

条件付き相互行為と話し言葉の会話分析に関する一考察 ～場面別の対話システム向上を目指して～

A Consideration on Conditional Interaction and Conversation Analysis of Spoken Language - Toward Improving Dialog System by Scene -

○ 太田 博三¹

Hiromitsu Ota 1¹

¹放送大学 教養学部

¹The Open University of Japan

Abstract: In recent years, interactive systems and dialogue generation in natural language processing have attracted attention. Particularly due to the spread of chat bots to call centers, accurate human interactive response is required. On the other hand, qualitative interactions in sociology's ethnomethodology and discourse analysis / conversation analysis are beneficial. Therefore, once again, using the Japanese language learner conversation data corpus of the National Institute of Japanese Language, it is a consideration aiming to verify the effect and apply it to the tendency of dialogue collapse and dialogue generation.

1. はじめに

1.1 研究の背景と目的

2020年のオリンピックに向けて、スマートスピーカーが家庭に普及し、自動運転が実用化されようとしている。人工知能が人間に代替し仕事を奪うとも言われている。既にPepperなどのロボットやチャットボット(Chatbot)が商用化されている。Pepperはイオンモールなどで店の案内をし、チャットボットはコールセンターの一部機能を代替しようとしている。ここで、データの数が足りないという切実な問題を有している。ディープラーニング(Deep Learning)による機械翻訳の精度は極めて高いが、対話応答の自動生成は不十分である。そこで本稿では、定性的なフレームワークのアプローチでサンプルサイズの不足分を補うことを試みたものである。

次に、エスノメソドロジーや会話分析での隣接対などの定性的な規則や傾向を、対話応答に適用してみたいと考え、追及してみた。

現行の対話システムはIF文による制御構造で開発されており、必ずしも隣接対や雑談の種類が明示的に、もしくは十分に活用しきれている訳ではない。

ディープラーニングの分野では、データの増幅(Data Augmentation)が発達し、転移学習も少しずつ発展してきた。これらでスケール化させ、実用化に結びつける動きがあるものの、画像認識のコンピュータビジョンに留まっているのが現状である。

1.2 研究の新規性

本研究の新規性はデータ量を増やすと大きな相違が生じるかを考察し、定性的な傾向規則性を対話システムに取り入れられるかを試みたことにある。

本稿では、誰もが入手可能なデータである日本語学習者の会話データを用いることで、統計的な有意性やサンプル数より、日常生活の感覚でわかることを重視したものである。次に、[質問]-[応答]や[申し出]-[受諾/拒否]などの隣接ペアの類型が上記のデータにどのくらいあてはまるかなど、計量化し、実証的な知見が得られれば、対話応答の次のトピックや発話文の予測が可能となる。隣接ペア以外に、主に以下の5つが有用と考えられる。

- 1) Yes/Noの応答詞 : あー, うん, えー, そう
- 2) あいづち : んー, はい, はー, えー
- 3) 言いよどみ : んー, あー, えー

4) 呼びかけ : ね, ねー

5) フィラー : あの一, その一, えーと, えっと

今後, これらを分析し, 構造化したものを, 対話システム等に追加することで, ユーザーの満足度の向上につながると思われる.

1.3 研究の主な手法

基礎集計を中心に行いながら考察する. 国立国語研究所の提供している「日本語学習者会話データベース」を用いて集計を行う. 隣接ペアは本稿で定義する種類のものに限定し計量化する. 次に, それらのペアが全体の会話の促進になっているかなどを考察する. また, その隣接ペアの前後, もしくは直後の発話が修復に向けてのものか, 完全に破綻しているが強引に会話を続けたものであるのかも含めて, 定性的な判断を行う. 日本語学習者の能力評価をするという特殊な場面設定のため, 難易度がワンランク上の「突き上げ (Thrust up)」が用いられているのが特徴である. この意図的に難しい面接試験での質問をすることで, 外国人の日本語能力を見極めるという方法論を用いている. いわゆる, とっさの対応が本能とみならず類のものである.

1.4 用いたデータセットについて

国立国語研究所が公開しているコーパスの中の1つである「日本語学習者会話データベース」(図1)を用いる. またKYコーパスも同様の趣旨で作られたものであり, 適宜, 用いた. 1990年の入管法の改正により, 日本の社会状況に応じて, 外国人受入れの適切な方策が必要となり, 日本語学習を必要とする住民(言語生活者)の需要に見合った言語教育の展開が期待されていた. ACTFL-OPI(全米外国語教育協会認定の面接式口頭能力テスト)を活用し, 日本語を用いた自然な会話に限りなく近い対話で構成されている.

ACTFL(全米外国語協会)によるOPI(Oral Proficiency Interview Test)に基づいており, 日本語OPIは1993年に発足し, 15年近く経過している. ここでの判断尺度は, 次の4つに区分されている.

- 1) 超級 (Superior)
- 2) 上級 (Advanced)
- 3) 中級 (Intermediate)
- 4) 初級 (Novice)

これは「日本語学習者会話データベースの利用手引き(平成22年5月国立国語研究所)」によれば, 言語運用能力は10種類の階級に区分されている(表1). 対話の SCRIPT は, インフォーマント(日本語学習者/データ提供者)とテスター(面

接者)とからなり, 30分ほどの対話形式で構成されている.

また上記の10段階のOPIレベルや性別, 年齢, 出身国などを選択することができる. 検索条件を設定してダウンロードすると, 文字化(一部, 音声化)されたSCRIPTが入手でき, 有用である.

表1-. OPI 能力区分表

区分	OPIレベル	階級	OPI評価
1	超級 (Superior)	1	超級
2	上級 (Advanced)	2	上級-上
	〃	3	上級-中
	〃	4	上級-下
3	中級 (Intermediate)	5	中級-上
	〃	6	中級-中
	〃	7	中級-下
4	初級 (Novice)	8	初級-上
	〃	9	初級-中
	〃	10	初級-下

2. 先行研究

本考察では, 下記の3つ区分した. 1つ目は, エスノメソドロジーや会話分析などの社会学である. 言語学も多分に含まれている. 2つ目は, 対話システムを支える自然言語処理, 3つ目は, 深層学習, すなわちディープラーニングである.

2.1 エスノメソドロジー・会話分析

坊農・高梨他(2009)では, 隣接ペアとは, [質問]-[応答]の対をなす発話の連鎖を指すものとして, 対話システムにおける対話モデルに発話連鎖構造の土台としているとある. さらに, 隣接ペアの概念には, [質問]に対し, [応答]がなされなかった場合には, どのような修復連鎖や挿入連鎖構造が生起しながら会話が進行するかを述べている. 魏(2015)は「あの一」や「まー」などをフィラーと定義し, 発話者が何らかの心的操作を行っている最中に発するもので, 場をつなぐ機能を持つ言葉と定義している. 多くは「感動詞」や「間投詞」に区分される. このフィラーを使いこなすのも, あいづちなどと同じく, 会話をつなぐ言葉として, 留意したいと考えている.

2.2 自然言語処理

対話システムに実装される可能性は示している。また、徳永・乾・松本(2005)及び徳永(2014)は、チャット対話の収集からコーパス作成、そしてチャット対話の構造モデルを提案している。このチャット対話の質問や返答などの談話機能を担う構成単位が交換行為である。交換単位は「働きかけ」、「応答」、「補足」の3種類に区分され、さらに2,3の枝葉に分かれている。また、素性に関する考察は有益であり(表2)、本研究ではこれらを精緻化することが具体的な目標でもある。素性の組合せと継続関係の同定や再現率は2人の場合でも3人の場合でも、86%と高く、素性も厳選されている。発言間の結束度は次の式で求めている。 $\langle n(\text{名詞}), \text{rel}(\text{助詞}), v(\text{動詞}) \rangle$ の共起確率 $P(\langle n, \text{rel}, v \rangle)$ を求める。この確率 $P(\langle n, \text{rel}, v \rangle)$ は、Probabilistic Latent Semantic Indexing(PLSI)で推定する。単語の共起を潜在的な意味から同時発生とみなす手法である。PLSIにおける共起確率 $P(\langle n, \text{rel}, v \rangle)$ は次の式で与えられる。

$$P(\langle n, \text{rel}, v \rangle) = \sum P(\langle \text{rel}, v | z \rangle) P(\langle \text{rel}, v | z \rangle) P(n | z) P(z).$$

ここで、 z は共起の潜在的な意味クラス(隠れクラス)を指しており、パラメータの $P(\langle \text{rel}, v | z \rangle)$, $p(n | z)$, $p(z)$ はEMアルゴリズムで推定している。

表2 素性一覧徳永・乾・松本(2005)

素性	素性の説明
発言の末尾の表層表現	各発言の末尾が句点、読点、クエスチョンマークであるか否かの2値
CRRuとPREu間の発言時間の差	CRRuとPREu間の発言時間の差が2分以上であるか否かの2値
発言間の結束度	共起確率に基づくCRRuとPREu間とCRRuとNBNUs間の結束度の強い方を1とする2値
交換行為の対話クラス	対話行為辞典に各交換行為(20種類 ≡ 隣接対)のクラスに分類したもの
交換行為の末尾の表層表現	同一人物の複数読み取れる発言の一番最後の末尾がクエスチョンマークであるか否かの2値
交換行為の発言時間の差	CRRuとPREmの先頭の発言における発言時間の差が5分以上であるか否かの2値

2.3 対話自動生成のディープラーニング

対話応答の自動生成に関しては、ICML Workshop(2015)で Vinyals et al(2015)の Google のチームが NIPS2014 で発表された Sequence to Sequence model を基としている。多層の Long-Short term memory (LSTM)を用いて文章をベクトル化(エンコード)し、別の多層 LSTMを用いてベクトルをデコード(復元)するものである。これは「日本語-英語」間の機械翻訳でよく用いられているアーキテククトであり、従来と比べて、自然な会話を生成するようになった。Ghazvininejad et al(2018)は、上記のモデルを拡張発展させたものである。会話型だけでなく非会話型データも組み合わせることによ

り、Seq2seqにおけるNeural Conversation Modelを発展させたものである。

3. 基礎集計と分析による考察

日本語学習者会話データベース全体的にデータを見渡してみると、全データは390個ある。インフォーマント(日本語学習者)の属性は、20代が圧倒的に多く、女性が男性の2倍近くおり、大半を占めている。日本語学校生や大学・大学院生が半分を占めている(図1)。

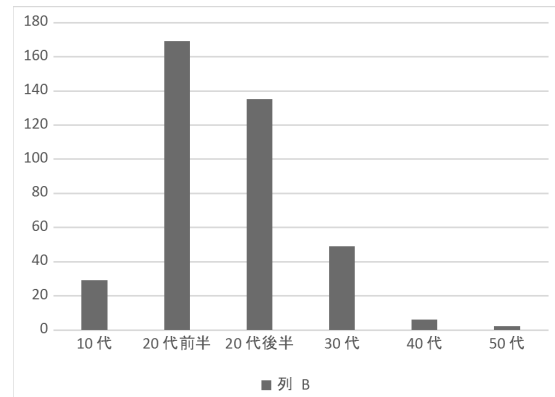


図3.1 インフォーマントの年代別分布

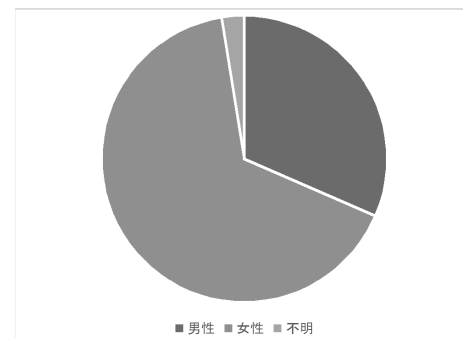


図3.2 インフォーマントの性別

図3.3 インフォーマントの日本滞在時間

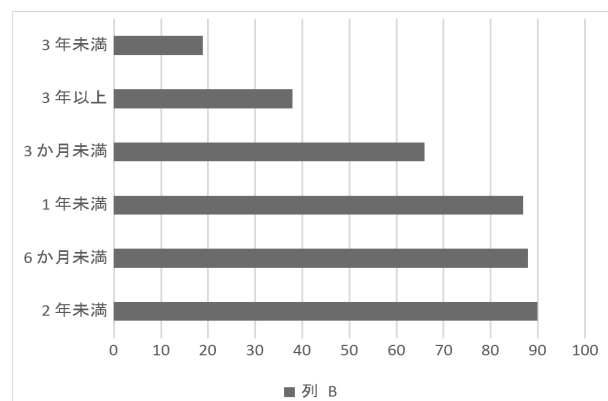


図3.4 インフォーマントの日本語学習期間

かけ) や同意要求 (呼びかけ) とそれに対応して、理解機能 (応答) や同意要求 (応答) が成立しており、高度な対話がなされている。これらを、類型化したものが表 3.5.3 である。いずれも隣接対が多く見受けられ、またフィルターのようなつなぎ言葉も相手への同意を示すなど、発話促進の機能を果たしていると言える。

表 3.5.3 発話連鎖組織の一例

＜連鎖＞	
01A	情報要求= 語り要求
02B	情報提供= 語り(結論・評価)
03A	評価
04B	語り(詳細)
05A	理解・評価

4. データの増加の考察

本研究での新規性の一つとして、国立国語研究所の「日本語学習者会話データベース」(以下、国語研と略す) と同じ設計で取得された KY コーパスを追加することで、まずは、「品詞の分布」や「フィルター」、そして「笑い」の集計をし比較してみた。生産現場やコールセンターなどの商用の現場であると、定量的な根拠が求められる。少しでも、サンプルサイズについて勘どころが得られればと思われる。

4.1 国語研のデータと KY コーパスの比較

表 4.1.1 に、国語研 (339 個) と KY コーパス (44 個) のスクリプト数の一覧を示したものである。KY コーパスを追加すると、全体で 383 個となり、13% の増加となっている。表 4.1.2 から、超級と上級の上のデータの追加が大きい。出身国や年齢等の属性を加味しなければ、データの増加はマイナスにはならないと言える。

表 4.1.1

	国語研	KY-corpus	国語研+KY-corpus	増加率
超級	9	10	19	111%
上級-上	24	18	42	75%
上級-中	34	0	34	0%
上級-下	52	0	52	0%
中級-上	68	4	72	6%
中級-中	84	2	86	2%
中級-下	36	4	40	11%
初級-上	21	2	23	10%
初級-中	10	2	12	20%
初級-下	1	2	3	200%
	339	44	383	13%

図 4.1.1 国語研のスクリプト数の分布

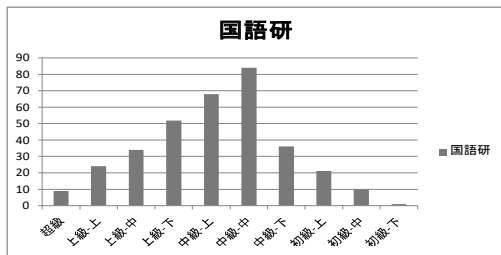


図 4.1.2 KY コーパスのスクリプト数の分布

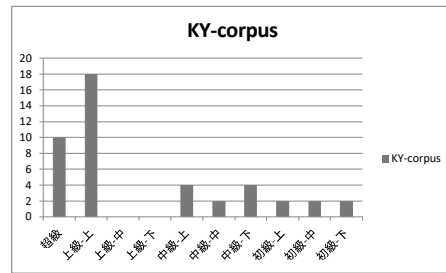


図 4.1.3 KY コーパスの追加差分 (赤棒)

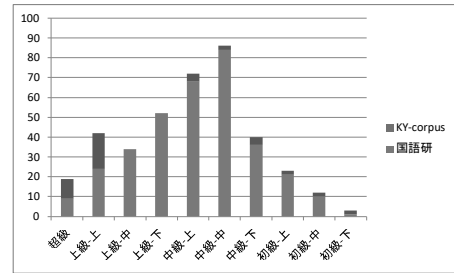
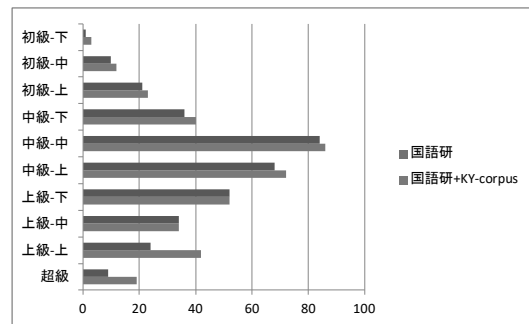


図 4.1.4 KY コーパスの増加分の比較



4.2 コーパスを増やす前と後との品詞の数量での比較考察

まず、表 4.2.1 に、国語研の上級者のスクリプトを形態素解析した後の品詞上位 30 個を示す。

表 4.2.1 国語研超級

1 補助記号-読点	6223
2 助動詞	6061
3 助詞-格助詞	5622
4 補助記号-括弧開	5350
5 補助記号-括弧閉	5348
6 名詞-普通名詞-一般	4822
7 感動詞-一般	4445
8 副詞	4035
9 感動詞-フィルター	3327
10 動詞-一般	3118
11 助詞-終助詞	2855
12 助詞-非自立可能	2618
13 補助記号-一般	2411
14 助詞-接続助詞	2235
15 助詞-係助詞	2109
16 助詞-副助詞	1691
17 記号-文字	1622
18 助詞-準体助詞	1412
19 名詞-普通名詞-サ変可能	1259
20 名詞-普通名詞-副詞可能	1218
21 代名詞	1159
22 連体詞	988
23 名詞-数詞	598
24 接尾辞-名詞的-一般	589
25 名詞-固有名詞-地名-国	553
26 形容詞-非自立可能	536
27 形容詞-一般	503
28 接頭辞	353
29 接続詞	327
30 名詞-普通名詞-助数詞可能	324

表 4.2.2

順位	基本形	個数
1	ん	2150
2	はい	2070
3	です	1820
4	か	1545
5	の	1521
6	と	1479
7	そう	963
8	えー	899
9	も	894
10	んー	741
11	T	703
12	あの	677
13	た	671
14	な	634
15	し	614
16	を	614
17	ます	566
18	[565
19]	565
20	笑	506
21	ない	487
22	その	464
23	あのー	458
24	まあ	451
25	[449
26]	449
27	ー	428
28	って	421
29	*	407
30	あ	384

次に、KY コーパスの上級者のスクリプトを形態素解析した後の品詞上位 30 個を示す。

表 4.2.3 KY コーパス超級 表 4.2.4

順位	品詞	個数	順位	書字形(=表層形)	個数
1	助動詞	6364	1	です	1744
2	助詞-格助詞	5959	2	らん	1307
3	名詞-普通名詞-一般	4525	3	はい	916
4	感動詞-一般	4371	4	いう	914
5	補助記号-括弧閉	3378	5	そう	801
6	補助記号-括弧開	3377	6	あの	750
7	副詞	3221	7	ええ	702
8	動詞-一般	3078	8	あー	562
9	感動詞-フィラー	2707	9	あのー	561
10	動詞-非自立可能	2693	10	ます	530
11	助詞-終助詞	2658	11	ない	505
12	助詞-接続助詞	2407	12	えー	493
13	補助記号-一般	2268	13	から	448
14	助詞-係助詞	2029	14	その	423
15	記号-文字	1572	15	けど	376
16	助詞-副助詞	1571	16	それ	324
17	助詞-準体助詞	1371	17	まあ	314
18	代名詞	1331	18	んー	299
19	名詞-普通名詞-副詞可能	1277	19	こと	294
20	名詞-普通名詞-サ変可能	1221	20	こう	275
21	連体詞	1151	21	なん	258
22	接尾辞-名詞的-一般	647	22	笑い	249
23	名詞-固有名詞-地名-国	583	23	人	209
24	名詞-数詞	519	24	ちよつと	199
25	形容詞-一般	518	25	まし	199
26	形容詞-非自立可能	487	26	さん	198
27	名詞-普通名詞-助数詞可能	444	27	にほん	193
28	接続詞	405	28	ある	192
29	形状詞-一般	346	29	じゃ	192
30	接頭辞	330	30	たら	174

最後に、国語研に KY コーパスを追加した後の上級者のスクリプトを形態素解析した後の品詞上位 30 個を示す。

表 4.2.5 国語研+KY コーパス超級 表 4.2.6

順位	品詞	個数	順位	書字形(=表層形)	個数
1	助動詞	12425	1	です	3564
2	助詞-格助詞	11581	2	はい	2986
3	名詞-普通名詞-一般	9347	3	いう	1902
4	感動詞-一般	8816	4	そう	1764
5	副詞	7256	5	あの	1427
6	動詞-一般	6196	6	えー	1392
7	感動詞-フィラー	6034	7	らん	1327
8	助詞-終助詞	5513	8	ます	1096
9	動詞-非自立可能	5311	9	んー	1040
10	補助記号-一般	4679	10	あのー	1019
11	助詞-接続助詞	4642	11	ない	992
12	助詞-係助詞	4138	12	その	887
13	助詞-副助詞	3262	13	あー	877
14	助詞-準体助詞	2783	14	って	864
15	名詞-普通名詞-副詞可能	2495	15		852
16	代名詞	2490	16		852
17	名詞-普通名詞-サ変可能	2480	17	まあ	765
18	連体詞	2139	18	から	750
19	接尾辞-名詞的-一般	1236	19	ええ	703
20	名詞-固有名詞-地名-国	1136	20	それ	637
21	名詞-数詞	1117	21	けど	611
22	形容詞-非自立可能	1023	22	こう	604
23	形容詞-一般	1021	23	こと	589
24	名詞-普通名詞-助数詞可能	768	24	人	528
25	接続詞	732	25	笑	518
26	接頭辞	683	26	ちよつと	486
27	形状詞-一般	625	27	なん	468
28	名詞-普通名詞-形状詞可能	542	28	日本	455
29	形状詞-助動詞語幹	340	29	じゃ	391
			30	やっぱり	374

5. 突き上げ(Thrust up)と相互行為の考察

職場や医療現場、男女別の会話等の様々な場面における相互行為による研究がなされてきた。ここで、OPI (Oral Proficiency Interview Test) での試験監と外国人受験者での相互行為は、[名前の質問応答]などの定量的なパターンマイニングによって、

以下のような、ある一定の談話構造がパターンが読み取れる。3つ目までは、あらかじめ、事前に対策して準備しておけば答えられるものであるが、この間で試験管は、超級か上級か中級か、または初級かの4つに予測し分類し評価を行う。そして4つ目以降に、1つ以上難易度を高めた質問を行う。これが「突き上げ」である。受験者に対して、1対1の[質問-応答]や勧誘等を行い、即座の記憶ではない、内発的な日本語能力を評価する特殊や相互行為であると考えられる。

また、ベイズ統計学のように、初めての受験者の場合と2回目、または3度目の受験者とは、「突き上げ(Thrust up)」の効果は対数のように減少することが考えられる。いわゆる、条件付き相互行為と呼べるであろう。

- 1つ目: お互いの名前の[質問-応答],
- 2つ目: 出身国の[質問-応答],
- 3つ目: 趣味の[質問-応答]
- 4つ目: 日本についてどう思うか?[突き上げ]

条件付き相互行為の例

例) P(2度目のインタビュー|1度目のインタビュー)

6. 今後の展望

本研究はチャットボットやコールセンターなどでの対話システムを対象としたものである。考察の中で興味深いのは次の2点であった。1つは、データの増加の影響とサンプルサイズについて、2つは、条件付き相互行為である。TOEICやTOEFLなどの同じ形式の対話では、1回目に比べて2回目が、2回目に比べて3回目が試験監の突き上げ(Thrust up)の効果が対数関数のような現象を示すということである。この効用関数はコールセンター等でも同じく確認できるものと考えられる。

参考文献

- [1] 国立国語研究所(2010)「日本語学習者会話データベースの利用手引き」
- [2] 鎌田・山内「タグ付き KY コーパス」<http://jhlee.sakura.ne.jp/kyc/corpus/>
- [3] 徳永・乾・松本(2005)「チャット対話における発言間の継続関係と応答関係の同定」自然言語処理言語処理学会