

## 三者会話データの収録方法及び分析枠組の概要

高梨 克也

井佐原 均

通信総合研究所

{takanasi,isahara}@crl.go.jp

## 1. はじめに

対話進行のメカニズムの解明にとって、対話コーパスからの帰納的・一般化は不可欠な手法の一つである。しかし、従来の対話コーパスは二人間の課題指向的対話を収録したものが大半であり、対話の基本構造を定式化する上で、二人対話に基づく分析では不十分な点が多いと考えられる。本発表は三人での自然な会話のデータについて、その収録方法及び課題デザインの概要を紹介するとともに、分析のために必要な基本的観点について整理する。

## 2. 三者会話コーパスの必要性

従来の多くの対話システムは人間とコンピュータの間の二人間会話であると見なすことができる。これに対して、例えばヘルパーエージェント[中西00]を含む環境は人間二人とエージェントからなる三者会話状況となる。しかし、対話システムにおいても、ヘルパーエージェントにおいても、システムの会話能力はごく限定的であるというのが現状であり、まずは目標とする課題を人間同士の間で共同で解決するプロセスを模擬対話として収集し、そのデータからの帰納的・一般化を通じてシステムに必要な要件を解明していくという手法が不可欠となる。しかし、二人間の課題指向的対話については、すでにさまざまな対話コーパスが構築・公開されているが[堀内99][CBRDD00]、三者での会話を体系的に収録・分析する試みは[石崎99][松坂01]など極めて少ない。

また、より理論的にも、二人間対話コーパスに基づいて、会話における話者交代規則などを構築しようとする試みも存在するが[小磯00]、二者会話の場合と三者以上の会話の場合とでは、次話者の決定や聞き手に向けた発話のデザインなど、会話進行にとって基本的な特徴に関しても相違があることが予測される。

このように、三者以上の参加者を含むコミュニケーションツールの開発や対話規則についての理論的研究のための基礎的データとして、三者会話コーパスの構築は重要なものである。

## 3. 収録デザイン

## 3.1 収録環境

収録は2001年10/1(月)~5(金)の5日間、京都のデジタルシティ研究センター二条ラポで行われた。被験者数は15組45人で、各組とも二課題を行い、計30対話を収録した。1対話あたりの平均所要時間は約20分である。

収録環境の機材面に関して[図1]、収録媒体は音声はDAT、映像はミニDVを使用した。音声収録にはDATデッキ3台を用い、デッキ1は被験者A,B、デッキ2は被験者B,C、デッキ3は被験者C,Aの音声を、それぞれL,Rチャンネルに収録した。映像は三人分の映像を4分割器で合成し、1台のDVデッキで収録した。

DVの音声チャンネルにはLに被験者A,C、Rに被験者BとオペレータDの音声を記録した。DVでの記録は精確な分析用というよりも記録用であり、1本だけでなるべく会話の全体像が把握できるものであることを目指した。

各被験者の音声を独立のチャンネルに収録するため、各被験者はそれぞれの別の防音ボックスに入った。そのため、対面状況ではないので<sup>1</sup>、各被験者へのフィードバックとして、例えば被験者AはヘッドフォンのL,RからそれぞれB,Cの音声を聞くとともに、音声の方向と一致するように、2台の小型モニターを設置し、左のモニターにはB、右にはCの映像が映るようにした[図2]。

## 3.2 課題・被験者デザイン

被験者は全員大学生である。同性の3人を1組とし、組み合わせ方は、(a)ABC友人、(b)AB,AC友人、(c)AB友人、(d)全員初対面、の4通りである。各組とも以下の2種類の課題を連続して行い、その際の会話音声や身振り・表情を上述の方法で収録した。課題選択の基準は、初対面の組み合わせの場合にも、自然な活発な会話が行われることを促進するものであることである。また、事前に研究目的が会話自体の収録であることを告げると会話がちこちこちなものになることが予想されたため、実験の目的は社会的IQに関する簡単なテストであると説明し、収録終了後にデブリーフを行い、本実験の結果に基づき被験者の社会的IQについて分析することはないことを告げた。各被験者にはデータ公開に関する同意書に署名していただいた。

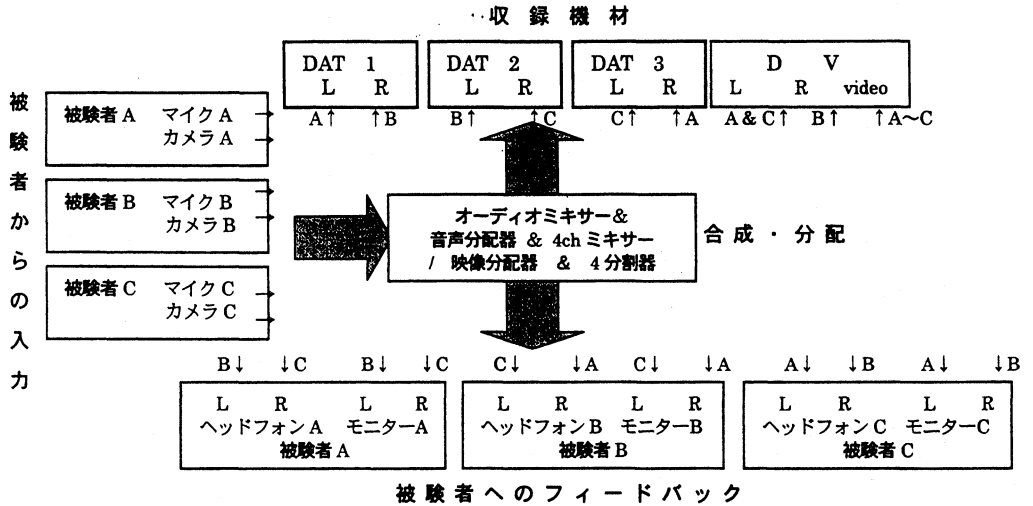
写真課題<sup>2</sup>

**教示(要約):** 社会的IQに関する簡単なテスト。写真についての問題[図3]に対して、自分以外の二人の被験者と話し合いながら、共通の解答を選択する。問題は全部で3問。順序や話し合いの方法は被験者の自由。

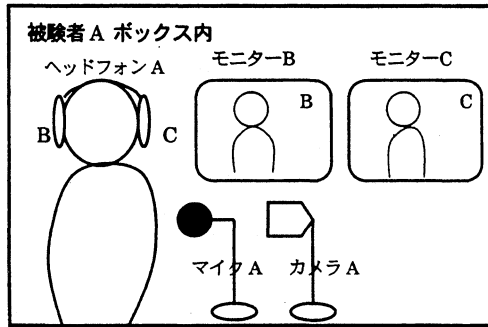
被験者のうちの一人ないし全員に対して各問題についてのヒントを提供した。ヒントを設定した理由は、1. 特に初対面のグループでは話し合いが活発にならない、全員の最初の答えが一致した場合には会話がほとんど行われなまま課題が終了してしまう、という難点が予想されること、及び、2. 会話の主導権との相関を分析するためである。ヒントは、一人が持ち、他の二人が持たない

<sup>1</sup> [松坂01]では視線の動きなどの分析に焦点があるため、対面状況の自然さが重視されている。これに対して、本収録では三者の音声を分離して収録することが重視されているため、対面状況でなく、従って、非言語情報などの詳細な分析には適していない。その意味で、両環境で収録されたデータは相補的であると考えられる。

<sup>2</sup> 写真課題は[矢野96]と同じものであり、課題の特徴もまとめられている。また、そこで得られた会話データの特徴の分析については[矢野99]参照。本収録との主な相違点は、本研究では参加者が三人であること、会話の活発化のために「ヒント」を用いていることなどである。



【図1】収録機材配線図



【図2】対話環境



【図3】写真課題

- この二人は
- a. きょうだいである
  - b. 結婚四年の夫婦である
  - c. 初対面の他人同士である

場合(条件 FEE)と、三人が平等に持つ(内容は異なる場合(条件 ABC)の二通りを準備し、これを上記の被験者組み合わせ(a)~(d)の間でバランスするようにした。

**自由課題**

**教示(要約):** お互いに自由に会話する。いつ誰が話すか、どのような話題について話すかは被験者の自由。「話題リスト」の内容を自由に利用してもよい。

話題リストを用意したのは写真課題の場合と同様の理由による。話題は[中澤 02]を参考に、大学生活、趣味や好みなど、被験者が話題にしやすく、自己紹介的な項目を選択した。

**4. 分析のためのデータ整理—転記・タグ付与**

収録された音声データは、AB, BC, CAをそれぞれL, Rチャンネルとして、16kHz, 16bitステレオで電子データ化し、各対話ごとに3つの音声ファイルを作成した。各ファイルを1チャンネル1話者のステレオファイルとしたのは、以下の理由による。

1. 転記作業の際に重複発話箇所が特定しやすい
2. 転記者が同時に再生して話者の聞き分けを行うためには2チャンネルが望ましい。

転記作業基準は[堀内 99]及び[CBRDD00]を参照しつつ、適宜改良を加えたものを用いる。転記は【図4】のようになる。転記ファイルにはさまざまな談話タグ[荒木 99][CBRDD00]が付与される。そのうち特に談話行為タグは対話進行過程を反映するものであるが、三者以上の会話においては、これらの談話行為が「誰へ」のものであるかを体系的に記述できることが望ましい。

**5. 分析のための基本的観点**

**5.1 分析の第一段階—会話分析の装置系とその詳細化**

対話データからの帰納的 一般化においては発見のための作業仮説が不可欠であるが、そのための手がかりとして、会話分析

09:43:960-09:51:970 A: \*1 でもな、なんか外書とかさ:全然さ; もうなんか<in>もうほんま、なんか、その場で直訳直訳直訳\*2 状態や  
\*3 からの、全然わからへん。

09:50:160-10:00:700 C: \*2<in>  
\*3 そうそう:そう。\*4 とゆうか、あれ、<noise>読んで、読んだ時点から詳しく説明してほしいよね、\*6 あの、訳した時点から。

09:50:560-09:51:270 B:  
\*3 う:ん

09:52:720-09:55:170 B:  
\*4(? んん)で、そんなん、しゃあないんちゃう?契約法とかわから\*5 んもん。

09:54:330-09:55:630 A:  
\*5 わからへん。

09:58:620-09:59:400 B:  
\*6 そうそう  
そうそう。

#### 【図 4】 転記データの例 (03\_2d)

分野で発見・定式化されてきた話者交代システムや隣接ペアなどの知見を利用することができる<sup>3</sup>。しかし、特に三者以上の会話を分析する場合、これらの会話分析の枠組みについて、以下に述べるような精緻化が必要であると考えられる<sup>4</sup>。

対話データ分析の第一段階では、発話間の関係を認定していくことを通じて、対話を階層的に構成する作業が有効である。発話間の関係としては、会話分析において提唱されている隣接ペア *adjacency pair* などがある<sup>5</sup>。隣接ペアとは、例えば、「依頼」と「受諾」のように、異なる話者による連続した二つの発話からなる発話対で、参加者がある目的を共同で遂行するための基本単位と見なすことができる。基本的に、隣接ペアは 2 人の参加者によって形成されるため<sup>6</sup>、三者以上の会話においては、当該の隣接ペアの共同構築に直接関わらない参加者がいることになる。この点を話者交代システム *turn-taking system* との関連で考察する。[Sacks74]によって定式化された話者交代システムは概略以下のようなものである。

- ターン構成単位における最初の移行適格場所 TRP で、
- (1a) IF 「現行話者が次の話者を選ぶ」ならば、  
THEN 選択された参加者が次のターンを取る権利を持ち、義務を負う。移行はここで行われる。
- (1b) IF 現行話者が次話者を選ばないならば、  
THEN 次の話者についての「自己選択」が可能。最初に話し始めた者がターンを得て、移行はここで行われる。
- (1c) IF 現行話者が次話者を選ばず、かつ他の者が自己選択を行わないならば、  
THEN 現行話者が話し続けることが可能であるが、義務的ではない。
- (2) IF 最初の TRP において、ルール(1a)や(1b)が機能せず、

(1c)の条項に従って現行の話者が話し続けるならば、THEN 以降の TRP において、移行が行われるまで (a)-(c)が循環的に再適用される。

理論上隣接ペアの第二部分は(1a)「現行話者による次話者選択」によって生成されるものである。しかし、二者会話では聞き手は常に一人であるため、現行話者のターンの完了点が自動的に次話者のターン開始時点となることが多く、ターンの移行に関する(1a)「現話者による次話者選択」と(1b)「次話者による自己選択」の区別が問題とされることは少ない。

また、これに応じて、厳密な意味での隣接ペアの場合と、隣接ペアの第二部分とは区別されるべき種類の応答との間の相違も不明確となる。「現話者行による次話者選択」の場合に比べて、「次話者による自己選択」による発話にはさまざまな種類のものが考えられる。

- (1) A: もうこんな時間か。  
B: そんなにイライラしないで。

この会話の流れはきわめて自然であるが、この例は厳密には隣接ペアではない。なぜなら、

- ◆ 隣接ペアは第二部分の応答を待って完結する単位であるが、Aの発話は自己完結的であり、独り言でも構わない。
- ◆ 第一部分が関連する第二部分を要請するということには第二部分を発話するのに適切な次話者も選択されているということが伴うが、Bの発話内容はこの状況に居合わせたものならば誰が発話しても構わないものである。

このように、話者交代規則における選択の区別や隣接ペアとそれ以外の発話間関係を明確に区別するためには、発話の持つ「聞き手デザイン」を反映した記述体系が必要となる。

#### 5.2 聞き手デザイン

[Clark82]によれば、三人以上の参加者がいる会話においては、聞き手の役割は様々ではない。Aによる隣接ペアの第一部分によって選択された第二部分の話し手 B は A の発話の受け手 *addressee* としての地位を有するのに対して、もう一人の参加者 C は傍参加者 *side-participant* として区別される。三人以上

<sup>3</sup> 本稿で言及する会話分析の諸装置に関しては[石崎01] [高梨01a]参照。  
<sup>4</sup> 以下で提示する理論的枠組みは[高梨01b]をもとに、その要点を簡潔に述べたものである。  
<sup>5</sup> 類似の概念には、バーミンガム学派の発話交換構造がある。両者の比較に関しては[石崎2001]の第6章参照。  
<sup>6</sup> 後述のように、三人目の参加者がいずれかの発話の産出を補助・代行する場合もあり、こうした場合には、三人目の話者は他の参加者のうちのいずれかと「チーム」を形成していると考えられることができる。

の参与者を含む会話では、話し手はさまざまな言語的・非言語的手段によって、どの参与者がどの聞き手役割(受け手/傍参与者など)を引き受けるべきかを割り当てる必要がある。こうした工夫を発話の持つ聞き手デザイン audience design と呼ぶ。ある発話の時点で次話者となることが意図されていない傍参与者も当該の発話を理解できている必要がある<sup>7</sup>。

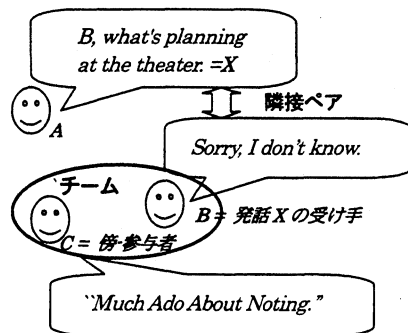
また、三者会話においては、三人が対等な参与者である「一人が一人ずつに向けて話す」状態以外にも、「二人が一人に向けて話す」状態や「一人が二人に向けて話す」状態などの多様な位相が現われることもあり、こうした状況では二人以上の参与者が状況に応じて柔軟に「チーム」を編成しているといえる。[図5]の例では、BとCが「共語り手」としてAに應對している。チーム編成という観点から重要なのは、これが三者以上の会話においてある時点までは傍参与者だった者が発話する資格を得るメカニズムの一つを解明する手がかりとなるためである。

従来の二者対話コーパスにおいては、発話者は記されているものの、誰がその発話の直接の受け手であるかが明記されることはなかった。二者会話の場合には、現行話者でない方の参与者=受け手であると自動的に想定されていたためである。また、ある発話の直後の発話が元の話し手によって意図された受け手による発話である場合と、次話者が受け手以外の立場から発話を開始している場合とを区別した分析が行われることもほとんどなかった。三者会話を分析する際には、発話に含まれる聞き手デザインから意図された次話者を特定する方法を確立し、予測される次話者と実際の次話者が一致しない場合、この次話者の言語行為にはどのような特徴があるかを分析することが必要である。

## 謝辞

三者対話収録に関して、最初にその意義を示唆していただいた榎本美香氏(千葉大学大学院)、予備収録の環境を提供していただきました石崎雅人氏(北陸先端科学技術大学院大学)、収録環境を見学させていただきました松坂要佑氏(早稲田大学)、収録機材等に関してさまざまな相談に乗っていただきました野口広彰氏(ATRメディア情報科学研究所)、柴田雅光氏(株)インターグループ)、課題についてご教授いただきました矢野博之氏(通信総合研究所)、収録作業の運営に携わっていただきました大原裕子氏(株)インターグループ)、機材オペレータの村上誠氏、収録会場の利用を許可していただきました石田亨氏(京都大学/デジタルシティ研究センター)の各氏、及び、収録場所に関してご協力いただきましたデジタルシティ研究センター二条ラボの皆様我感到いたします。本収録作業は科学技術進歩事業団CREST「デジタルシティのユニバーサルデザイン」と通信総合研究所の共同研究の一環として行われた。

<sup>7</sup> 仲介エージェントを含む三者会話において、エージェントの傍参与者としての会話能力が低い場合に生じる問題については[中澤02]参照。



【図5】 チーム編成

## 参考文献

- [荒木99] 荒木雅弘・伊藤敏彦・熊谷智子・石崎雅人. 発話単位タグ標準化案の作成. 人. 知能学会誌 14(2): 251-260.
- [Archer80] Archer, D. *How to Expand Your S.I.Q.* M. Evans and Company, Inc. (藤原・市村英次訳. ポディーランゲージ解析法 誠信書房, 1988)
- [CBRDD00] 人. 知能学会 談話・対話研究におけるコーパス利用研究グループ. 様々な応用研究に向けた談話タグ付き音声対話コーパス. 人. 知能学会研究会資料 SIG-SLUD-9903-4 19-24.
- [Clark82] Clark, H.H. & Carlson, T.B. Hearers and speech acts. *Language*, 58: 332-373. also in Clark, H. H. 1992. *Arenas of Language Use*. U. of Chicago Press & Center for the Study of Language & Information.
- [堀内99] 堀内靖雄・中野有紀子・小磯花絵・石崎雅人・鈴木浩之・岡田美智男・仲貞純子・土屋俊・市川薫. 日本語地区課題対話コーパスの設計と特徴. 人. 知能学会誌 14(2): 261-272.
- [石崎01] 石崎雅人・伝藤輝. 談話と対話言語と計算③. 東京大学出版会.
- [石崎99] 石崎雅人・加藤恒昭. 多人数対話の特徴分析—情報交換パターンと主導権の観点から—. 人. 知能学会研究会資料 SIG-SLUD-9901-3 13-19.
- [小磯00] 小磯花絵・伝藤輝. 円滑な話者交代はどのようにして成立するか—会話コーパスの分析にもとづく考察—. *認知科学* 7(1): 93-106.
- [松坂01] 松坂要佑・小田勇一郎・小林哲則. グループ会話におけるマルチモーダル会話データの収録・解析システムの構築. 人. 知能学会第15回全国大会論文誌 3B2-02.
- [中西00] 中西英之・Isbister, K.・石田亨・Nass, C. 仮想空間でのコミュニケーションを補助するヘルパーエージェントの設計. *インタラクシオン2000*. 107-114.
- [中澤02] 中澤諭・中西英之・石田亨・高梨克也. 社会的エージェントのバランス理論. *インタラクシオン2002*.
- [Sacks74] Sacks, H., Schegloff, E.A. & Jefferson, G. A simplest systematics for organization of turn-taking for conversation. *Language*, 50(4): 696-735.
- [高梨01a] 高梨克也. 社会的相互行為を「見る」方法. 人. 知能学会誌 16(6): 799-805.
- [高梨01b] 高梨克也. 三人会話原理—聞き手デザインと参与の編成. *社会言語科学会第8回研究人会予稿集* 255-260.
- [矢野99] 矢野博之・伊藤敏. 創発的な対話に関するコーパスの構築. *自然言語処理* 6(4): 117-137.
- [矢野96] 矢野博之・伊藤敏. 協調型タスクにおける非言語情報の使われ方. *情報処理学会ヒューマンインターフェース研究会* 65(3): 9-14.