

話し言葉対話データベース構築のための 実データの収集について

Collecting Real Dialogues for a Spoken Dialogue Database

熊本 忠彦 伊藤 昭 矢野 博之 小嶋 秀樹
Tadahiko KUMAMOTO Akira ITO Hiroyuki YANO Hideki KOZIMA
{kuma, ai, yano, xkozima}@crl.go.jp

郵政省 通信総合研究所 関西支所

Kansai Advanced Research Center, Communications Research Laboratory, MPT of Japan

〒 651-24 神戸市西区岩岡町岩岡 588-2

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

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][2] 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研究機関で構築されている [3]-[5]。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組織的、体系的に収集し、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関する議論は少ない [5][6]。

特に、話し言葉における言語現象の分析や話し言葉を対象とした処理手法の開発を目的として対話を収集するときには、その対話の品質が分析の結果や開発される処理手法のアルゴリズムに影響を与えてしまうため、対話データの利用目的に応じた対話収集法が必要とされる。そこで本稿では、収集される対話の品質を規定する要因について考察し、その結果に基づいて高品質な対話(本稿では一定の基準を満たす高品質な対話を実対話と呼ぶことにする)の収集法について検討する。また、その収集法に基づいて実際に我々が行っている対話収集実験についても述べる。

なお、本稿では対話の品質を対話そのものの品質と対話を構成する発話の品質とに分けて考え、それぞれにおいて spontaneous なデータを高品質、そうでないものを低品質と定義する。このように対話の品質を対話レベルの品質と発話レベルの品質とに分けて扱うという考え方は、従来あまりなされていないが、spontaneous speech からなる対話と spontaneous dialogue は異なるという我々の

考えから来ている。また、対話の品質を決めるのに spontaneous であるか否かという尺度を導入するのは、日常の対話場面では spontaneous speech からなる spontaneous dialogue が行われているからである。

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、それ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ている対話収集実験(3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

我々は、対話の品質を発話レベルと対話レベルとに分けて考えているが、それぞれのレベルにおいて対話の品質を規定すると思われる要因を表1および表2にまとめた。但し、我々は発話(speech)を音声(voice)という音響的側面と発話文(spoken sentence)という記号的側面とに分けて考えており、発話レベルの考察では発話の記号的側面だけに着目している。これは、音声処理では発話が native によるものか/non-native によるものかということや、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違いだけが問題となり、対話の品質とは無関係と考えられるからである。

我々は、ある発話や対話が spontaneous なものであるかどうかは、そのデータを収集するための実験の仕様によって規定されると考えている。表

表1 対話の品質を規定する発話レベルの要因

形態素に関する制約 (語彙の選択の仕方など)
構文に関する制約 (言い回しの決め方など)
意味に関する制約 (発話内容の決め方など)
発話単位に関する制約 (文単位/文節単位, 断片文の受理など)
発話の動機 (自発的か強制的/義務的か)

1に示された発話レベルの制約(4種類)のいずれかが被験者に課せられるとき, その発話は spontaneous であるとは言えない。また, 発話の動機が自発的なものではなく, 発話することを(例えば課題として)強いられているときも, その発話は spontaneous であるとは言えない。同様に, 表2に示された対話レベルの制約(2種類)のどちらかが被験者に課せられるならば, もしくは対話の動機が自発的なものでないならば, その対話は spontaneous であるとは言えない。逆に言えば, それぞれのレベルにおいて表1または表2に示されたような制約がなく, また発話/対話の動機が自発的なものであるならば, その発話/対話は spontaneous であるということになる。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[3]-[5]では, 発話レベルでは制約がないが, 対話レベルでは何らかの制約があり, 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, その結果に基づく収集法を提案する。この収集法は, 実験環境の設定が比較的容易なため, どこの研究機関でも実施することが可能である。

3.1 実対話収集法の検討

前章の考察より, 発話レベルで話し言葉を研究対象とするとき, すなわち音声認識や音声合成, 発話理解や発話生成に関する研究では, 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十分と言える。しかしながら, 対話レベルで話し言葉を研究対象とするとき, すなわち対話管理や対話制御, 話題管理などに関す

表2 対話の品質を規定する対話レベルの要因

対話構造に関する制約 (入れ子対話, 発話権の移動など)
対話の方向性に関する制約 (話題の展開, 対話の流れなど)
対話の動機 (自発的か強制的/義務的か)

る研究を行うときには, 発話レベルだけでなく, 対話レベルでも spontaneous な対話(実対話)を収集する必要がある。

実対話を収集するためには, 発話レベル, 対話レベルで共に制約がなく, また発話の動機, 対話の動機という面でも自発的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有効ではあるが, 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実対話を得ているし, また文献[7]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実対話を収集している。

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[7]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, 実対話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の対話収集法では話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, 話者(被験者)にとって対話が何かを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, その課題達成の手段として対話を用意し, このときの対話を録音する」という方法[8]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

3.3 および3.4で述べる対話収集実験の被験者

1日限りの アルバイト募集

日 給：7,000円（税引き後6,300円）
但し、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国語として話せること。
但し、高校生、高専生は不可

雇用主：郵政省 通信総合研究所 関西支所 知識地理研究室
神戸市西区宮原町西側、第二神明大久保1C近く
JR大久保駅からバスで約10分。近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
但し、実験の進み具合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、勤務時間のうち約1時間は待ち時間です。

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

勤務可能日：1月13日（月）～1月30日（水）
2月12日（水）～3月7日（金）
3月17日（月）～3月25日（水）
の日のいずれか1日。但し、土日祝を除く。

問い合わせ/受付：電話 078-969-2172
当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。
再雇用はできません。

図1 アルバイト募集のチラシ（原寸はB4判）

を募集するために、図1に示されたチラシ（原寸はB4判）を当支所最寄り駅から3駅離れた駅の周辺（1万世帯）に新聞折込みで配布した。その結果、65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。なお、個人の名前や住所などプライバシーに関する情報を含んでいる対話については、各人に対し事後承諾を得た。

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60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) 指定の時間に当支所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支所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、道順を携帯電話で当支所アルバイト受付係に随時聞けること、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

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こ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2 パーソナルナビゲーション対話の例

度であった。なお、実験の目的などについては一切しゃべっていない。また、教示内容に関係のない質問に対しては、実験終了後に説明すると回答した。

- (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支所へ移動した。但し、このときの車での移動ルートはバスの運行ルートとは別ルートであり、また研究所への進入にも別の門を利用した。
- (4) 最初の被験者が当支所に着いたら、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。
- (5) 実験終了後、携帯電話の利用経験や当支所の場所を知っていたかというこについてアンケートを実施した。

当実験によって得られた実対話の一部を図2に示す。但し、データ整理の途中のため、書き起こしミスや表記方法の揺れがあるかもしれない。

3.4 マルチメディア対話の収集

前節の実験を終了した2人にペアになってもらい、次のようなマルチメディア対話収集実験を行った。

- (1) 実験前の教示として、当実験の課題がナビゲータ側のみ提示されているブロック図（図3の上図参照）をオペレータが自分の作業画面（図3の下図）に完成させることであること、ナビゲータとオペレータの席は完全に隔離されていること、二人はマイクとイヤホンを用いてお互いに対話ができること、オペレータのみがマウスでブロックを動かせること、ナ

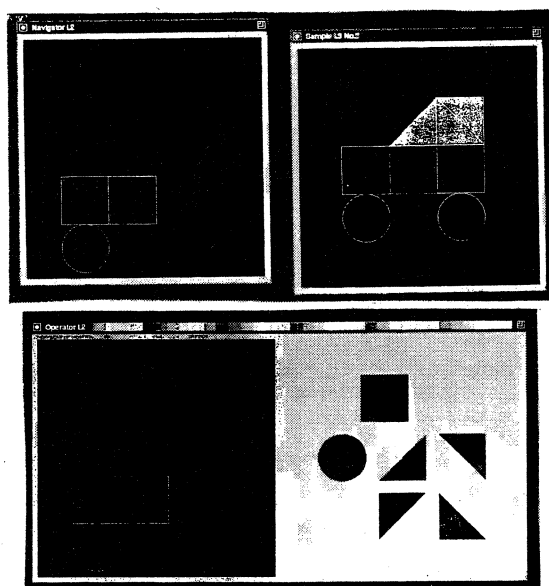


図3 積み木タスクの画面の一例
(上がナビゲータ側, 下がオペレータ側の画面)

ビゲータ側にはオペレータ側の操作の様子が表示されることを説明した。説明に要した時間は3, 4分であった。また, マウスの使い方を教え, 各自1回ずつ, 実際にブロックを動かしてもらった。

- (2) 一方がナビゲータ役, もう一方がオペレータ役になり, 4回実験を行った。1回毎にブロック図を変え, またナビゲータ役, オペレータ役という立場も入れ替えた。対話データとしては, このときの両被験者の顔画像および計算機画面の様子をビデオカメラで撮影し, 4分割画面という形で1本のVTRテープに録画した。

4. むすび

本稿では, 実対話を spontaneous speech からなる spontaneous dialogue と定義し, また spontaneous な speech や dialogue がどうしたら得られるかについても考察した。その結果として, 実対話を収集するための実験では,

- ・発話の仕方や内容に関する制約がない,
- ・発話することを強制されない,

- ・対話の構造や方向性に関する制約がない,
- ・対話することを強制されない,

ことが重要であることを示した。また, そのような条件を満たす対話収集方法についても検討し, 現在実施している実対話収集実験についても触れた。

今回の実験で収集される対話コーパスの規模は比較的小規模であり, 有意意味な統計データを抽出するには足りないかも知れない。しかし, 本稿で述べた対話収集法は他の研究機関でも容易に実施でき得るものであり, お互いのデータを持ち寄ることにより, 大規模な対話コーパスを作成することができるものと期待される。

参考文献

- [1] 佐川雄二, 大西昇, 杉江昇: “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] 中川聖一, 小林聡: “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,” 日本音響学会誌, 51, 3, pp. 202-210 (1995).
- [3] 江原輝将, 小倉健太郎, 篠崎直子, 森元逞, 樽松明: “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ADDの構築,” 情処学論, 33, 4, pp. 448-456 (1992).
- [4] 小林哲則, 板橋秀一, 速水悟, 竹沢寿幸: “日本音響学会研究用連続音声データベース,” 日本音響学会誌, 48, 12, pp. 888-893 (1992).
- [5] 田中和世, 速水悟, 山下洋一, 鹿野清宏, 板橋秀一, 岡隆一: “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).
- [6] 白井克彦, 竹沢寿幸: “音声対話処理,” 人工知能学会誌, 9, 1, pp. 48-56 (1994).
- [7] 黒岩真吾, 武田一哉, 井ノ上直巳, 山本誠一: “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).
- [8] Kumamoto, T., Ito, A. and Ebina, T.: “Design and construction of an advisory dialogue database,” IEICE Trans. Info. & Syst., E78-D, 4, pp. 420-427 (Apr. 1995).