

# 「対話の一体感」をもたらす音声インタラクションの時間的特徴

## —ロボットと人の漫才対話データの分析から—

本井佑衣 (立命館大学学部生) 岡本雅史 (立命館大学)

### 1. はじめに

これまで、話者交替がいつ行われるかや相槌を打つタイミングについての対話研究は数多く行われてきたが、話者交替や相槌のタイミングが対話全体の中で経時的にどのように変化しているのかについてはあまり考察されていない。こうした対話インタラクションの時間的特徴を捉える上で、二者間で発話ターンを交替させる際に各々の話者がもつ固有のリズムの相互調整に着目する必要がある。これまでの研究では、Szatrowski (1987) で発話に音楽的なリズムが存在することや、小磯・伝 (2000) らの研究で話者交替に規則的な法則があることが示されている。そこで本研究では、これらの研究を踏まえて、漫才対話における音声インタラクションが紡ぐ「対話の一体感」がどのような要因から生じるのかを明らかにする。

漫才対話を分析対象にする理由は2つある。一つは漫才が基本的に二者間での話者交替によって構成される音声インタラクションであるため、日常会話よりも環境要因に左右されずに分析を行うことができるからである。一方、テンポがいい対話としての規範的モデルとして日常対話との比較を行うことも可能である。漫才の演者はそれぞれ発話ターンや間を随時相手に合わせて変化させながら「息の合った」インタラクションを作り上げていると考えられる。お互いに「息の合った」対話、つまり「対話の一体感」を感じさせる要因は一体何であるのかを解明することで、最終的には日常場面におけるコミュニケーションでリズムやテンポの良い対話を行う手がかりとなることが期待される。

### 2. 関連研究

従来の研究では、我々が日常のコミュニケーションにおける発話の中にも音楽的なリズムがあることが示されているが、一体何がそうしたリズムを作っているのかは明確ではない。本井・岡本 (2019) では、漫才対話データの分析からリズム主導者が作るリズムをもう一方の対話者が合わせる同調の動きがあることが示された。ここでは、モーラや高低アクセントから生まれるマイクロリズムが分析されていたが、対話の重層的なインタラクションリズムを捉えるためには、発話長や交替潜時から作られる、よりマクロなリズムにも着目する必要がある。

小磯・伝 (2000) では、話者交替に規則的な法則があることが示されており、対話の自律モデルとして提案されていた。これは、話し手が発話をやめるときの行動傾向を文末度として数値に表したものである。文末度が高いほど話し手は発話を停止し聞き手は発話を開始する傾向にあり、非円滑な移行においても同じメカニズムで生じることも示唆されている。自律モデルによると発話者それぞれが異なる発話関数をもっていると考えられており、この関数が認知環境や対話相手などによって行動特性として現れる。

石本・榎本 (2017) は、話者移行適格場の到来を予測させる発話中の韻律変化の解明を行った。次話者となる聞き手は動詞部までの情報から発話末要素が後続するかどうかを音響的な要素から判断していることが判明した。また発話の最終アクセント句付近で話者の韻律が大きな変化を見せることも示している。これらは話者の移行がどこで起こるのかを予測するための研究であるが、本研究では、対話における円滑なタイミングを超えた、コミュニケーション全体を支配するインタラクション・リズムの存在を仮定している。

Erickson & Shultz (1982) は、会話の言語的・非言語的な行動は、音楽と同じような小節に分けることができ、メトロノームで計ることのできる正確なリズムを持って進展していくと述べており、発話を小節に分けることができるということを示した。また、ザトラウスキー (1989) は、この研究を踏まえ、日本語の対話を拍子付けする方法について示した。スロー再生した音声に適切な拍子を決め、さらにそれらを音符とリズム譜で表現し、原則として一語につき一音符を対応させている。これらの分析の中で、相槌を打つ適切なタイミングは前の発話が終わったすぐ後の拍子であり、相槌以外の応答タイミングは一小節分あけることが適切であると述べられている。

### 3. 研究目的と研究手法

本研究の目的は、対話のインタラクションリズムが二人の演者間でどのように生成・修復されているのかを分析すること

で「対話の一体感」を認知させる要因を解明することにある。漫才対話において二者はどのようにテンポの良さを共同的に作り上げているのかを分析することで、漫才に留まらず、日常場面のコミュニケーションにおいて息の合うやりとりの本質を探るヒントとなることが予想される。

テンポが良い対話とは両者が交替潜時を作らずに次々に発話インタラクションを行うことではなく、相手の固有のリズムに合わせていくものであると考えられる。マクロな視点から、対話を構成する発話ターンと交替潜時をそれぞれ一つの分析単位とみなし、それらの長さや組み合わせによってインタラクションリズムが生み出されていると仮定し分析を行った。

具体的には、(1) 発話長と交替潜時によって作られるインタラクションリズムが経時的に変化すること、(2) インタラクションリズムと相槌との対応関係、の両者によって「対話の一体感」が生まれることを示す。漫才対話を分析するにあたって、プロの漫才師「かまいたち」とヒューマノイドロボットPepperによる3本の映像音声データを利用し、音声分析ツールPraat<sup>2</sup>によって発話の書き起こしと音声データの分析を行った。データは2018年11月に立命館大学で収録したものである。Pepperの操作は対話データを収録する際の手法として知られるWizard of Oz (WOZ) 方式によって行った。WOZ方式では、ロボットになりすました人 (Wizard) が人間の発話に対して応答するため、二名の操作者が舞台裏から発話タイミングを操作した。長岡ら (2002) で、対話においては音声的情報が二者間の交替潜時を類似させる要因とであると示されていることから、今回は音声情報に絞って分析を行った。そして、それぞれボケ役がロボットである場合とツッコミ役がロボットである場合と人間同士での漫才対話の3つのパターンでの比較を行った。各パターンのデータ情報を表1に示す。

分析手順として、まずボケ役の発話、ツッコミ役の発話、交替潜時がどのような頻度で生起し、時間的に占有しているかを求めた。交替潜時は次話者が調整可能であるため、直後の次発話に依存するものと規定する。発話重複によって交替潜時がマイナスとなる場合は、次発話が遮ったところからを発話交替とみなし、その分だけ当該ターンの時間を削る。また、交替潜時は50ms以上の休止から生じるとし、それ以下の休止時間は存在しないものとみなした。以下、ボケ役の発話ターンをB、ツッコミ役の発話ターンをT、交替潜時をPとする。これらの配列で対話全体の時間的特徴を捉えることができる。対象とする配列は全部でTPB, TBP, BPT, BTP, PTB, PBT, TBT, BTB, PBP, PTPであり、配列ごとの合計時間の占有率と配列の生起頻度を計測した。また、B, P, Tそれぞれの要素からどの要素に遷移するのかも観察した。

## 4. 結果と考察

### 4.1 漫才対話要素の時間的占有率

分析の結果、人間 (Human) 同士の対話データ (以下, TH-BH) とツッコミ役が人間でボケ役がロボット (Robot) の対話データ (以下, TH-BR), ツッコミ役がロボットでボケ役が人間の対話データ (以下, TR-BH) の3つのデータでそれぞれ異なる特徴が見られた。表2より、TH-BH で時間的占有率が最も大きかった配列はTBTで次にBTB, BPTと続く。TH-BRで最も多かったのはBTBの配列で次にTBT, BPTと続く。この2つのデータでは全体を多く占有する配列が類似しており、全体的に交替潜時を作らずに連続した発話が行われる傾向がある。TH-BHとTR-BHではPからTへの遷移がTH-BRと比較して多い (表3)。つまり通常はツッコミ役の方が交替潜時を調整する傾向が強いが、ボケ役がロボットの場合はロボット自身がタイミングよく発話交替できないことを示している。このことは、TR-BHで最も多かったのがBPTの配列で次にPTB, TBPと続くことから、前の2つのデータと比較するとツッコミ役のロボットの発話ターンがほとんど交替潜時をおいてからの開始となっていることから示唆される。一方、表3を見るとTR-BHのP→Tの遷移頻度はTH-BHとほぼ同じように見えるが、Pを挟まずにB→Tとなる流れがかなり少ない。このことから、ツッコミ役の発話が全体を通して交替潜時をとってから行われていることが読み取れる。

表2: 配列の時間的占有率と生起頻度

		TPB	TBP	BPT	BTP	PTB	PBT	TBT	BTB	PBP	PTP
TH-BH	全体比 生起頻度(回)	6% (9)	10% (13)	13% (20)	3% (5)	10% (15)	3% (3)	<b>22%</b> (22)	<b>27%</b> (24)	4% (8)	3% (7)
TH-BR	全体比 生起頻度(回)	9% (10)	9% (9)	14% (14)	3% (5)	6% (7)	7% (7)	<b>23%</b> (15)	<b>21%</b> (15)	2% (4)	5% (6)
TR-BH	全体比 生起頻度(回)	12% (14)	16% (14)	<b>28%</b> (27)	2% (3)	<b>19%</b> (16)	1% (2)	3% (4)	5% (5)	6% (12)	7% (10)

表3: 遷移頻度

		B→T	B→P	T→B	T→P	P→T	P→B
TH-BH		27	21	38	11	22	9
TH-BR		22	14	24	12	14	12
TR-BH		7	27	20	14	27	14

<sup>1</sup> よしもとクリエイティブ・エージェンシー所属。

<sup>2</sup> Praat [ver.6.0.43] (<http://www.fon.hum.uva.nl/praat/>) 最終アクセス日: 2019/12/20

## 4.2 発話長と交替潜時によって作られるインタラクションリズム

次に発話長と交替潜時それぞれを一つのリズム構成要素とみなして分析を行った。前話者の発話長に対して次話者は交替潜時と発話長をどのようにとるかでグローバルなリズムの変化が見られると考える。そこでこのリズムの変動度を、現話者の発話長( $L_n$ )、次話者に依存する交替潜時( $P_n$ )、次話者の発話長( $L_{n+1}$ )の3つの要素を元に次式で定義する。

$$\text{リズム変動率 } x = \frac{P_n \text{ (次話者に依存する交替潜時)} + L_{n+1} \text{ (次話者の発話長)}}{L_n \text{ (現話者の発話長)}}$$

この式で求まる  $x$  値が大きくなるほど  $L_n$  に対しての次話者ターンが長く、 $x$  の値が 1 から 0 に近づくにつれて  $L_n$  に対しての次話者ターンが短くなる。また  $x$  が 1 に近づくほど  $L_n$  と次話者のターン長が等しくなっていることがわかる。さらに時系列でどのように変化しているかを見ることでリズム変動率がどちらに寄っているのかも観察することができる。 $P_n$ がないときは  $L_n$ 、 $L_{n+1}$  の比率となる。

この値を元に時系列でどのように変動率が推移しているかをグラフに表した。変動率  $x=1$  を基準としてどれだけの差異があるのかを見るために対数目盛りで表示する。変動率の値による横幅と縦幅が等しく続くと同じようなリズム構成で対話が行われていることを示す。TH-BH のデータ数が他の2つと比べておよそ 1.3 倍あったのでグラフの横幅もそれに対応するようにした。

TH-BH (図1) では変動率が前半 1/4 は大きい途中からプラスの変動率が続いており、 $L_n$  に対して次話者のターンが長いことを示している。ネタの内容と照らし合わせると、1個目のネタの開始と一致していた。前半のフリの部分で TH に対して BH の発話ターンが短かったのに対してここから BH の発話ターンが徐々に増えていく。およそ半分のところでマイナスの変動率が大きくなっており、2個目のネタが始まる箇所と一致している。後半は連続して変動しており、発話内容は TH のたたみかけである。TH が話す量が増えるため、BH は自分の発話ターンを短くすることで調整していると考えられる。後半もグラフの縦幅と横幅がほぼ同じようになっていることから同じリズムが数ターン繰り返されていることがわかる。変動率が高くなる部分の  $L_n$  は TH がマイナス変動率 15 とプラス変動率 5、BH はマイナス変動率 20 とプラス変動率 5 となっており TH がプラスのリズム変動をつくる確率のほうが高かった。TH-BR のデータ (図2) と TR-BH (図3) を見ると全体を通して変動の幅がシラバラであり、プラスの変動を作る傾向にある。TH-BH ほどの緩急は見られないが1個目のネタ導入前と2個目のネタ導入前の変動率の動きがほぼ一致していた。変動の緩急がないのはロボットが人間の作るターン長に合わせて交替潜時や自分のターン長を合わせたり崩したりする調整が行われていないからではないかと考えられる。また表1からもわかるように TR-BH では T と B の全体を通しての発話長が他の2つと比べて短く、分析対象時間が同じことから交替潜時がとても長いことがわかる。TR が発話内の間を含めた発話ターンを長めにとっているために BH は変動率が大きくなりすぎないように交替潜時をとったり、TR の発話に対して相槌を打つことで調整していると考えられる。

## 4.3 相槌と発話内休止によるセグメンテーション

インタラクションリズムの変動率の分析を元に、発話長は現話者ではない方の相槌によって分割されるのではないかと考え、相槌によってセグメントした場合のリズム変動率も求めた。また、対ロボットの対話になったときに人間が発話内休止

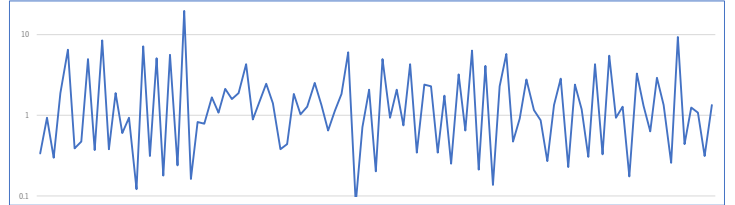


図 3 : TH-BH のリズム変動の推移

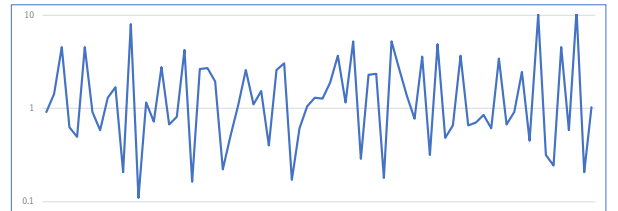


図 2 : TH-BR のリズム変動の推移

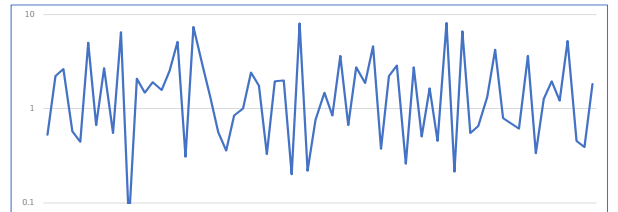


図 1 : TR-BH のリズム変動の推移

をとることが特徴的であったため、発話ターン内で50ms以上の休止が取られた場合も発話内休止としてセグメントを行った。3つとも変動率の縦幅は小さくなり、さらに細かい変動になった。これはセグメントによってx値の調整が行われたからであると考えられる。特にTH-BH(図4)では前半と中盤のリズム変動が主軸を基準に上下対称の動きになったことがわかる。図5,6を見ると前節のグラフよりも緩急がついたことによってロボット対人間の対話データが人間同士の対話データのリズム変動の形に近づいていることがわかる。これにより、人間がロボットの発話に対して相槌を入れたりロボットの発話に合わせた発話内休止を入れることでリズムを保とうとする動きがあるのではないかと考えられる。相槌や発話内休止が入ることによって前節の分析ではわからなかった変動率の差を埋めようとする動きがあることがわかった。

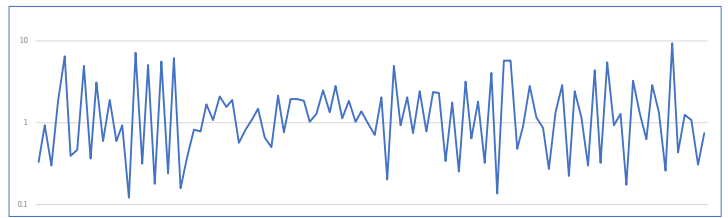


図4: TH-BH リズム変動の推移 (セグメンテーション後)

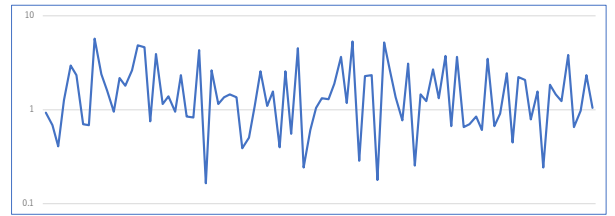


図5: TH-BR リズム変動の推移 (セグメンテーション後)

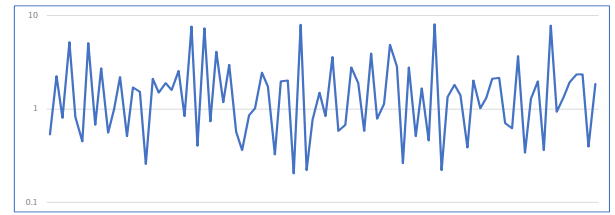


図6: TR-BH リズム変動の推移 (セグメンテーション後)

## 5. おわりに

対話におけるリズム変動率をグラフ化することで、対話の中で変調が生じる箇所や同じリズムが繰り返される部分が変動率の動きとして可視化が実現した。人間同士の対話データでは前話者の発話ターンの長さに対して交替潜時を含めた発話ターンの長さを一定に保とうとする収束の動きが対ロボットの対話と比べて特に強く存在することが判明した。

今後の課題として、マクロな視点からのリズムの分析をふまえたマイクロな視点で高低アクセントを含めた分析や、ジェスチャーなどの非言語情報も含めることで、コミュニケーション全体のリズムについての解明をすすめることができると考える。

## 謝辞

本研究はJSPS 科研費 17KT0143 の助成を受けたものである。またデータ収録において、吉本興業ホールディングス株、神奈川工科大学の大塚真吾研究室、ならびによしもと工学プロジェクトの協力を得たことを特記し、ここに感謝する。

## 参考文献

- Erickson, F. & Shultz, J. J. (1982). *The counselor as gatekeeper: social interaction in interviews*. NY: Academic Press Inc.
- 平出昌嗣 (2017). 発音とリズム—学校文法を超えて(2) 千葉大学教育学部研究紀要, **66**(1), 345-350.
- 石本祐一・榎本美香 (2017). 話者移行適格場の到来を予測させる発話中の韻律変化の解明 日本認知科学会, **2**(2), 31-37.
- 鹿島央 (2010). 日本語リズムの派生について—初級教材の分析と音声教育への応用をめざして— 名古屋大学留学生センター紀要, **8**, 5-14.
- 加藤雅代・古村光夫・橋本新一郎 (1994). 母音部エネルギー重心点に着目したリズム規則 日本音響学会誌, **50**(11), 888-896.
- 小磯花絵・伝康晴 (2000). 円滑な話者交替はいかにして成立するか—会話コーパスの分析にもとづく考察— 認知科学, **7**(1), 93-106.
- 本井佑衣・岡本雅史 (2019). 漫才対話の「テンポの良さ」を支える発話リズムの同期・変調パターン 社会言語科学会第43回大会発表論文集, 170-173.
- 長岡千賀・小森政嗣・中村敏枝 (2002). 対話における交替潜時の2者間相互影響 日本人間工学会誌, **38**(6), 316-323.
- 中村敏枝 (2009). コミュニケーションにおける「間」の感性情報心理学 音声研究, **13**(1), 40-52.
- 岡本雅史 (2008). 会話構造理解のための分析単位—実践: 漫才対話のマルチモーダル分析— 人工知能学会誌, **23**(4), 552-558.
- 羅希・定延利之 (2018). ボケ役、ツッコミ役の発話タイミングから見る面白さに対する認識—日本語母語話者、日本語学習者、中国語母語話者の比較から— 日本認知科学会第35回大会論文集, 943-951.
- 宿利由希子・ヴォーゲ ヨーラン・林良子・定延利之 (2018). ユーモアを生み出すための『間』—ボケとツッコミのタイミングに関する考察 日本認知科学会第35回大会論文集, 912-918.
- Szatrowski, P. (1987). A Discourse Analysis of Japanese Invitations, *Berkeley Linguistics Society*, **13**, 270-284.
- ザトラウスキー, ポリー (1989). 相づちとそのリズム 月刊日本語, **3**, 32-35, アルク出版