

実対話データベースの設計と構築

— アルバイト募集およびパーソナルナビゲーションにおける電話対話 —

熊本忠彦 伊藤 昭

郵政省 通信総合研究所 関西先端研究センター

〒651-24 兵庫県神戸市西区岩岡町岩岡 588-2

近年、話し言葉を対象とする自然言語処理の分野において対話コーパスを利用した研究が盛んに行われている。しかしながら、従来利用されている対話データは、模擬的な実験環境において収集された対話(模擬対話)であり、実際の状況下で得られるであろう対話(実対話)と言語的・対話的性質が同じであるとは限らない。そこで、我々は、模擬対話と実対話の定性的、定量的な共通点/相違点を明らかにするための材料として、2つの対話領域において模擬対話と実対話をそれぞれ収集した。本稿では模擬対話と実対話を規定する要因について考察し、実対話を収集するための実験環境を設計する。また、この設計に基づいて実際に行った実対話収集実験についても述べる。

Collecting Spontaneous Dialogues in an Experimental Environment

— Collecting telephone dialogues related to getting a part-time job and telephone dialogues for personal navigation tasks —

Tadahiko KUMAMOTO and Akira ITO

Kansai Advanced Research Center, Communications Research Lab.,
Ministry of Posts & Telecommunications

588-2, Iwaoka-cho Iwaoka, Nishi-ku, Kobe-shi, Hyogo 651-24, Japan

In recent years various kinds of dialogue corpora have been constructed and used for studies on dialogue and spoken language. In conventional dialogue collection, however, systematically collecting a large amount of dialogues and constructing a very large corpus are considered critically important, while the spontaneity of the dialogues collected is not a point to be carefully considered. "Spontaneous dialogue" does not mean that the dialogue itself is spontaneous but means that an utterance in the dialogue is spontaneous. However, it is obvious that the spontaneity of a dialogue determines the structure and aspect of the dialogue. Therefore, in this article, we consider the spontaneity of a dialogue and propose a method for collecting spontaneous dialogues in an experimental environment. We also describe the two kinds of experiments which we actually conducted for collecting spontaneous dialogues.

1. まえがき

近年、音声対話処理や自然言語処理の分野では、対話コーパスを用いた研究 [1]-[3] が盛んになってきており、そのためのコーパスもいくつかの大学・研究機関から提供されている [4]-[7]。しかしながら、従来の取り組みでは、対話データを如何にして体系的かつ大量に収集し、大規模なコーパスを構築するかという点に主眼が置かれており、収集される対話の品質についてはあまり考慮されていない [7][8]。実際、従来の対話データは、模擬的な実験環境において収集された対話(模擬対話)であり、実際の状況下で得られるであろう対話(実対話)と言語的・対話的性質が同じであるとは限らない。

模擬対話収集のための実験においては、対話参加者(被験者)は、自分の発話が録音されていることを意識しているであろうし、また時には何らかの役 — 例えば店員の役やお客の役、または事務員の役など — を演じている場合もある。このような模擬的な実験環境で行われた対話が実際の場面で行われている対話と同じであるとは考えにくい、その反面、何がどう異なるのか明らかでない。

そこで我々は、模擬対話と実対話の定性的、定量的な比較を行うための材料とすべく、2つの対話領域において模擬対話と実対話をそれぞれ収集した。本稿では模擬対話と実対話を規定する要因について考察し、実対話を収集するための実験環境を設計する。また、この設計に基づいて我々が実際に行った実対話収集実験についても述べる。

本稿では実対話と模擬対話の違いを「対話の品質」という観点で捉え、自然度、自由度という2種類の尺度を導入して、論じる。具体的には、自然度、自由度をそれぞれ「自然(自由)である」、「比較的(自然(自由)である)」、「どちらとも言えない」、「比較的(不自然(不自由)である)」、「不自然(不自由)である」の5段階評価とし、自然かつ自由な発話を実発話(spontaneous speech)、自然かつ自由な対話を実対話(spontaneous dialogue)と呼ぶことにする。従来、spontaneous speechという言葉とspontaneous dialogueという言葉

はほぼ同義で用いられているが、本稿では明確に区別し、前者を発話レベルの用語、後者を対話レベルの用語として用いる。これは、模擬対話の中にもspontaneous speechからなる対話がありうることを考えれば、自然な発想と言える。

本稿の以下の構成は次の通りである。まず、2.で対話の品質を規定する要因と、各要因が対話に与える影響について考察する。次に3.で実対話の収集法について検討し、その結果に基づいて行われた実対話収集実験(2種類)について述べる。最後に4.で本稿のまとめと今後の課題について述べる。

2. 対話の品質に関する考察

対話の品質を規定すると思われる要因を、自由度を尺度とする「制約」と自然度を尺度とする「対話状況の設定」とに分け、分類した。その結果を図1に示す。

「制約」とは対話収集実験において被験者に陽に課せられる制約のことであり、その種類としては発話(音声)レベルの制約、発話(テキスト)レベルの制約、対話レベルの制約の3種類が考えられる。発話(音声)レベルの制約とは、発話がnativeによるものか/non-nativeによるものかといったことや、自由発声によるものか/朗読によるものかといった音響的な制約のことであり、音声認識や音声合成のような音声レベルの研究に際し、問題となる。発話(テキスト)レベルの制約は、自然言語処理分野における研究テーマの分類に対応して、更に細かく分類されている。そのため、研究テーマに対応して考慮すべき制約の種類も変わってくる。対話レベルの制約も図1に示されたように細分類されている。対話の深さに関する制約とは入れ子対話のあり方に関する制約のことであり、対話の幅に関する制約とは話題の展開の仕方、例えば話題を唐突に変えることができるか否か。に関する制約のことであり、また、対話の量に関する制約とは対話時間や発話権の移動回数、発話文字数などに関する制約のことであり、発話(テキスト)レベルの制約の場合と同様、対象とする研究テーマの違いによって考慮すべき制約の

- 制約
 - 発話 (音声) レベルの制約
 - 発話 (テキスト) レベルの制約
 - * 形態素レベルの制約 (語彙の選択の仕方など)
 - * 構文レベルの制約 (言い回しの決め方など)
 - * 意味レベルの制約 (発話内容の決め方など)
 - * 発話単位に関する制約
 - * 発話タイミングに関する制約
 - 対話レベルの制約
 - * 対話の深さに関する制約
 - * 対話の幅に関する制約
 - * 対話の量に関する制約
- 対話状況の設定
 - 対話参加者への動機付け
 - * 発話の動機
 - * 対話の動機
 - 対話の前提条件
 - * 対話参加者の役割
 - * 対話の収録環境

図1 対話の品質を規定する要因の分類

種類も変わってくる。

一方、「対話状況の設定」とは対話収集実験における対話環境(被験者を含む)をどう設定するかということであり、設定の種類としては対話参加者への動機付けに関する設定と対話の前提条件に関する設定の2種類が考えられる。対話参加者への動機付けは発話レベルと対話レベルで区別されている。対話参加者の発話に対する動機が自然であるとき、すなわち発話することを強制されたり、義務として発話するのではなく、自発的に自由意思で発話するとき、その発話は発話の動機という点において自然であると言える。同様に、対話参加者の対話に対する動機

が自然であるとき、すなわち対話することを強制されたり、義務として対話を始めるのではなく、自発的に自由意思で対話を始めるとき、その対話は対話の動機という点において自然であると言える。特に、この対話の動機という面については今まで議論されることもなかったが、この動機が自然か不自然かということが対話全般に与える影響は大きいものと考えられる。実対話を収集するための実験環境の設計においては、この対話の動機を如何にして対話参加者(被験者)に自然に与えるかという点に最も留意した。

対話の前提条件は、対話参加者の役割設定と対話収録環境の設定とに分けられる。これらの設定は、対話参加者の心理状態を大きく左右し、対話全般に影響を及ぼすものと考えられる。例えば、対話参加者に何らかの役が与えられるとき、その人はその役に応じた典型的な、別の言い方をすると作為的な振舞いをしてしまうかも知れない。また、対話が収録されていることを対話参加者が意識しているとき、その振舞いは通常のものとは異なった、ある意味で改まった振舞いになるかも知れない。

以上で述べた項目に対し、自由かつ自然な発話を spontaneous speech と呼び、自由かつ自然な対話を spontaneous dialogue と呼ぶ。但し、ここで注意が必要なのは、対話が spontaneous でなくても、発話が spontaneous であることは十分可能ということである。従来の模擬対話収集実験 [4]-[7] では、spontaneous speech が収集されてはいるが、spontaneous dialogue とは言えず、この一例となっている。

3. 実対話の収集について

本章では、実対話の収集法について検討し、その結果に基づいて設計された2種類の実対話収集実験について述べる。

3.1 実対話収集法の検討

前章での考察により、音声認識/合成のような音声レベルの研究や自然言語理解/生成のようなテキストレベルの研究で spontaneous speech を

7日限りの アルバイト募集

日 給 : 7,000円 (税引き後6,300円)
昼食、交通費等は一切支払われません。

雇用条件 : 18歳以上の男女で日本語を母語として話せること。
昼食、高次生、高専生は不可

雇用主 : 郵政省 通信総合研究所 調査支所 知識処理研究室
神戸市西区宮岡町西宮、第二神明大橋1C近く
 JR大久保駅からバスで約10分、西神中央駅からバスで
 約30分、バス停から徒歩約10分。

勤務時間 : 10時から約3時間 もしくは 14時から約3時間
昼食、実験の都合等によっては多少早く終ることも
 遅く終ることもあります。
 なお、勤務時間のうちの1時間は待機時間です。

仕事内容 : 携帯電話を用いた実験および
 TV電話を用いた実験の被験者

勤務日 : 1月13日(月)～1月20日(水)
 2月12日(水)～2月17日(金)
 3月17日(月)～3月25日(金)
の間のいずれか1日、昼食、土曜日を除く。

問い合わせ/受付 : 電話 078-969-2172
通信研究所のアルバイト受付係まで。
 なお、電話以外での問い合わせ/受付には
 一切お答えすることができません。

その他 : 1日に雇用可能なアルバイトは4名まで。
料金はかかりません。

図2 アルバイト募集のチラシ (原寸はB4判)

対象とするときには、従来の模擬対話収集実験で得られる spontaneous speech からなる模擬対話でも十分と言える。しかしながら、対話レベルで観測される言語現象の分析や対話管理、対話制御、話題管理などの対話レベルの研究を行うときには、spontaneous dialogue を収集する必要がある。

spontaneous dialogue を収集するためには、図1に示したような制約がなく、また対話状況の設定も自然であることが要求される。このような要求を満たすためには、「日常のいろいろな場面でなされている実際の対話を録音する」という方法がもっとも確実ではあるが、情報通信倫理などの問題から対話を録音できる状況は限られており、何らかの工夫が必要である。例えば、文献[1]ではラジオ番組で放送された電話相談対話を録音することによって spontaneous dialogue を得ているし、また文献[9]では実働の対話システムを用いたフィールドテストによって spontaneous dialogue を収集している。

しかし、文献[1]の方法では対話ドメインの決め方が受け身であり、対話システムの利用価値が高い対話ドメインを選択することができない。また、文献[9]では実働の対話システムを用いているため、現在のシステム水準に応じた対話データしか得られず、人間同士が行うような対話を実現するためのノウハウや知識を得ることは難しい。

そこで我々は、対話が必然的に必要となるような場を作り出すことによって、spontaneous dialogue を収集することを試みている。具体的な実験仕様については次節以降で述べるが、普段の生活でもよく見掛けるような対話の場面を設定している。特に、3.2の対話収集法では対話参加者は自分が被験者となっていることも意識していない。なお、これらの対話収集法で重要なことは、対話参加者(被験者)にとって対話何かが遂行するための手段であり、目的ではないことである。このような状況は、「被験者には対話すること以外の課題を与え、その課題を達成するための手段の一つとして対話を用意する」という方法[10]によって容易に作り出せる。

3.2 アルバイト受付対話の収集

ある実験の被験者を募集するために、図2に示されたチラシ(原寸はB4判)を当センター最寄り駅から3駅離れた駅の周辺(1万世帯)に新聞折り込みで配布した。その結果、66名からアルバイトの申し込み/問い合わせの電話があったが、このときの電話対話をミニディスク(MD)に録音した。アルバイト申込者(66名)の内訳(性別/年齢構成)は表1の通りである。また、電話対話の書き起こし例を図3に示す。なお、図3において読点「、」は短いポーズを、3点リーダ「…」は長いポーズを表している。また、句点「。」は文の終りを表している。

従来の対話収集実験では被験者は大学生であることが多いが、大学生は一般社会における平均的な人物より高い問題解決能力を有しており、対話も比較的論理的かつ効率的に行われるものと考えられる。今回我々が行った実験では、アルバイト申込者は年齢層も幅広く、また新聞折

表1 アルバイト申込者の性別と年齢

性別	年齢	申込者数
女性	18～19	4
	20～29	19
	30～39	13
	40～49	9
	50～59	2
	不明	7
男性	20～29	3
	30～39	1
	不明	8
合計		66

秘書: はい、通信総合研究所アルバイト受付係です。

申込者: あっ、すいません。

秘書: はい。

申込者: あの…広告で、見たんですけどね。

秘書: あっ、はい、ありがとうございます。

申込者: あはっ、あの…わたくし

秘書: はい。

申込者: あの、50歳…になってるんですけど。

図3 アルバイト申込者と秘書の電話対話

り込みという無差別的な手段で集められているため、より平均的な対話が収集されたものと考えられる。

なお、本節および次節で収集された対話に対しては、その使用許諾を実験終了後に得た。使用許諾を得られなかった一部の対話(表1において年齢不詳になっている対話)については破棄した。

3.3 携帯電話によるパーソナルナビゲーション対話の収集

前節の方法で採用された計58名のアルバイト(電話申込者の友人、配偶者、親子という場合もあった)に対し、以下のような実験を行った。

- (1) 指定の時間に当センターの最寄り駅に集合してもらう。但し、一度に2人ずつ集合し

てもらう。

- (2) 実験前の教示としてその場で、当実験の課題が集合場所から当センターへの単独行であること、可能な移動手段が徒歩とバスであること、道順を携帯電話で我々の秘書に随時聞けること、および携帯電話の使い方を説明した。説明に要した時間は大体2～3分程度であった。なお、実験の目的などについては一切話していない。また、教示内容に関係のない質問に対しては、実験終了後に回答すると説明した。
- (3) 被験者2名のうちの1人が実験を開始し、もう1人は実験者の車で当センターに移動した。但し、このときの車での移動ルートは被験者が通るであろうルートとは異なるルートであり、また研究所への進入にも別の門を利用した。
- (4) 最初の被験者が当センターに到着したら、次の被験者を集合場所まで連れていき、同じ実験を行った。

当実験によって得られた電話対話の一例を図4に示す。また、当実験に参加した被験者の性別と年齢を表2に示す。

4. むすび

本稿では、模擬対話と実対話の違いを「対話の品質」という観点で捉え、自然度、自由度という2種類の尺度を導入することによって spontaneous speech と spontaneous dialogue のあり方を定義した。また、spontaneous speech や spontaneous dialogue を収集するための実験環境を設計し、この設計に基づいて2種類の実対話収集実験を実施した。1つは、アルバイト募集のチラシを新聞に折り込み、それに応募してきた人たち(66名)と秘書との電話対話を録音するというものであった。もう1つは、以上の電話対話の結果採用された人たち(58名)に当センターの最寄り駅に集合してもらい、そこから当センターまでを、秘書の携帯電話による誘導を

秘書: はい通信総合研究所です。
 被験者: あっすいません市バスのバス停まで来たんですけども。
 秘書: あっはいえーとそしたらそのバス停に国道大久保って書いてありますよね。
 被験者: えーと…こえっ
 秘書: 国道大久保。
 被験者: あっあ書いてあります。
 秘書: はいえーとそうしましたら
 被験者: はい。

図4 パーソナルナビゲーション対話の例

表2 ナビゲーション対話参加者の性別と年齢

性別	年齢	被験者数
女性	18 ~ 19	5
	20 ~ 29	22
	30 ~ 39	15
	40 ~ 49	9
	50 ~ 59	2
男性	20 ~ 29	3
	30 ~ 39	2
合計		58

受けながら一人で実際に移動してもらおうというものであった。そして、このときの電話対話が録音された。

現在までに実対話だけでなく、模擬対話の収集も終えている。但し、模擬対話とは、アルバイト募集タスクにおいては、別途雇用した被験者に新聞に折り込んだチラシと同じものを見せ、そのアルバイトに応募するフリをしてもらうというものであり、もう一方の道案内タスクにおいては、被験者に当センターの最寄り駅から当センターまでを移動するフリをしてもらうというものであった。どちらの対話領域においても、実対話収集実験と模擬対話収集実験の違いは、実際の行動なのか、それともフリなのかと点だけであり、その他の実験条件は同じであった。また、フリをするにあたって、被験者は実際には実験用の個室に座しており、そのことはもう一方の対話参加者である秘書も知っていた。現在

は、以上のようにして収集された実対話および模擬対話の対話データ(ミニディスクに録音された音声データ)の書き起こしを行っている。今後の課題としては、以上のようにして収集された実対話および模擬対話の定性的、定量的な比較・分析が挙げられる。

参考文献

- [1] 佐川, 大西, 杉江: “自己修復を含む日本語不適格文の分析とその計算機による理解手法に関する考察”, 情処学論, 35, 1, pp. 46-52 (1994).
- [2] 中川, 小林: “自然な音声対話における間投詞・ポーズ・言い直しの出現パターンと音響的性質”, 音響誌, 51, 3, pp. 202-210 (1995).
- [3] 伝, 春木, 石崎: “音声対話コーパスを用いた言い直しの分析”, 情処研報, SLP 16-2, pp. 7-12 (1997).
- [4] 江原, 小倉, 篠崎, 森元, 樽松: “電話またはキーボードを介した対話に基づく対話データベースA D Dの構築”, 情処学論, 33, 4, pp. 448-456 (1992).
- [5] 小林, 板橋, 速水, 竹沢: “日本音響学会研究用連続音声データベース”, 音響誌, 48, 12, pp. 888-893 (1992).
- [6] Thompson, H. et al.: “The HCRC map task corpus: Natural dialogue for speech recognition”, *Proc. DARPA Human Language Technology Workshop*, pp. 25-30, San Francisco, USA (1993).
- [7] 田中, 速水, 山下, 鹿野, 板橋, 岡: “RWC計画における音声対話データベースの構築”, 情処研報, SLP 11-7, pp. 37-42 (1996).
- [8] 白井, 竹沢: “音声対話処理”, 人知学誌, 9, 1, pp. 48-56 (1994).
- [9] 黒岩, 武田, 井ノ上, 山本: “機械との対話における発話分析—内線電話受付システムにより収集した対話データの分析—”, 信学技報, SP94-30, pp. 57-64 (1994).
- [10] Kumamoto, T. et al.: “Design and construction of an advisory dialogue database”, *IEICE Trans. Info. & Syst.*, E78-D, 4, pp. 420-427 (1995).