 書き起こした内容の形態素解析及び音声情報取得
3. 解析又は取得されたものに対して本手法の適用

利用する言語情報は、発話内容を書き起こしたものから利用できる情報のことを指す。会話の始めから順に書き起こしを行うので、発話の順番なども言語情報から得ることができるものとして扱う。また、主に使用するのは形態素解析の結果得られた品詞情報やその活用形で、これらを利用することで発話の意図を推測することができる [2]。本手法は中でも接続詞と助詞に文を説明する機能があると考えて、活用をしている。接続詞は後文が前文に対して持っている意味的關係を予告する機能を持つとされている。この機能はは前文と後文の発話者が変わっても同じであると考えられるので、文頭に接続詞がくるこ

とにより、前の発話者に対しての態度を示すものであると考える。助詞には語の關係構成や文の成立に大きく寄与することが分かっている。中でも終助詞はそれが所属する文に対して意味的關係を持っている。この終助詞を手法に取り入れることで、文の意図を汲み取ることができると考える。接続詞と終助詞以外にも文の意図を表す語句の研究は進められており、それらの知見も手法の中には取り入れている [3]。

利用する音声情報は言葉の基本周波数、特に終助詞のイントネーションを指す。言葉の基本周波数に関連する研究は数多くされており、中にはイントネーションやアクセントの推測やそれらの変化による受け手の印象の変化の研究なども含まれている。本手法では基本周波数を用いて文末に現れる上昇句調や下降句調と言った文末句調の違いによる会話の TurnTaking を研究したものに注目し、手法の中に取り入れている [4]。

上記の要素を組み合わせる本稿で提案する符号化手法が表1である。

表1 符号化の方法

構成要素		符号名												
		対話開始	発話継続	受容	拒絶	転換	思考	言及	示唆	発話交代	自己疑問	単純疑問	複雑疑問	対話終了
対話関係	符号化の部分	文頭句	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		文末句								✓	✓	✓	✓	✓
	次の発話意図	受容										✓	✓	
		拒絶										✓	✓	
		転換											✓	✓
		不明												✓
次の発話者	同じ		✓							✓	✓	✓	✓	
	異なる									✓	✓	✓	✓	
間が一定以上					✓									
形態素	接続詞	順接の機能			✓									
		逆接の機能			✓									
		転換の機能				✓								
	終助詞	全般									✓	✓	✓	
		なねよか									✓	✓	✓	
		終助詞でない									✓	✓	✓	
フィラー								✓						
その他	石井らの知見がけど								✓					
句調	上昇句調									✓	✓	✓	✓	
	短い平坦句調									✓	✓	✓	✓	
	上記以外									✓	✓	✓	✓	

表1では、横に会話の符号を、縦に会話中の要素を配置している。符号化の際には、縦の符号名に対してその符号の成立に必要な要素を横にして、必要な要素にチェックマークが入っているかどうかを確認すればよい。そのチェックされた場所に全て当てはまるならば、その符号名を付けることができる。受容と拒絶に関しては、接続詞以外にも肯定の意図を含む“うん”や“はい”、拒絶の意図を含む“いいえ”といった誰が読んでも意図を読み取れると考える場合も符号を付与できるとする。発

[†] 長岡技術科学大学

話継続の符号は前発話と次発話が同じであればつけることが出来るが、それ以外にも発話の間に受容である“うん”や“はい”といった相づち等の文頭のみで判断できる符号しか入っていない場合も発話継続の符号を付けることが出来る。

3 実データでの符号化手法検証

提案した手法で符号化した会話をコーパスによってその妥当性を検証する。今回使用するコーパスには同立場の3人による会話が収録されている [5]。本手法が目指す適用場はゼミやディスカッションと言った意見交換の場なので、コーパスからそのような場面を抜き出し適用した。

3.1 適用例

以下ではコーパス中のいくつかの発話例をもとに、どのように符号化を行っているのかの説明する。

この場合は、Bの発話の文頭部が“もう”になる。これは品詞が副詞になるため、今回のモデルでは不明と符号化される。またBの発話の文末部は“ね”になる。これは品詞は助詞、その活用形が終助詞となるため、イントネーションでの判断を行う。このときのイントネーションは下降型なため、この会話の符号は文頭部“不明”文末部“言及”となる。また、Aの発話は“うん”の応答のみである。そのため、これは文頭部と考え、明らかな肯定なので、“受容”と符号を付けることができる。また、文末部は存在しないため、“-”が入る。そのため、符号の構成としては以下のようなようになる。

発話内容	符号名
B: もういいないもんね	B: 不明 言及
A: うん	A: 受容 -

3.2 時系列での変化

本手法によりコミュニケーション場内で行われた会話を記述し、それを先行研究のフレームワークに組み込むことで、発話から読み取れることができる場の変化を示す。

表2 発話から見られる特徴的变化

話者	発話意図		時間 [sec]	
	文頭	文末	前発話からの経過時間	発話時間
B	不明	示唆	0.36	2.44
C	受容	-	0.34	0.63
A	受容	-	-0.58	1.04
B	受容	-	-0.65	0.22
C	受容	-	1.52	0.34
B	不明	不明	0.14	0.98
C	転換	言及	0.08	1.16
B	受容	-	-0.08	0.36
C	発話継続	不明	0.70	0.895

表2は話が一度止まり、その後再び会話が始まり、Cが説明を

行う様子を示す時系列データである。この時系列データから読み取ることが出来るのは、会話の中でどのような意図を持った発話が行われているかである。この会話では、始めから中頃にかけてBの発話に対してAとCが“受容”の反応を示したのが分かる。話が止まり、話題を変えるためにCが“転換”の意図を示し、少し他と比べると長く発話しそれに対してBが“受容”の反応を示し、それを受けてCが“発話継続”の意図を示した。これら一連の流れから、一度は会話が全員が受容する形で終了し、その後新しい話題をCが提示しBがその話題を受け止めたことで話が新しく進むという、コミュニケーションの形が見いだせる。

4 結言

今後この符号化手法を厳密にすることで、会話の流れがどのようなものかを時系列で表すことができるようになり、コミュニケーションの変化を容易に把握することが可能になると考える。また、これを前述したコミュニケーションのフレームワークに組み込むことで、対話時の音声特徴と言語情報以外にも身体動作や雰囲気といった要素を含めたコミュニケーション全体の時系列データを記述することが出来るようになり、会話中にどのような要素を含んだことを行えば今より活発な会話になるのか、話題内容が変わっていくのか、そういったコミュニケーション変化を引き起こす要因を知るための手がかりを導くことができるようになる。本稿では記述する方法を示したが、それが実際の場においてどのくらいの精度を示すのかを検証を行っていない。その検証を今後行っていく予定である。

参考文献

- [1] 五十嵐 涼, 中平 勝子, 北島 宗雄: トリオンによるコミュニケーション場の雰囲気記述., 情報処理学会第76回全国大会講演論文集, 4-241-242, 2014
- [2] 後藤 義貴, 五十嵐 涼, 中平 勝子, 北島 宗雄: ヒューマンコミュニケーション場の普遍特徴量の検討, FIT2014 J-019, 2014
- [3] 石井カルロス寿憲, ニック・キャンベル: 句末の機能的役割, 日本音響学会 2004 年春季研究発表会, Vol. 1, 235-236, 2004
- [4] Carlos Toshinori ISHI: Perceptually-Related F0 Parameters for Automatic Classification of Phrase Final Tones, IIEICE transactions on information and systems E88-D(3), 481-488, 2005-03-01
- [5] Den, Y. & Enomoto, M: A scientific approach to conversational informatics: Description, analysis, and modeling of human conversation. In Nishida, T. (Ed.), Conversational informatics: An engineering approach, pp. 307-330. Hoboken, NJ: John Wiley & Sons, 2007