

行為を伴う実対話と伴わない模擬対話の比較

An Analysis of Real and Simulated Dialogues in a Task Domain

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

〒 651-2401 神戸市西区岩岡町岩岡 588-2

^{††} 山形大学 工学部

〒 992-8510 米沢市城南 4 丁目 3-16

1 はじめに

近年、音声対話処理や話し言葉処理の分野では、対話コーパスを用いた研究 [1]-[4] が盛んに行われており、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 [5]-[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するかという点やクリアな音声データを如何にして収録するかという点に主眼が置かれており、収集される対話の自然性についてはあまり考慮されていない [8]。実際、従来利用されている対話データは、その多くが模擬的な対話環境において収集されたフリだけの対話 (模擬対話) であり、対話参加者 (被験者) が何らかの役 (例えば店員の役や事務員の役など) を演じたり、何かをするフリをしたりしている。このような人工的な実験環境で収集された模擬対話が実際の状況下で行われる行為を伴う対話 (実対話) と言語的・対話的性質においてまったく同じであるとは考えにくい、その反面、何がどう異なっているのかわからない。

そこで我々は、実対話と模擬対話を携帯電話を用いた道案内というタスク領域において収集し、実対話参加者と模擬対話参加者の意識の差をアンケート調査に基づいて比較する。また、両対話の言語的・対話的性質に関しても分析を行い、その定量的な比較を試みる。

2 携帯電話を用いた道案内タスクにおける対話の収集

本章では、携帯電話を用いた道案内タスクにおいて行われた実対話と模擬対話の収集実験について述べる。

被験者 105 名のうちの 57 名に対し実対話収集実験を、残りの 48 名に対し模擬対話収集実験を

行った。それぞれの実験に参加した被験者の性別と年齢を表 1 にまとめる。

[実対話収集実験]

- (1) 各被験者に当センターの最寄り駅に来てもらった。
- (2) 実験前の教示として、その場で以下のことを 2～3 分程度で説明した。
 - 実験の課題がその駅から当センターへの単独行であること
 - 移動手段が徒歩とバスに限られること
 - 道順を携帯電話で秘書に随時質問できること
 - 携帯電話の使い方
- (3) 実際に移動してもらった。

[模擬対話収集実験]

- (1) 各被験者に当センター内の 1 室に入ってもらった。
- (2) 実験前の教示として、その室で以下のことを 2～3 分程度で説明した。
 - 実験の課題が当センターの最寄り駅から当センターへの単独行のフリであること
 - 想定されている移動手段が徒歩とバスであること
 - 道順を携帯電話で秘書に随時質問できること
 - 携帯電話の使い方
- (3) 移動するフリをしてもらった。

以上の実験において行われた秘書と被験者の電話対話をミニディスクに録音した。実対話の例を図 1 に、模擬対話の例を図 2 に示す。なお、対話データの利用に際しては、各被験者から承諾書を得ている。

表 1: 被験者の性別と年齢

性別	年齢	実対話	模擬対話
女性	18～19	5人	0人
	20～29	21人	11人
	30～39	15人	19人
	40～49	9人	13人
	50～59	2人	3人
	60～70	0人	1人
男性	20～29	3人	0人
	30～39	2人	0人
	60～70	0人	1人
合計		57人	48人

3 アンケート調査に基づく比較

本章では、各被験者に対して実施したアンケート調査に基づいて、実対話参加者と模擬対話参加者の意識の差を比較する。

今回のアンケート調査では、「携帯電話による誘導について」という項目で、「わかりやすかった」(5点)から「わかりにくかった」(1点)までの5段階評価をしてもらい、その評価の理由を自由記述形式の欄に書いてもらった。以上のアンケートの結果を表2にまとめる。表2には道案内に対する5段階評価の平均得点と標準偏差、および自由記述欄への記入率(記入した人の数÷被験者数)、一人当りの文字数(記入された文字の数÷記入した人の数)が示されている。但し、文字数の計数において句読点や?マークは数えなかった。

表2より、自由記述欄への記入率は模擬対話の方が高いが、記入された文字数は実対話の方が多くことがわかる。これは、模擬対話参加者の方が被験者であるという意識が強く、アンケートに回答することを責務として捉えたためと考えられる。その反面、書きたいことや書けることは実際に移動した実対話参加者の方が多く、文字数の差となって現れたものと考えられる。一方、道案内に対する評価は実対話の方が良かった。これは無事移動できたという事実が評価を良くしているためと考えられる。

次に、自由記述欄に記入された評価理由をその内容に応じて分類した。この分類において、実対話参加者は言及したが、模擬対話参加者が言及しなかったものとしては、「携帯電話の利用」、「道に迷う可能性」、「移動距離」の3項目があった。特

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: あっあっ書いてあります
 秘書: あっはい、えーと、そうしましたら
 被験者: はい

図 1: 実対話の例

秘書: はい、通信総合研究所アルバイト受付けです
 被験者: あっもしもし
 秘書: はい
 被験者: すいません、今大久保駅の
 秘書: はい
 被験者: 北口タクシーのりばなんですけれども
 秘書: あっはい

図 2: 模擬対話の例

に「道に迷う可能性」や「移動距離」には「初めて駅に降り市バスに乗ったが迷わずに來れた」や「徒歩キョリが少ないせいもあると思うが」といった実際に移動した人しか思い付かないような記述が含まれていた。

一方、実対話参加者は言及しなかったが、模擬対話参加者が言及したものとしては、「説明の長さ」、「説明のスピード」、「経路知識の有無」、「移動可能性」の4項目があった。「説明の長さ」には「一挙に説明するのではなく3段階で説明してくれたから」や「何回にも分けておしえてもらったので」といった記述が分類されており、駅から当センターまでの経路を一度に説明するのではなく、想定されている現在地に応じて説明を分けたことが評価されている。しかしながら、実対話ではこのような状況に即した説明は当然かつ自然なことであるため、誰も言及しなかったものと考えられる。また、「説明のスピード」には「説明の仕方が早くて頭の中に思い浮かべる時間よりも聞く時間の方が短くて忘れてしまう事が所々あったので電話では少しわからないところがありました」という記述が分

表 2: アンケートの結果

実対話参加者 (57 人)	
携帯電話による誘導について (5 段階評価)	4.74 (平均得点) 0.43 (標準偏差)
自由記述欄の記入率	63.2%
自由記述欄の文字数	38.2 文字/人
模擬対話参加者 (48 人)	
携帯電話による誘導について (5 段階評価)	4.48 (平均得点) 0.77 (標準偏差)
自由記述欄の記入率	70.8%
自由記述欄の文字数	27.2 文字/人

類された。実際に移動しながらの対話であったなら、説明と状況が一致するので、「頭の中に思い浮かべる」必要もなく、このような記述はされなかったものと考えられる。一方、「経路知識の有無」には「その場所を見たことがないから」や「場所を知っているから想像がついたのでわかりやすかったと思う」といった記述が分類され、「移動可能性」には「多分あの説明でいけると思う」や「初めての場所はわかりにくいかも知れない」のような記述が分類された。この 2 項目に分類された記述は、実際に移動しなかったために生じたコメントと言える。

実対話では駅から当センターまで移動することが被験者の主目的であり、そのため、説明の長さやスピードといった体裁よりも、説明の内容や聞き取りやすさなど実質的なことに関心が寄せられている。一方、模擬対話では対話すること自体が主目的であるため、内容だけでなく、体裁にも言及されている。また、実際に移動したらという観点でのコメントも多かった。

4 言語的・対話的性質に関する分析

本章では、携帯電話を用いた道案内タスクで収集された実対話と模擬対話を分析し、言語的・対話的性質の定量的な比較を試みる。

まず、実対話・模擬対話のそれぞれにおける秘書と被験者の平均発話文字数を調べた。その結果を表 3 に示す。表 3 より、秘書の発話文字数は、実対話であるか模擬対話であるかに関係なく、ほぼ同じであったが、被験者の発話文字数は実対話の方が 2 割以上多かったことがわかる。これは、模

表 3: 平均発話文字数

文字/人	実対話	模擬対話
秘書	1180	1162
被験者	466	376

表 4: 間投表現の出現頻度

間投表現	発話者	実対話	模擬対話
はい	秘書	46.1 回/人	49.7 回/人
	被験者	50.1 回/人	48.9 回/人
えっ	秘書	1.6 回/人	0.5 回/人
	被験者	2.9 回/人	1.3 回/人
えっと	秘書	12.6 回/人	9.4 回/人
	被験者	2.1 回/人	1.2 回/人
あっ	秘書	9.5 回/人	5.3 回/人
	被験者	11.6 回/人	7.5 回/人
あの	秘書	8.9 回/人	7.2 回/人
	被験者	5.0 回/人	4.3 回/人

擬対話では曖昧な説明や誤解があっても、被験者自身困らないが、実対話では道に迷ってしまうかも知れず、それゆえ実対話参加者の方が状況説明や情報獲得に対し、より積極的であったためと考えられる。

次に、実対話と模擬対話における間投表現の出現頻度を調べた。調査の対象とした間投表現は、肯定や相槌の意で用いられる「はい」と、文献 [1] や文献 [2]、文献 [10] において利用頻度が高いことが示された「え」、「えっと」、「あっ」、「あの」の 4 種類であった。但し、「え、えっ、えー」などを「えっ」として扱うなど、表層的または意味的に似ている表現は同じものとして扱われている。

以上の調査の結果を表 4 に示す。表 4 より、「はい」は、実対話であるか模擬対話であるかに関係なく、同程度に用いられていることがわかる。また、対話データ中の「はい」の使われ方を調べてみると、肯定や相槌としてだけでなく、相手に対する主導権を渡したり、対話のリズムを整えたりするためにも用いられており、対話の自然さを演出するのに役立っているように感じられる。

一方、「えっ」、「えっと」、「あっ」、「あの」は実対話の方が高頻度であった。特に、驚きを表す「えっ」において、出現頻度の差 (実対話における出現頻度 ÷ 模擬対話における出現頻度) は顕著であり、秘書発話の場合で約 3.1 倍、被験者発話の場合で約 2.1

倍の差があった。これに対し、ためらいや呼び掛けを意味する「あの」では秘書発話、被験者発話ともに約 1.2 倍程度の差しかなかった。間投表現の出現頻度において、実対話の方が高頻度なのは、模擬対話では秘書、被験者ともに心構えができており、ゆとりをもって対話に参加したため、つなぎの語である間投表現を必要とする場面が比較的少なかったものと考えられる。逆に、実対話では被験者は移動しながらの対話であり、状況に即した質問／応答を実時間で要求されたため、考えながらの発話文生成となり、つなぎ語としての間投表現がより多く用いられたものと考えられる。また、「はい」を除く間投表現(4 種類)の中で出現頻度が高かったのは、実対話であるか模擬対話であるかに関係なく、秘書発話の場合で「えっと」、被験者の発話の場合で「あっ」であった。「えっと」は思案するときや言葉に詰まったときに用いられる語であり、秘書が被験者の状況を考えながら、適切な指示／応答を行おうと努力している様子が伺える。一方、「あっ」は何かを急に思い出したり、驚いたりしたときに思わず発する語であり、被験者が秘書の指示／応答をその場の状況と照らし合わせながら、対話している様子が伺える。

5 まとめ

本稿では、携帯電話を用いた道案内タスクにおいて行われた実対話と模擬対話の収集について述べ、実対話参加者と模擬対話参加者の意識の差をアンケート調査に基づいて比較した。その結果、模擬対話では説明の内容だけでなく、説明の長さやスピードといった体裁にも言及されていたが、実対話では体裁よりも説明の内容や聞き取りやすさなど実質的なことに関心が寄せられていることがわかった。また、実対話と模擬対話を分析し、その言語的・対話的性質を比較した結果、実対話参加者の方が発話文字数が多かったことや、「はい」の出現頻度は、実対話であるか模擬対話であるかに関係なく、同程度であること、「え」や「えっと」、「あっ」、「あの」は実対話の方が高頻度であり、移動の状況に即した発話文生成を実時間で行う際のつなぎ語として活用されていることがわかった。また、「はい」は、肯定や相槌としてだけでなく、相手に対話の主導権を渡したり、対話のリズムを整えたりするためにも用いられており、対話の自然さを演出するのに役立っているように感じられた。なお、実対話収集実験では各被験者に当センター

の最寄り駅から当センターまで実際に移動してもらい、模擬対話収集実験では当センター内の 1 室で移動するフリをしてもらった。このとき、もう一方の対話参加者である秘書は、被験者が実際に移動しているのかフリだけなのかを知っていたが、同じように誘導するよう要求された。

現在、音声データの書き起こし時に生じるエラーや揺らぎを削除するための努力を行っており、これが終了した後、各対話の言語的・対話的性質に関する本格的な分析を行う予定である。なお、今回収集された実対話・模擬対話をコーパスにして、公開・配布する計画を立てている。

参考文献

- [1] 村上, 嵯峨山: 自由発話音声における音響的な特徴の検討, 信学論, J78-D-II, 12, pp. 1741-1749 (1995).
- [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, Mo. 3, pp. 202-210 (1995).
- [3] 熊本, 伊藤: 支援対話の解析に基づく対話処理方式の提案, 信学論, J77-D-II, 8, pp. 1492-1501 (1994).
- [4] 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).
- [5] 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).
- [6] 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).
- [7] Pao, C. and Wilpon, J.: "Spontaneous speech collection for the ATIS domain with an aural user feedback paradigm," *Proc. of DARPA Speech and Natural Language Workshop*, pp. 43-48, New York, USA (1992).
- [8] 白井, 竹沢: 音声対話処理, 人工知能学会誌, 9, 1, pp. 48-56 (1994).
- [9] 妙木, 松本, 長尾: 汎用日本語辞書および形態素解析システム, 第 42 回情処全大講演論文集, 1C-9, pp. 3-17-3-18 (1991).
- [10] 熊本, 伊藤: 支援対話に現れた言語現象の分析, 情処研報, 97-NL-119, pp. 91-96 (1997).