

## 日本語の自由会話における発話間の因果関係の推定

西澤 信一郎

横浜国立大学工学部

中川 裕志

横浜国立大学工学部

## 1 はじめに

一般に、談話理解を行なうためには広範囲な情報が必要とされる。発話中に登場する各単語の意味情報や概念の階層情報など語彙レベルでの情報から、会話における話題に関する背景知識や発話のつながりなどいわゆる文脈情報までを、我々は他人との会話において用いており、特に文脈情報を適切に扱うことが計算機上での談話理解には必要であると考えられる。ところで、我々が会話を行なう際には、いわゆる文脈情報を会話自身の中から抽出・生成していることが多い。特に二つ以上の発話間の関係(順接であるのか逆接であるのか、など)つまり談話構造[Hob90]からは、話者と聞き手との間での思考の流れや発話中で述べられる事柄の間の関係(因果関係にある事柄なのか、など)を抽出することが可能であると考えられる。このことから、会話の持つ談話構造を理解することは、談話理解、特に文脈情報の生成、という点から必要になるものと考えられる。そこで本稿では、因果関係を記述するとされる接続(助)詞によって作られる談話構造を対象とし、複数の参加者による日本語の自由会話中でそれが実際にどのような構造となるのか、またそれを機械的に会話中から発見するにはどのようにすればよいのか、について検討することとする。

## 2 因果関係を表す談話構造

## 2.1 自由会話コーパス

現在、対話コーパスとして、地図課題・スケジュール設定・パズル解法などのある決まった目的(task)の元での会話を書き起こしたものが存在する。これに対して我々は今回「飲み会」の席上での雑談(自由会話)を書き起こしたテキストをデータとして利用した<sup>1</sup>。その例を図1に挙げる。この会話データは大学のあるサークルでの飲み会の席上で録音されたものを、そこに同席した者がテキストに書き起こしたものである。書き起こす際に、次のような処理をしている。

1. 発話の切れ目の認識。
2. 会話内容によるセグメント分け。
3. 話し手と聞き手のデータ追加。

<sup>1</sup>今後このテキストを「コーパス」と呼ぶ。

## 会話 24

1. O → P あのね、これでもいいんじゃない、
2. P → O わかった、
3. O → P えー、嘘で言ったんだよ、
4. E → O 何、
5. O → E だって、牛乳入れろって言ってたらさー、
6. G → G 何か酒飲みたいなー、
7. K → G あっ、ありますよ、
8. G → K それ何、
9. E → G モルツ、
10. P → G ウイスキー、
11. E → G うまいよ、

図1: 自由会話のテキスト例

## 4. 発話の末尾の調子のデータ追加。

例えば、上の例では、この一連の発話群が“会話24”というセグメントに属することを示す。また、「K → G あっ、ありますよ、」という発話データの場合、文頭にある“K → G”が、この発話の話し手がKで聞き手がGという人物であることを示している。また、文末の“、”が、この文が下がり調子で発話されたことを示している。

## 2.2 因果関係の表現

本稿では、接続(助)詞によって述べられる因果関係表現について、(1):因果関係の前件(原因/理由)および後件(結果)を述べる発話と、因果関係を表す接続(助)詞とがどのような位置関係(構造)をとるのか、(2):因果関係を述べている発話部分を会話中から推定するためにはどのような手順が必要か、を検討する。今回用いたコーパス中には、例えば次のような会話が因果関係を表現している例として存在する<sup>2</sup>。

- a. P → \* 布団ひいてきます、
- b. B → P うん、
- c. A → \* いや、ベットの部屋よかったな、
- d. ΔE → A ベットの部屋いいよね、
- e. ΔE → A いったら寝れんどもん、
- f. E → A ばたって、

このような会話のうち、本稿では、次の条件に当てはまるような会話を対象とする。

<sup>2</sup>この例では、第四文が後件を、第五文が前件を表す。

1. 接続詞「だから/だって」および接続助詞(相当)「から/のだから/のだもの」<sup>3</sup>が用いられており、これによって因果関係が表現されている。
2. 因果関係の前件および後件が同一人物によって発話されている。

これらの条件に当てはまる発話は、「だって」や「から」などの語によって二つの発話が因果関係を表現するような談話構造をとっていると考えられる。そこで、まず、これらの接続(助)詞による因果関係表現がどのような談話構造をとるのか、について検討する。

まず、接続詞が用いられる場合であるが、接続詞が、接続すべき文と文との間におかれる性質の語であることから、接続詞を含む発話の意味内容を考慮に入れた場合、(a):「前件. 接続詞-後件」あるいは(b):「後件. 接続詞-前件」、つまり接続詞を含む発話とその前方に位置する発話とが因果関係をなす、という構造がまず考えられる。

ところで、これらから外れる構造として、前後の発話の位置関係を入れ換えた、(c):「接続詞-後件. 前件」あるいは(d):「接続詞-前件. 後件」という場合が考えられるが、これらは完全に不適格な構造なのであろうか。

この二通りの構造のうち、(d)は因果関係を記述する際に意味的に自然であると考えられる「前件(原因/理由)-後件(結果)」という、前件と後件の生起の時間順序に従った発話である。そのため、人間がこの順序の発話を聞いた場合には、解釈の余地がある程度は残っているのではないかと考えられる。一方、(c)では、接続詞が二つの発話を結合する位置に存在せず、また二つの発話が「前件(原因/理由)-後件(結果)」という順序にあるわけでもない。よって、因果関係を述べる発話をこの構造のようにとらえることは(d)の場合よりも難しく、ほぼ不可能なのではないかと考えられる。

次に、接続助詞が用いられた場合は、接続助詞を含む発話の前件を述べることから、前件および後件を述べる発話の順序として、(A):「前件-接続助詞-後件」あるいは(B):「前件-接続助詞. 後件」、(C):「後件. 前件-接続助詞」という構造が考えられる。このうち、本稿では発話の末尾に接続助詞が現れる(B)および(C)の形式について扱うこととする<sup>4</sup>。さて、接続助詞の構文的な機能は、二つの節を何らかの関係で結びつけることである。(B)および(C)では前

件と後件が二つの発話にわかれるが、接続助詞が使われているためこれらの発話は文ではなく節だと考えられる。よって、前件と後件の順序が入れ替わる(C)はいわゆる倒置現象となり、非常によく見られる現象だといえる。つまり、接続助詞は前件と後件の位置関係に影響を与えることはなく、(B)と(C)の構造は両方とも因果関係として認識可能なものであるといえる。

以上の考察は次の原則にまとめることができる。

**原則1** 因果関係の表現に接続詞を用いた場合、因果関係の前件(原因/理由)および後件(結果)を述べている発話の位置関係は、次の条件を満たす。

1. 接続詞を含む発話は、その前方にある発話と組になって因果関係を表現する。前件/後件がどちらの発話となるかは接続詞の種類による。
2. 接続詞を含む発話とその後方にある発話は、これらの発話が意味内容として「前件(原因/理由)-後件(結果)」の順に並んでいる時に限り、組となって因果関係を表現することが可能である。

一方、因果関係の表現に用いる接続助詞は、前件および後件を述べている発話の位置関係を特に制限しない。

この原則について、実際にコーパスを用いて検証してみる。図1に例を示したコーパスは全部で1980の発話からなる。このうち、先に述べた条件に当てはまる発話は45例存在した。ところで、原則1および接続詞などの性質[森田80]から、これらの例は以下のような構造をとっている必要があると考えられる。

- |                   |                    |
|-------------------|--------------------|
| (a1): 前件. だから-後件. | (b1): 前件-接続助詞. 後件. |
| (a2): 後件. だって-前件. | (b2): 後件. 前件-接続助詞. |
| (a3): だって-前件. 後件. |                    |

そこで実際に上で述べた45例を対象とし、著者を含む数人で手作業により分類した結果を次に示す。

- |            |            |
|------------|------------|
| (a1): 4例.  | (b1): 5例.  |
| (a2): 11例. | (b2): 24例. |
| (a3): 1例.  |            |

なお、原則1によって排除されるような構造である「だから-後件. 前件」という会話は今回用いたコーパス中から発見することはできなかった。

### 3 談話構造の推定

接続(助)詞によって表される会話中の因果関係は、原則1に従った構造をとることを前節で述べた。これを利用すると、本稿で扱っている「だから/だって」

<sup>3</sup>本稿では、「から/のだから/のだもの」をあわせて接続助詞として扱うこととする。また、「から」などの表現には、そもそも因果関係を表さないような用法が存在する[白川95, 野田95]ため、ここでは[白川95]で述べられている方法に準じてテストを行ない、45例全てが因果関係を記述していることを確認した。

<sup>4</sup>これらのうち(A)は通常の複文である。

および「から / のだから / のだもの」が用いられた発話の場合、因果関係を記述している部分をコーパス中から取り出すための手順を次のように仮定することができる。

仮定 1 コーパス中から因果関係を記述している部分を見出すための手順を以下のように仮定する。

1. 接続(助)詞を含む発話をコーパス中から発見する。この発話を  $U_0$  とし、この発話の話者を  $S_0$  とする。
2.  $U_0$  を検索の起点とし、 $S_0$  による発話を以下のようにしてコーパス中から検索する。
  - (a)  $U_0$  が「だって」あるいは接続助詞を含む発話の場合は、そこから前後へコーパスを検索する。
  - (b)  $U_0$  が「だから」を含む発話の場合は、起点から前方へのみの検索を行なう。
3.  $U_0$  からコーパス中で最も近い位置で発見された発話が、 $U_0$  と因果関係をなす発話となる。

この手順がどのように働くのかについて、接続(助)詞別に以下で述べる。

### 3.1 「だから」による因果関係

前述の(a1)「前件. だから-後件」に当てはまる例が4例存在する。これらはすべて、仮定1に従って該当部分を見出すことができる例であった。なお、これらのうち、次に挙げる例は原則1を裏付ける例であると考えられる。

- (2) a.  $\Delta A \rightarrow B$  木のアンバランスだとどこ取ってもいいじゃん
- b.  $B \rightarrow A$  ああ
- c.  $\Delta A \rightarrow B$  だからねー、どんどんどんどん高くなるんだよ
- d.  $\nabla A \rightarrow B$  クロのだとあれじゃん、サイコロ転がしてやったじゃん
- e.  $B \rightarrow A$  うん
- f.  $\nabla A \rightarrow B$  だからねー、相当低い時でも駄目な時は駄目なんだよ
- g.  $A \rightarrow A$  あれまああったよ

この例では、「 $\Delta$ 」同士、「 $\nabla$ 」同士が因果関係を表している。なお、第三文から最も近いAの発話は第四文であり、第六文から最も近いAの発話は第七文であるが、これら同士は因果関係を表していない。つまり、原則1にあるように、「だから」を含む発話とその後方の発話とは因果関係を表すことができない、ということである。

### 3.2 「だって」による因果関係

前述のように、(a2)「後件. だって-前件」に当てはまる例が11例、(a3)「だって-前件. 後件」に当てはまる例が1例存在する。会話例を次に挙げる。

- (3) a.  $E \rightarrow I$  磯ちゃん、飛びたい
- b.  $I \rightarrow E$  いやー
- c.  $Q \rightarrow I$  私がさ、びよーんってやってあげる
- d.  $I \rightarrow Q$  じゃ、一緒にやろう
- e.  $\Delta E \rightarrow *$  大丈夫だよ
- f.  $\Delta E \rightarrow *$  だって、那須いった時は女の子が嫌やりました

なお、12例のうち、仮定1では人間によって認識される因果関係が発見できない例が3例ある。これらは、「だって」を含む発話の直後に別の発話者による「～よ」という発話がなされている、という共通した特徴がある。次にその一例を示す。

- (4) a.  $\Delta M \rightarrow M$  何か、去年はごくごくごくって飲んだ気がする
- b.  $B \rightarrow M$  いや、一気だよ、一気
- c.  $B \rightarrow M$  こんぐらい
- d.  $\Delta M \rightarrow B$  瓶がだって、3人で10分ぐらい
- e.  $B \rightarrow M$  3人じゃねーよ
- f.  $B \rightarrow M$  5人だよ
- g.  $M \rightarrow B$  えっ、違う、違う

「 $\Delta$ 」同士が因果関係を表すが、第四文から等距離で逆方向の第七文に、やはりMの発話があり、これらの間には、Bの「～よ」という発話が存在している。このような例から、次のような補足手順が得られる。

「だって-前件」という発話(話者  $S_0$ )から後方への検索中に、 $S_0$  以外の話者による「～よ(+な/ね)」という発話が発見された場合には、後方への検索を打ち切る。

おそらく、終助詞「よ(+な/ね)」の力によって、因果関係を表すような結び付きが切断されるのであろう。

### 3.3 接続助詞による因果関係

前述のように、(b1)「前件-接続助詞. 後件」に当てはまる例が5例、(b2)「後件. 前件-接続助詞」に当てはまる例が24例存在する。例を次に示す。

- (5) a.  $M \rightarrow K$  俺、まだ入ってるの、これ
- b.  $K \rightarrow M$  ほら、4種類もある
- c.  $\Delta K \rightarrow M$  あっ、これ、これ飲んで下さい
- d.  $\Delta K \rightarrow M$  おいしいから
- e.  $M \rightarrow K$  俺が酔っ払うとどうなるか知らないな

なお、会話例のうち、仮定1では人間によって認識される因果関係が発見できない例が5例存在した。まず最初の例は、接続助詞を含む発話の先頭に「(それ)で」という接続詞が用いられている次のような例である。

- (6) a. N → F カルピスと焼酎じゃない  
 b. ΔN → F で、これ濃いよーって言われて、ソーダ入れちゃったから  
 c. E → N カルピスソーダ割り  
 d. ΔN → E さっきのソーダの味がするでしょ

同一話者の発話として隣接しているのは第一文と第二文だが、因果関係を表すのは第二文と第四文である。おそらく、第二文の「で、これ濃いよーって…」の「(それ)で」が、『第一文と第二文の流れが逆転せず、従って因果関係を表現しない』、ということを示しているものと思われる。これから、次のような補足手順が得られる。

「前件-接続助詞」という発話の先頭に「それで/そして」など話題の転換を表す語が用いられているならば、そこから前方への検索は行なわなくてもよい。

これは、「それで/そして」によりそれ以前での発話における話題から、因果関係を表すまとまりでの話題へと話題がシフトするためであると考えられる。

次の例は、接続助詞を含む発話から最も近い位置にある発話が「から」を含む、という例であり、同じ特徴を持つ会話例が全3例存在する。

- (7) a. Δ△E → E 止めてくるかな  
 b. N → F 磯さんはおねんねしてんの  
 c. F → N 寝てる  
 d. A → \* 明日弾けるんじゃない、明日  
 e. Δ E → \* 磯チャン弾けると大変だからな  
 f. ΔE → \* 弾けると気持ち悪くなっちゃうからね、磯さん

上記の会話例は、第六文で一度会話が途切れる。つまり、第六文以降にはEの発話が存在しない。このことから、第五文から最も近いEの発話で、「から」を含まない発話は第一文である。また、第六文から最も近いEの発話で、「から」を含まない発話も第一文である。これらから、次のような補足手順が得られる。

「前件-接続助詞」という発話(話者 $S_0$ )から最も近い位置にある、 $S_0$ による発話が「から/のだから/のなもの」を含む場合は、その発話を後件とはせず、さらに検索を進める。

さらに、発話の聞き手が関係するような、次のような例も1例存在した。

- (8) a. ΔI → R はーい、大丈夫  
 b. ΔI → R ちゃんとやるから  
 c. I → I パーテンドーと化してきた私だった

第二文に隣接しているI自身の発話が第一文と第三文と二通りある。ただし、第二文と同じくRへの発話であるのが第一文であるから、第二文と第一文とが因果関係を表すものと考えられる。これから、次のような補足手順が得られる。

「前件-接続助詞」という発話 $U_0$ (話者 $S_0$ , 聞き手 $H_0$ )から前後へ同一距離の位置に、 $S_0$ による発話が存在する場合は、それらのうち聞き手が $H_0$ である発話が後件となる。

以上のように、仮定1で述べた手順がコーパス中で実際にどのように働くのかについて検討した結果、前節で述べた45例のうち、35例についてはこの手順によりコーパス中からその位置を決定することが可能であった。さらに、上述した補足手順を考慮すると、45例全てのコーパス中での位置が決定可能であることが判明した。

#### 4 おわりに

前節で検討した手順(仮定1)は、コーパスの表層情報を用いるものである。また、四種の補足手順についても、発話の意味内容は問わない。このことから、計算機での処理によりこれらの手順を実現することは比較的容易なのではないかと考えられる。

ところで、コーパス中には、接続(助)詞が用いられている次のような例も存在する<sup>5</sup>。

- (9) a. A → B 何か、3種類くらい売っててこれが一番良かったみたい  
 b. ΔF → B 人体実験やんないの? 今日  
 c. ΔB → F だってもう眠いもん  
 d. B → F 何か最近酒弱いみたい

このように、異なる人物による発話が組となって因果関係を述べるような例はコーパス中に45例存在する。このような場合の談話構造や、その発見の手順などについての議論は、次の機会に譲る。

#### 参考文献

- [Hob90] Jerry R. Hobbs. *LITERATURE AND COGNITION*, pp. 83-114. No. 21 in CSLI Lecture Notes. CSLI, 1990.  
 [白川95] 白川博之. 理由を表さない「カラ」. 仁田義雄(編), 複文の研究(上), pp. 189-219. くろしお出版, 1995.  
 [野田95] 野田春美. 「のだから」の特異性. 仁田義雄(編), 複文の研究(上), pp. 221-245. くろしお出版, 1995.  
 [森田80] 森田良行. 基礎日本語2. 角川小辞典. 角川書店, 1980.

<sup>5</sup> “Δ” 同士が質問応答であり、質問文の内容が後件、返答文の内容が前件を表す。