日本語学習者会話データベースにおける 隣接ペアの談話促進に関する一考察
～対話破綻検出や対話システムへの適用を目指して～

A Study on Promotion of Discourse of Language Action and Adjacent Pair in Japanese Learner's Conversation Database

○ 太田 博三[1]

Hiromitsu Ota 1[1]

[1]放送大学 教養学部
[1]The Open University of Japan

Abstract: In recent years, interactive systems and dialogue generation in natural language processing have attracted attention. Due to the spread of the chat bot to the call center, an 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 provided by the National Institute for Japanese Language, we examine the effect and aim at applying to the tendency of dialogue breakdown and dialogue generation.

1. はじめに

1.1 研究の背景と目的

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中，従来から発展し商用化されているロボットの Pepper や各種チャットボット（Chatbot）は，人間と比べて，小さくない乖離があると指摘され，4 年以上が経過している．チャットボットのコールセンター等への導入も 2 年以上経過している．ここでは，制御文による対話応答が第 1 義的に実装され，第 2 義的にディープラーニング（Deep Learning）による運用が，もくされた．しかし，いずれも不完全燃焼に終始している．どちらか一方，もしくは折衷でも，人間に代替する品質に昇華できていないと考え，調べ始めたのが本考察のきっかけである．

次に，エスノメソドロジーや会話分析の勉強会に参加した際，対話システムのような粗い応答では不十分だという趣旨のご発言と勝手ながら解釈した機会があり，追及してみようと考えた．

そこで，現行の対話システムに定性的な視点で考察し，定量的な分析に持ち込むことで，スケール化させ，実用化に結びつける第一歩にしたいというのが目的である．全体的にディープラーニングに適用したらよいか，部分的かも検討したい．

1.2 研究の新規性

本研究の遠い新規性となるが，もっぱら，数量データによるディープラーニングに定性的な要素を取り入れたいという点であるが，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，計量化してみた．ここで，実証的な知見が得られれば，話者交代の予測や対話の破綻をしても修復する発話を学習させるなど，次につなげられる．具体的には，隣接ペアの次には，主に以下の 5 つが考えられる．

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 ：あのー，そのー，えーと，えっと

今後，これらを分析し追加することで，更なる対話システムの質的向上につながる可能性があると考えられる．

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．

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．

図1．属性別の日本語教育会話データベースの検索画面

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)．対話のスクリプトは，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）されたスクリプトが入手でき，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，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「あのー」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．く n(名詞)，rel(助詞)，v(動詞)〉の共起確率 $P(<n, rel, v>)$を求める．この確率 $P(<n, rel, v>)$ 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 $P(<n, rel, v>)$は次の式で与えられる．

$$P(<n, rel, v>) = \Sigma P(<rel, v|z)\ P<rel, v|z)P(n|z)P(z).$$

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 $P(<rel, v|z>)$，$p(n|z)$，$p(z)$)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	各発話の末尾が句点，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NBNu_s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話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の先頭の発話における発話時間の差が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「日本語ー英語」間の機械翻訳でよく用いられているアーキテクトであり，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．

2.4 対話システムと会話ユーザーインターフェース

狩野(2017)では，現在に至るまでの対話システムと将来の展望を簡潔にまとめられている．1960年代に開発されたELIZAや人工無能から，「○」「△」「×」などのアノテーションによるチューリングテストを経て，現代の雑談対話システムの1つである2016年に発表された論文に基づくMicrosoft社の「りんな」まで網羅している．「りんな」では，発話ペアデータと教師付き機械学習は統計的な対話システムの多くに共通していることが少なくない．また，対話データを正解データとしてくんれんする強化学習では，状態遷移の訓練になるため，会話の流れを学習することになり，未知の対話に対応することが期待されている．

2.4 国内外での取り組み

対話破綻検出チャレンジ(2015-2017)やDTFC7, NTCIR-14など，年次でハッカソンのような国際的な大会として開催され，集合知となっている．

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）．

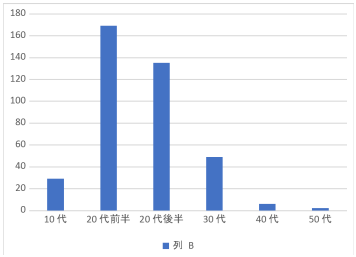

図3.1 インフォーマントの年代別分布

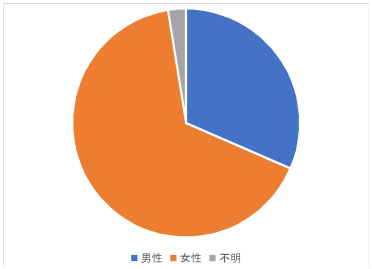

図3.2 インフォーマントの性別

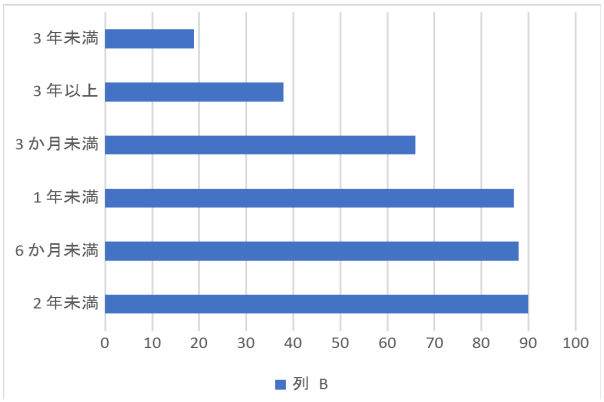

図 3.3 インフォーマントの日本滞在時間

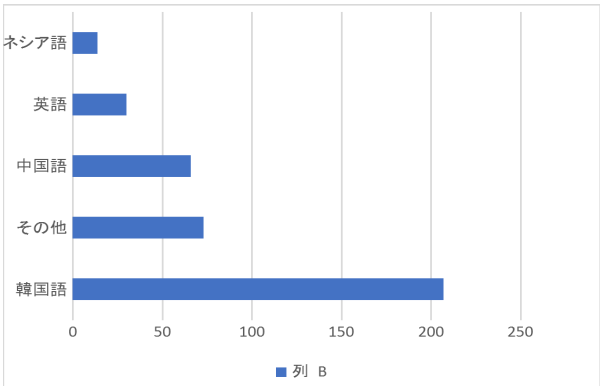

図 3.4 インフォーマントの日本語学習期間

3.1 インフォーマント属性間の比較について

本考察では，自然な対話文の自動生成を目指していることから，次の 2 つを比較考察する．OPI の判断尺度は，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」に準じ，超級と中級とで比較考察した．超級は人間と仮定し，中級はチャットボットなどの機械としてみた．主な選択した要因は次の 2 つである．

1) 流暢さ
2) 語用論的能力

超級では，「流暢さ」とは，会話全体がなめらかであること，これに対して，中級では，「流暢さ」とは，つかえることが多く一人で話つづけるのは難しいと定義されている．一方，超級の「語用論的能力」とは，ターンテイキングや間の取り方，相づちなどが巧みにできると定義されているのに対して，中級は，相づちや言い換えなどに成功するのはまれとされている．

3.2 超級と中級のデータについて

中級は年齢が 20 代半ばとまとまって分布しているのに対し，超級は 23 歳から 49 歳とばらつきが大きい．これはデータ数が 9 つと少ないためであるが，出身国と母国語は超級のハンガリーのインフォーマントを除いて，韓国である．滞在期間が大きく異なり，
超級は 5-10 年が大半であるのに対し，中級は 3 か月から 6 か月の間に分布している．

表 3.1 超級のインフォーマント属性

	データ番号	OPIレベル	年齢	性別	出身国	母語	職業等	日本滞在期間	日本語学習期間(参考)	日本語能力試験(参考)
1	10	超級	28	女	韓国	韓国語	会社員	5年	7年	－
2	15	超級	26	男	韓国	韓国語	専門学校生	3年	18年	－
3	76	超級	34	女	韓国	韓国語	主婦	5年	4年	－
4	202	超級	35	女	韓国	韓国語	講師	10年	6年	1級
5	230	超級	26	男	中国	中国語	大学院生	5年1ヶ月	7年	－
6	258	超級	43	女	ブルガリア	ブルガリア語	教師	18年	22年？	－
7	323	超級	23	男	韓国	韓国語	大学生	5年	8年	－
8	338	超級	49	男	韓国	韓国語	会社員	22年	2年	1級
9	349	超級	40	女	韓国	韓国語	大学教員	15年	2年～3年	－

表 3.2 中級のインフォーマント属性

	データ番号	OPIレベル	年齢	性別	出身国	母語	職業等	日本滞在期間	日本語学習期間(参考)	日本語能力試験(参考)
1	2	中級-下	22	女	韓国	韓国語	大学生	3ヶ月	11ヶ月	－
2	12	中級-下	25	男	韓国	韓国語	日本語学校生	5ヶ月	5ヶ月	－
3	26	中級-下	27	女	韓国	韓国語	日本語学校生	5ヶ月	18ヶ月	－
4	8	中級-中	24	女	韓国	韓国語	日本語学校生	3ヶ月	7ヶ月	－
5	9	中級-中	25	女	韓国	韓国語	日本語学校生	2ヶ月	8ヶ月	－
6	22	中級-中	28	女	韓国	韓国語	日本語学校生	5ヶ月	1年	－
7	6	中級-上	24	女	韓国	韓国語	日本語学校生	6ヶ月	1年6ヶ月	－
8	7	中級-上	28	女	韓国	韓国語	専門学校生	2年	23ヶ月	－
9	11	中級-上	26	女	韓国	韓国語	日本語学校生	6ヶ月	9ヶ月	－

表 3.3 比較に用いたデータセット数

OPIレベル	母数	使用したデータ数	合計
超級	9	9	9
中級-下	36	3	
中級-中	84	3	9
中級-上	68	3	

3.3 隣接ペアとその計量化の検討

隣接ペアの重要な特性に，第 1 部分(First-Pair-Part: FPP)が産出されると，それに対応する特定の型の第 2 部分(Second-Part-Pair: SPP)の産出が条件的に適切になると前川・小磯他(2015)は言及している．本節では，試みの一環として，形態素解析した後に，同じ語句がでてきたら，その合計の半分として数量化した後に，目視で確認をすることにした(表 3.3.1，表 3.3.2，表 3.3.3)．隣接ペアが見出しやすい次の 4 つの品詞に焦点を当てて考察することにした．

1)名詞
2)感動詞
3)間投詞
4)応答詞

表 3. 3. 1

隣接ペアである例		
‘chiba-1232:514.5590-516.4996		
A1:	選べんだ	←第1部分(FPP)
B2:	選べる	←第2部分(SPP)
A3:	へえ	

表 3.3.2

隣接ペアでない例		
‘chiba-0332:437.2296-441.4541		
C1:	私も動物飼いたいな:	←働きかけ(I)
C2:	植物でもいいや	←働きかけ(I)
A3:	うん	←応答(R)

中級データセット
1) 頻出名詞(上位 10 件)
ん(2254) ー(1717) 笑(649) 私(294) こと(268) 音(249) 人(247) 息(246) 日本(183) お(169) 今(168) , (168) 韓国(167)

1) 頻出詞(上位 10 件)
2)感動詞
はい(2594) あー(944) あ(530) えー(309) そうですね(61) ありがとう(55) はーい(52) え(48) ま(41) ふーん(33) うん(30)

超級データセット
1)名詞
ん(2869) ー(1778) 笑(546) の(376) こと(342) 人(310) それ(292) 日本(261) 今(199) 韓国(184)

2)感動詞
はい(2114) えー(1038) あー(389) あ(273) そうですね(183) ま(159) なるほど(97) え(89) ふん(65) ふーん(54)

3.4 結果の考察
中級の名詞では，ん(2254) ー(1717) 笑(649) 私(294)などが多く，主観性が見受けられた．その一方で，超級の名詞では，人(310) それ(292) 日本(261) 今(199) 韓国(184)などのようぬ，代名詞や国柄を表す語句が見受けられ，客観性が見受けられた．
また，中級の感動詞では，はい(2594) あー(944) あ(530) えー(309) はーい(52)など，あいまいさが見受けられた．一方で，超級の感動詞では，はい(2114) えー(1038) そうですね(183) ま(159) なるほど(97) など確かな返答が見受けられた．

4. 今後の展望

本研究はチャットボットなどの対話システムを対象としたものであるが，今後は次のような視点で，ロボテックスを対象とした研究につなげたい．秋谷・丹波・久野・山崎他(2007)では，介護ロボットの実現に向けて，介護者と高齢者との相互行為を深く分析したものである．今後はロボットに搭載されることが予見される．相互行為の視点が，より人間的になると考えられ，期待されている．

謝辞

本研究の一部は，学部時代にマレーシア語や生成変形文法を教えて頂いた正保勇先生(東京外国語大学名誉教授)の雑談の中での教えが大きく影響している。また計量社会科学に関しては，博士課程時代にご指導頂いた聖学院大学大学院の松原望先生(東京大学名誉教授)を想起しながら試行錯誤できた．ここに謝意を表したい．

参考文献

[1] 国立国語研究所(2009) 「日本語教育ネットワーク」https://nknet.ninjal.ac.jp/nknet/ndata/opi/
[2] 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」
[3] 鎌田・山内「タグ付き KY コーパス」 http://jhlee.sakura.ne.jp/kyc/corpus/
[4] 坊農・高梨他(2009)「知の科学 多人数インタラクションの分析手法」3 章, 人工知能学会編集, オーム社
[5] 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
[6] 徳永(2004)「チャット対話における発言の継続関係と応答関係の同定」修士論文 奈良先端科学技術大学院大学 情報科学研究科 情報処理学専攻
[7] 牧野成一他(2001)「ACTFL-OPI 入門」アルク
[8] Vinyals,et al(2015) Quoc Le. A Neural Conversational Model, arXiv
[9] Ghazvininejad(2018) A Knowledge-Grounded Neural Conversation Model. Microsoft
[10] 喜連川他(2017)「暗黙の発話状況を考慮したニューラル対話モデル」. 言語処理学会 第 23 回年次大会 発表論文集
[11] 串田, 平本, 林(2017)「会話分析入門」 勁草書房
[12] 対話破綻検出チャレンジ 2015 https://sites.google.com/site/dialoguebreakdowndetection/
[13] 船越・東中他(2016)「対話破綻検出チャレンジにおける対話破綻データと破綻検出結果の分析」 言語処理学会 第 22 回年次大会
[14] 東中・船越(2016) Project Next NLP 対話タスクにおける雑談対話データの収集と対話破綻アノテーション 人工知能学会 SLUD
[15] DSTC7(2017) Dialog System Technology Challenges http://workshop.colips.org/dstc7/index.html
[16] NTCIR-14(2018-19) - 国立情報学研究所

http://research.nii.ac.jp/ntcir/ntcir-14/index-ja.html
[１７] EMCA 研究会 エスノメソドロジー・会話分析とはなにか - http://emca.jp/learn
[１８] 狩野(2017) 「コンピューターに話が通じるか対話システムの現在」 情報管理 Vol.59 no.10
[１９] 石崎・伝(2001)「談話と対話」東京大学出版会
[２０] 藤田 自然言語表現の言い換え http://paraphrasing.org/~fujita/paraphrasing-ja.html
[２１] 魏(2015)談話におけるフィラー「ま（一）」の待遇差に関する予備的考察, 山口大学東アジア研究学術雑誌論文
[２２] Smith et.al(2000) Conversation Trees and Threaded Chats, CSCW
[２３] 秋谷・丹波・久野・山崎他(2007)「介護ロボット開発に向けた高齢者養護施設における相互行為の社会学的分析」電子情報通信学会論文誌 D Vol.J90-D No.3 pp. 798-807