

日英語対照研究のための名称なし日本語地図課題対 話コーパスの作成：予備実験と転記テキスト

著者	吉田 悦子
雑誌名	人文論叢：三重大学人文学部文化学科研究紀要
巻	18
ページ	173-195
発行年	2001-03-25
その他のタイトル	ニチエイ タイショウ ケンキュウ ノ タメノ メイ ショウ ナシ ニホンゴ チズ カダイ タイワ コーパ ス ノ サクセイ ヨビジッケン ト テンキ テキスト Constructing an Unlabelled Japanese Map Task Dialogue Corpus for a Contrastive Study of English and Japanese : A Pilot Experiment
URL	http://hdl.handle.net/10076/1876

日英語対照研究のための名称なし日本語 地図課題対話コーパスの作成 — 予備実験と転記テキスト — *

吉田悦子

要旨 本稿は、1997年三重大学において収録された「名称なし日本語地図課題対話コーパス」（以下、「名称なし日本語コーパス」）作成のための予備実験の概要および収録対話の転記テキストを資料として公表するものである。名称なし日本語コーパス作成の目的は、1. 日英語対照研究を可能にする「より自発的で自然な」日本語の対話コーパスを作成し共有できるデータとして電子化すること、2. すでに千葉大学で完成した日本語コーパスでは出現が限られている指示表現に注目し、その分布と用法を記述、考察、議論し、被験者間の相互作用的なコミュニケーション活動として機能していることを例証すること、である。予備実験および転記テキスト作成作業全体を通して、実質的かつ技術的な問題点として、音声情報の欠落および不正確さ、収録条件の不統一性、転記テキスト記述上の不正確さが指摘された。この3つの問題点をふまえて、より信頼できうる対話収録とコーパス作成のために改善すべき点は、2チャンネル音声データの確保、必要最低限の厳密な音声情報の記述、収録条件の統一、目的に応じた転記テキストの基準の確立などである。

1 はじめに

本稿は1997年三重大学において収録された「名称なし日本語地図課題対話コーパス」（以下「名称なし日本語コーパス」）作成のための予備実験の概要および収録対話の転記テキストを資料として公表するものである。さらにこの転記テキストの作成について留意すべき点について考察し、本実験への橋渡しとしたい。

1.1 背景

近年さまざまな研究領域において対話への関心が高まっている。人のコミュニケーション活動を探るという目的にとって、対話のコーパスは信頼できるデータとして欠かせない。にもかかわらず自発的かつ自然な対話の日本語データの作成と利用環境は未だ十分ではない。千葉大学ではエジンバラ大学 Human Communication Research Centre (HCRC) で作成された英語による対話コーパス（以下「エジンバラコーパス」）¹ に刺激を受け、日本語地図課題対話コーパス（以下、「千葉大コーパス」）を完成させた。² この千葉大コーパスはさまざまな研究領域において具体的なデータを提供するだけでなく、日本語と英語という異言語間での対照研究を可能にした。しかしながら、このコーパスでは出現がかなり制限される言語現象もある。そのひとつが指示表現の多様な用法である。指示表現の分布と談話構造との関係に焦点を当てた言語研究の遂行のためには、対話者のやりとりから読み取れる音声言語情報のよりどころとなるコーパスが不可欠であり、小さな規模でもより指示表現の現れやすい環境での言語データの

収集と蓄積が求められている。したがって、すでに進行中の英語のコーパスプロジェクトとほぼ同じ方法と規模で実験を行い、そのデータを日英語対照研究のために共有できる形で構築しておくことが必要となった。

1.2 日本語対話データ収集の目的と意義

名称なし日本語コーパス作成の目的は、以下の通りである。

1. 日英語対照研究を可能にする「より自発的で自然な」日本語の対話コーパスを作成し共有できるデータとして電子化すること
2. 千葉大コーパスでは出現が限られている指示表現に注目しその分布と用法を記述、考察、議論し、被験者間の相互作用的なコミュニケーション活動として機能していることを例証すること

もちろん指示表現の全体像をこのコーパスのデータだけから把握することは不可能である。しかし、対話というジャンルに出現する指示の分布と用法を観察し、そのコミュニケーション上の役割を談話構造と関連づけ、具体的なデータで例証できることが最優先される。

名称なし日本語コーパス作成の意義は、自然な対話を記録した日本語の資料が極めて限られている中において、小規模ではあっても日本語の話しことばのデータ蓄積に貢献できること、そしてデータの共有が実現することである。したがって将来的に作成する日本語コーパスは、エジンバラ、千葉大のオリジナルコーパスの基本設計を踏まえながらも、あえて地図上の目標物の名称を省いた地図課題を用いた実験をおこなう。このため被験者がより自由に地図上の目標物を指示する余地が生まれてくることになり、指示の種類と分布形態に関してオリジナルコーパスとは明らかな違いが出てくる結果として予想される。

2 予備実験の概要

2.1 地図課題について

地図課題は英語、日本語共通のものがのぞましい（図1参照）。今回の予備実験では、地図上の目標物の位置と経路はオリジナルと同じものを使用した。目標物のイラストは異なるものを使用した。堀内ほか（1997：3）から千葉大コーパスで使用した地図課題の概要を引用する。

地図課題は二人の実験参加者により共同で達成される課題である。まず、各参加者には一人一枚ずつ地図が与えられる。そのとき、一人の実験参加者（以下、情報提供者：Instruction Giver）の地図には経路が記されているが、もう一人の実験参加者（以下、情報追随者：Instruction Follower）の地図には経路が記されていない。二人の実験参加者に課される課題は、相手の地図が見えないように向かい合い、お互いに会話を交わしながら、情報提供者の地図上の経路を情報追随者の地図上に再現することである。なお、情報提供者の地図には経路以外に出発地点と目標地点、そしていくつかの目標物とその名称がえがかれている。一方、情報追随者の地図には、出発地点と目標物およびその名称だけが描かれており、目標地点と経路は描かれていない。二つの地図は完全に同一ではないが、そのことは実験の前にあらかじめ実験

この箇所の図版は著作権の
関係で削除した

図 1

参加者に明示的に教えられている。しかし、二つの地図がどのように異なっているかは伝えられておらず、課題の遂行過程で解決すべき問題の一つとされる。(アンダーラインは著者)

オリジナルコーパスで用いた地図課題と今回の予備実験で用いた地図課題の重要な違いは、「地図上の目標物に名称をつけていない」という点にある。すべての実験参加者は、お互いを視認できるが、相手の手と地図を見ることはできない。³ また身振り手振りを相手に示すことは禁じられている。それぞれの対話は異なる地図課題に基づいており、3におけるように A から D の 4 種類の対話テキストが得られた。

2. 2 実験参加者

実験参加者は 20 歳から 32 歳までの男女計 8 人の三重大学人文学部の学生である。

4 種類の対話 A から D における実験参加者の性別、年齢、出身地、および親近性条件⁴ は以下の通りである。

A

情報提供者 : Giver : 男 25 才 出身地 群馬

情報追随者 : Follower : 女 32 才 出身地 三重

親近性なし

B

情報提供者 : Giver : 女 21 才 出身地 愛知

情報追随者 : Follower : 女 32 才 出身地 福井

親近性なし

C

情報提供者：Giver：女 20才 出身地 三重

情報追随者：Follower：女 20才 出身地 愛知

親近性あり

D

情報提供者：Giver：女 21才 出身地 愛知

情報追随者：Follower：女 21才 出身地 岐阜

親近性あり

2.3 収録の実施

対話の収録は小型カセットテープレコーダー一台で行った。

それぞれの収録時間は以下の通りである。

A 6分21秒

B 13分33秒

C 7分33秒

D 4分50秒

2.4 転記テキストの方法

収録した対話をもとに転記テキストの作成を行った。基本的に転記方法は、厳密な音声情報を除いて、千葉大コーパスに従った（堀内ほか：1997）。⁵ 話しことばの表記であることを考慮して、句読点の使用は一切禁止した。読みやすさという点を重視して、漢字仮名交じり文による記述を行った。また外国語のカタカナ表記については慣用となっているものに使用を制限した。延ばす音の表記も必要最低限におさえた。明らかに疑問を呈している発話にはクエスチョンマークを記した。オーバーラップ（二人の話者の発話が重複する状態）の箇所は「*」で記した。⁶ 笑いながら語彙を発しているものは、通常どおりに転記をおこない、丸括弧内に「(笑)」と記述した。イントネーション・ユニット⁷ がはっきりしているものについては一字分スペースをとり、発話の切れ目を明示した。また2秒以上のポーズ（無音区間）については一律して連続3ドット「…」で記述することとした。

3 転記テキスト

テキスト中の番号は対話者交代に対して用いた。番号の右横に情報提供者（Instruction Giver）はGを情報追随者（Instruction Follower）はFを付記した。

A

1G：えっとじゃー説明します＊

2F： ＊はい

3G：んーとスタート…の地点の下にんーとそばっていうあのそばの店みたいなの見えますか＊

- 4F : * はい
5G : んとその左側スタート地点から左側に下がってください
6F : あはい
7G : でその下にあるうさぎ*
8F : * はい
9G : …の横も通りますずっと下に下がるんですけども*
10F : * あっはいはい
11G : でうさぎの下を*
12F : * はい
13G : うんとう右に曲がってください*
14F : * はい
15G : で…少し行ったらんーと田んぼ*
16F : * はい
17G : 田んぼがありますよね*
18F : * はい
19G : 田んぼーとうさぎの間
20F : はい
21G : を上にあがります*
22F : * あはい
23G : …そんでーあのそばそばー…の建物というか*
24F : * あはい
25G : の辺りを右にまがって*
26F : * はい
27G : んーとしばらく行きます
28F : はい
29G : …あっ右にーなんですけども*
30F : * はい
31G : いやそんで上になんかこう煙の出てる工場みたいな*
32F : * はい
33G : があるんですけども*
34F : * あはい
35G : あのそちらの右にたどり着くように右側に…横を通るんですけども工場の右ってこと
なんですけれどもー
36F : …えーとそそばのお店のー
37G : はい*そばのお店のあたり田んぼとー*
38F : * ところ
39F : * えー
40G : そばうさぎってありますよね
41F : あはいそばうさぎ
42G : 右左と分けたらー

- 43F : はい
- 44G : んーとだからそのー田んぼのー横をこえてー *
- 45F : * はい
- 46G : あのーそばの高さくらいのところに行ったら
- 47F : はい…右に曲がるんですね *
- 48G : * はい右に曲がるん *
- 49F : * でーあのそのまままっすぐいくとー
- 50G : はい
- 51F : 煙の煙の出ているうちがあるんですけどー
- 52G : はいはい
- 53F : そのそのうちは通り過ぎるんですか *
- 54G : * あーそのうちあの右側を少し * 通り過ぎるん軽く
- 55F : * はい
- 55F : 通り過ぎると
- 56G : はいあっほんでさ通り過ぎて左折するんですけども *
- 57F : * あっ左折するん *
- 58G : * はい
- 58G : あの * 工場と煙の間ぐらいのところを通過してもらえると
- 59F : * あっ…うふっ
- 60G : あの *
- 61F : * もう一度行きますけど
- 62G : はい
- 63F : そば屋のところを右に曲がって *
- 64G : * そば屋のところを右に曲がりますよね
- 65F : はい
- 66G : でこれはまったく右にってもらって *
- 67F : * ええ
- 68G : で工場の右があっ右側を通過できるように *
- 69F : * はい
- 70G : …上へあがります
- 71F : ポストと工場の横ですか
- 72G : はい
- 73F : はいはいどうぞ
- 74G : ほんでふうふっ右側通過したら *
- 75F : * はい
- 76G : すぐに左折するんですけども *
- 77F : * はい…は * い
- 78G : * はい
- 79F : はい
- 80G : うふそのときにあのー

- 81F : はい
82G : あそんで*
83F : * はい
84G : えっと…こう少しあの工場は通過して
85F : はい
86G : うーんと少し下がるんですあの左下に
87F : はい
88G : 下がって*
89F : * はい
90G : んでまたこうまっすぐ*
91F : * はい*
92G : * 左側に*
93F : * はい
94G : ほんでごみ見えますよね
95F : 見えます
96G : うーんとごみの左側を
97F : はい
98G : 通るように上に上がっ*てください
99F : * 上がるはい
100G : でごみは通過して*
101F : * はい
102G : …右折します
103F : はい
104G : ほんで何か家みたいなのみえますよね
105F : …うーん* (笑)
106G : * 家というか建物というか横が横に木のある
107F : 横に何がある
108G : 木
109F : 木のある*
110G : * はい*
111F : * うーん建物…ないですけど*
112G : * あっ建物ないんですか
113F : 踏み切りあります うさぎいます 山見えます
114G : …はいあ ということはこの絵がないんですねーうー*んとそうなるにあのうさぎの
*
115F : * はい * はい
116G : うーんと…少し左下ぐらいに*家があるんですけど*
117F : * はいなるほど * はいはいはい
118G : はいあのそこの家の左側を通過して
119F : はいはい…ふんはいはい*

- 120G : *うさぎの左下ぐらいなんですけども*
- 121F : *はいはいはい
- 122G : でそこ…家を*
- 123F : *はい
- 124G : その左側を通過して*右折します
- 125F : *えー
- 126F : それはうさぎの下に行くということですか
- 127G : うさぎの下あ*あの一
- 128F : *家を通り過ぎて*
- 129G : *はい*
- 130F : *右折すると*
- 131G : *はい
- 132F : うさぎの下に行く
- 133G : うさぎの下までは行かないんですけども*
- 134F : *ああはいはい
- 135G : もうちょっと下なんですけれども*
- 136F : *あはい
- 137G : でそのあたりから郵便ポストが見えますか
- 138F : あーはい
- 139G : あの一*に向けてその家*をこえたら行くんですけども
- 140F : *うふっ *うふっ(笑)はいうひひ*
- 141G : *うふっ
- 142F : えっポストに行けばいいんですね
- 143G : はい*ポストの下に向けて行ってください
- 144F : *ポスト はいどうぞ(笑)ポストの下*はい
- 145G : *でポ はい ほんでポストの下を通過して
- 145G : *うーんと上へ上へというか販売機みたいな見えますか
- 146F : *はい う*ー
- 147G : *ありませんか
- 148F : うー
- 149G : やまーは二つありますか
- 150F : やまは二つあります
- 151G : やまーの
- 152F : はい
- 153G : 右側のやまの下*に
- 154F : *はいはい
- 155G : あの販売機が*あるんですけども
- 156F : *はい
- 157G : そこに向けて上にあがっ*てください
- 158F : *あっはいはい

- 159G : …でその販売機の左側を*
160F : *はい*
161G : *通るようにして*
162F : *はい
163G : あの二つ目のやまに向けて*
164F : *はい
165G : …行ってください*
166F : *はい
167G : で二つ目のやまあ左側*の
168F : *はい
169G : やまの*
170F : *はい
171G : うーんとふもとまできたら*
172F : *はい
173G : …ゴールです
174F : ゴール
175G : はい…できません*か
176F : *たどりたどりつきましたが*
177G : *という
178F : はい…終わりです でも絵がちがうたぶんちがうかな

B

- 1G : じゃあいきます
2F : はい
3G : えっと…神社の
4F : うん
5G : えー…北北…北北西からスタートします
6F : 神社の北北西から? (笑)
7G : はい えっ北北西1センチぐらいから (笑)
8F : 北北西1センチぐらいから?
9G : はい 神社を中心に北北西の*
10F : *うん
11G : 1センチぐらいからスタートして えーとそこから南に…南に南に向かっ
11G : てください えへへ
12F : ずっと真っすぐ下に?
13G : えーと ちょっと南に えー 1, 2, 3, 4, えー5センチぐらい
14F : 5センチぐらい
15G : 5センチぐらい少し (笑) えーと少し東側にカーブをしながら
16F : 東側にカーブをするよ*うに
17G : *少しゆるいカーブをえがきながら5センチ

- 18F : あの神社の下のところにあるこれは このあそこにたっするの？
- 19G : あーたっしませんたっしませんえーとそれでこの教会みたいなのを*
- 20F : *うん
- 21G : ぐらいまでその5センチっていうのは教会ぐらいまでのことで
- 22F : うん
- 23G : そこに向かってとりあえず南に 真南にゆるいカーブをえがき ふっ
- 24F : ゆるいカーブをえがきながら*
- 25G : *あっ 東側 東 あ ちょっと東東より？
- 26F : うん
- 27G : 左っていうか あー右右よりにこう描いて下さい 描きながら5センチぐ
- 27G : らい 5センチぐらいそこからーその下の学校に平行に…その学校ぐら
- 27G : いに平行に*
- 28F : *うん
- 29G : 2センチか3センチていうところは学校の時計のちょうど真上あたりま
- 29G : で*
- 30F : *うん
- 31G : えー平行ー そのその平横に横っていうか…右に真っすぐ その時計の
- 31G : 真上ぐらいまで行って下さい
- 32F : えっごめんちょっと待って*えっとー…スタート地点が*神社の所の*絵よりも上の
所やんね
- 33G : *はい *うん *うん
- 33G : 神社の*北 少し上
- 34F : *うん 少し上やんね
- 35G : そこからー少し*
- 36F : *そこから教会の方に
- 37G : え そ 位置的にはまっすぐで*平行 あ 真っすぐの線を引い スタートからとり
あえずそこまでのゴールまでは*真っすぐなんだけど
- 38F : *うん*うん
- 38F : あっ真っすぐなの うん
- 39G : そのそのまっすぐの達するのはまっすぐなんだけど その2点の線直線で
- 39G : 結ぶあー線で結ぶとまっすぐなんだけど*少し右側にカーブをつけて*そこ
- 39G : まで達して
- 40F : *うん *うん
- 40F : 教会の これは教会のどこまで達すればいいの？上の十字の所？
- 41G : ではなくって*このま あのー窓みたいなのぐらいまで
- 42F : *うん ああのでかい窓の所まで？
- 43G : 大きい窓 あのー格子のある窓まで行ってえーそれで*そこからまっ そ
- 43G : の地点からまっすぐ…右へ*右でちょうど時計の全然時計のその学校みた
- 43G : いなやつとは全然距離があるんだけど*そこっえーとそこ それに平行
- 43G : に2センチくらい行って…

- 67G : 左手のバス停*バス停の左手で*えーとこの え今この一海岸線ぞいにつ
67G : ていうのは分かった？
68F : *うんうん *うん
68F : うーん海岸線ってやつ？
69G : たぶんその海岸線ってやつ それで今度は その地点から*そのパラソル
69G : のとかき氷の間を通って*
70F : *うん *うん
71G : そこ 間を通ってほしい そこからぐるーっとカーブで…ちょっと半そ
71G : の半島の*ラインにそうようそうようなカーブを*かき氷とパラソルの間
71G : を通って欲しい
72F : *うん *うん
72F : パラソルってさあ*かき氷のどどころら辺にある？
73G : *うん
73G : あっ パラソルは ない*あっ えーと じゃあどうしよう あのーバ
73G : バス停はある？
74F : *ないうん
74F : うん バス停はある
75G : バス停の所から* 1つぼこってなってるでしょ 海が海っていうか海
75G : 岸線が ちょっと
76F : *うん
76F : えっと バス停の左側の所？
77G : んーと 下の方バスの真下に少しこうなって*うん うんうん そしたら
77G : もう1つもう1つゆるーくなってるでしょ
78F : *窪んでる所？
78F : ゆるーいゆるーい えっと 私のこっちの地図だったら バス停があっ
78F : て…ひだりっ側 バス停がこの紙の右の隅の方にあるんさ
79G : あー 2つない バス停？
80F : 1個しかない*うんうん*
81G : *1個しかない
81G : *そっかぁ
81G : そのバス停ではない* (笑) 私のバス停は えーと じゃあサルまで分かる？* サ
ルのさあ 左上に信号機ある？
82F : *そのバス停ではない
82F : *うん
82F : ない ふふ
83G : んと じゃあサルから サルまでは分かる？
84F : うんていうかまずこいのぼりがなかったんさ*それで この教会の所の一
84F : 応上を通ってったんやけど
85G : *あー
85G : ほんのちょっと上でいい

- 7G : そのちょうど真ん中くらいからです。そ*こがスタート地点です*
- 8F : * 真ん中から * うん
- 9G : でちょっと一右と*か左とかいっていいんだよね*
- 10F : *うん *いいんじゃない?
- 11G : 左のほうにいくと*一なんか変*な モスクみたいな モスクみたいなん
- 12F : *うん *ろうそくみたいな? モスク?
- 13G : なんかよくわかんないけど
- 14F : 長細いの?
- 15G : 長細くなーい 下のほう
- 16F : 下?
- 17G : うん 下のほうに
- 18F : ビニールハウス?
- 19G : ビニールハウスの一よ 左
- 20F : なんにもないよ うっはっはっは (笑)
- 21G : え なんにもないの?
- 22F : え? ビニールハウス?... ビニールハウスってさー細長いやつ?
- 23G : 長細い三つみたいなやつ
- 24F : 三つ?
- 25G : うん
- 26F : なんか ふつー 半円みたいなのがまるまるまるってあってえ
- 27G : んーそうそうそう それがビニールハウス
- 28F : うん えー私 かみーかみちがう? 方向 たてむきだよ
- 29G : ……え?
- 30F : や か かみぜんたいさぁ
- 31G : かみ全体? たてむきだよ
- 32F : んーだよ
- 33G : あーこれ図はいっしょだねえ きっと
- 34F : だよねえ えー? スタート地点がそのビニールハウスと*一ちょっと右上の
- 35G : * ビニールハウスと
- 36F : 建物の真ん中なんだよね んでそこからどこい*くって?
- 37G : *そうそうそうそう
- 37G : で ビニールハウスの左のほうにさぁー*小さな建物みたいなのない?
- 38F : *うん
- 38F : ビニールハウスの隣?
- 39G : 左
- 40F : 左?
- 41G : ずーっととなりーで 左のとなり
- 42F : ないよ うっはっは (笑)
- 43G : えっ じゃあ どう
- 44F : え? (笑)

- 45G : 11 個図がある？
46F : 1, 2, 3, 4, …11 うん あるある
47G : 図いっしょのはずだよねえ
48F : うん…ビニールハウスがねえ一番下にぼつんであってえ その上に小
48F : さな建物とおー*
49G : *うん
49G : ろうそくがのってそうなの？
50F : んでそのまわりにねえなんかトンネルみたいな
50G : あーこれろうそくがのってそうなの？にみえんのかなーとにかくその左の
50G : やつ？*
51F : *うんうん
52G : そこから下を回ってー左ずーっと左にすすんでー* その下を回*って
53F : *うん *うん
54G : そのトンネルみたい*なをーくぐる形に？す*すむ
55F : *うん *うん
56G : でー… そっちいったらさっきの建物みたいなところ？*その前をとお
56G : って*ー
57F : *うん *うんうん
58G : ちょっと上のほうにーすすんでーいき*つつ テントみいたいのがあるよね
59F : *うん
59F : うん
60G : でーぐるうーっとすす…んーなんていうのかなあ？ちょっと上のほう
60G : に進んでいってーテントみたいなやつ 行くのに下がって行ってテント
60G : みたいなやつの前をとおる
61F : んーなんかさあ 谷みたいなものがあるやんね
62G : あもう 下まで行かないけどーあそこのほうに進んでいくように見えて
62G : テントの下をとおる
63F : あー
64G : んーテントの下をとおるそんでーカルガモかなんか分かんけどペリカ
64G : ンみたいなやつが三匹ー
65F : ペリカン？（笑）
66G : ちょ 目ーみえてる？
67F : なーに？ペリカンって
68G : 鳥みたいなやつ
69F : 鳥？んー？テントはあるよ
70G : テントあるよねえ あーじゃー一緒だねえ これ わかんないけど ま その
71F : ペリカン… とり……
72G : テントの下にあるんですけ*どー まいーやとにかくテントの左下の方を
73F : *な ないよ
74G : ぐるっと迂回するように

- 75F : テントの左下をぐるっと迂回* んで？
76G : *うん
76G : そのように進んでいってー * ごーとずーと上に行くんだけど * 左手に
77F : *うん *うん
78G : 右手に山が見えるやんねえ？
79F : うんうんうんうんうん
80G : んでその横をずーと行って*
81F : *うん
81F : 山 山を右手に見るの？
82G : 山を右手に見るように進んで行*って んでまあとにかく気にせずに進んでい
83F : *うん
84G : ってー*左手にー竜巻みたいのが みえるやんかぁ * みえないかなぁ？
85F : *うん *うん
85F : あのねえー竜巻は一番右端にある
86G : 二つあるんだってえ 私のやつ
87F : ないよ？
88G : ないの？
89F : (笑)
90G : まーとりあえずーまーその右手にある山をちょっとばかしいったところ
90G : でー くっと*右側に曲がってえ でー*また山あるよねえ *その山
91F : *うん *うん *うん
92G : を こえていく*
93F : *こえる*
94G : *うん*
95F : *上を？
96G : 上を山の上を越えていく* で 山の 山を降りて*いく 山を登り
97F : *うん *山を
98G : 山を降りる
99F : うん
100G : うんで…そのーさっきさ*ぁ 谷底みたいになんかあるっていつ
100G : たじゃんかあ*
101F : *うん
101F : *うんうんうん
102G : それをちょっと大きく 迂回する形で
103F : 迂回*
104G : *うん ぐるーっとまわって*
105F : *どっち 右回り左回り
106G : えーっとなんていうの うん 最初からみて左っかわを通っていく
107F : 左っかわっていうんかなあ
108G : ぐるっ…だから山を降りた時点でぐるっとまわる ぐるっとまわると

- 108G : すると えーっなんていうんだろ 山そのまま *
109F : *うん
110G : 紙をふつうに見た感じで左っかわを通るんだよ *
111F : *紙をふつうに見た感じ
111F : で左っかわ…
112G : 左っかわを通る左っかわを通して *
113F : *紙の？
114G : ぐるーって迂回して *
115F : *迂回して
116G : ちょっと行ったところに竜巻あるやんね それより手前でストップ おわり *
117F : *ふん
118G : と これでいいやんね そこフィニッシュ書いて 最初のところスタートって書いて

D

- 1G : はいうんとスタートから神社の鳥居の左側をまっすぐ下におりてで右手に
1G : まがっていくとお墓があるからお墓とあと西郷さんらしき人のまんなかを
1G : 上に…上がってって
2F : え？お墓？
3G : お墓お墓らしきものお墓じゃないのかなこれ
4F : お墓らしきものはない *ないよ
5G : *え？えーへへへ（笑）ない？スタートから神社の鳥
5G : 居わかる？
6F : 鳥居はうん鳥居はわかる
7G : 鳥居の *左っかわを *下におりていくやんね *
8F : *うん *うん *うん下におりていくと
9G : そんでなんか…なんか…にここマーク *と鳥居の真ん中を *右 *に曲がって *いく
とそうするとなんか富士山の上んとここにお墓みたいなのやつ あ ない？鳥居の右下に
お墓じゃないの？ビル？お墓？ははは *え？ *
10F : *うん *右に *うん
10F : *えー？ *お墓？
11G : これお墓じゃない？ビル？
12F : わからない
13G : 富士山の上 え西郷 *さん
14F : *え なんかね 夫婦岩があるよ
15G : え ちがう 富士山の上
16F : ふじ *
17G : *一番上の…富士山の上
18F : 富士山の上には
19G : うん
20F : 一番上の富士山？

- 21G : あー上のほうの富士山の上 上の富士山の上 西郷
22F : 西郷さんはわかるんやって*
23G : *その左は?
24F : その左には*
25G : *何?
26F : 鳥居
27G : …ふふ (笑)
28F : (笑) なんで? だってさ西郷さんは西郷さんはわかるやん ここになんか
28F : 犬もっとる人おるやん*
29G : *うん犬がおる*
30F : *その右っかわにさ古墳みたいなのあるや
30F : んそのこと?*
31G : *ひだ
32F : かぎ*あなってゆーのこれ?
33G : *左?
34F : うん西郷さんの*右っかわ*変な首がうによーって長くて*
35G : *うん *うん *うんそれじゃな
35G : くって左 西郷さんの*
36F : *西郷さんの左は*
37G : *鳥居しかない?
38F : 鳥居しかないんだってー
39G : あ そーなんだ じゃいいや* (笑)
40F : * (笑)
41G : で鳥居の横を もちよっと待って 最初からいくと 鳥居の横を*
41G : 下にいくよね*で*
42F : *うん
42F : *うん*にこここマークと鳥居とそのあいだをいくんだよ
42F : ね?
43G : うん で*右に曲がっていくんだってー
44F : *うん
44F : うんうんわかる うん
45G : そうすると西郷さんの*上を*こうやって囲むように*
46F : *うん*うん *西郷さんの上を囲む
46F : ように?
47G : にー*だんだん上にいって右にいって
48F : *はいはい
48F : うん右にいって
49G : 下にいってって
50F : うん下にいって
51G : で ろくろみみたいなやつと西郷さんのあいだを通過って で今度一番上っか

- 51G : わの富士山の下を通過して
- 52F : … (笑) 一番上っかわの富士山って? *
- 53G : *あ 富士山と上にあるやんね
- 54F : 一個しかない
- 55G : (笑) んーじゃあじゃ ろくろくびのところを*左に行くよね 左にいくと
- 55G : 夫婦岩があるよね
- 56F : *うん
- 56F : あーがーってななめにつつきっていいの?
- 57G : うん ななめにつつきって*で夫婦岩の左を通過して
- 58F : *がー
うん あーなんか にこにこマークのあいだにね
- 59G : そうそう そこを夫婦岩側にかーって通過してって*そうする
- 59G : と下に富士山がある?
- 60F : *うん
- 60F : うん あるある
- 61G : その上のほうを
- 62F : うん
- 63G : 右に*
- 64F : *夫婦岩とにこにこマークのあいだをわーって行って下にかーって
- 64F : って*
- 65G : *下にいて
- 66F : 富士山の上まで?
- 67G : 富士山の上のほう 上と富士山と夫婦岩の真ん中へん*を右に曲がって
- 68F : *うんうん
- 68F : 真ん中へんを右に曲がって
- 69G : で 田んぼと*なんか鐘つき…がある?…ない? 右手のほうに
- 70F : *うん
- 70F : 木 木ならある
- 71G : 木の上ってなにもない?
- 72F : うん なにもない
- 73G : その上もなにもない?
- 74F : 木いーがあって*あ なんか空間があって*田んぼみたいなところ
- 75G : *うん *うん
- 75G : じゃ 田んぼの下を通過して*で 下におりてって*木の右手を通過して
- 76F : *うん *うん
- 76F : 木の右手を通過して
- 77G : で 木の下を通過して 左に曲がって
- 78F : うん…うん
- 79G : ってすると お地蔵さんが…ある?…ない?
- 80F : お地蔵さんは富士山のひだりにある*

- 81G : *あ そうそうそうそうそうそう
82F : そこまでいって ねねこは無視?
83G : …うん (笑) ねこは無視 (笑) でお地蔵さんの下を通過して
84F : うん
85G : で 左下にいくとフィニッシュ*
86F : *なにして? フィニ*
87G : *お地蔵さん*
88F : *お地蔵さんと富士
88F : 山のあいだの…
89G : ううん ちがうちがう で 木 木のところを左に曲がったよね*んで
90F : *うんうん
90F : 木のところをがーっていくと富士山につくやんか
91G : そうそう で 富士山につきあたる前に*下のほうにぐーって行って*で
91G : お地蔵さんの下通って
92F : *うん *うん
92F : お地蔵さんの下通ればいいの?*
93G : *うん
うん お地蔵さんの下通って*で お地蔵さんの左をちょっと上にいって*
94F : *うん *うん
95G : おわり もうフィニッシュ?

4 問題点

予備実験および転記テキスト作成作業全体を通して、実質的かつ技術的な問題がさまざまに生じてきた。特に重要で改善が必要とされる問題点は以下の3つである。

1. 音声情報の欠落および不正確さ

一台の小型テープレコーダによる録音という簡便な対話収録方法のため、対話者の発話内容それ自体の収録は可能であったが、より厳密かつ微妙な記述が求められる音声情報や周辺言語情報のほとんどは失われてしまった。同様に音声パワー情報不足のため、千葉大コーパスのような発話単位 (Utterance Unit)⁸ の導入は不可能であり、客観的な記述のための単位としては話者交代を利用するのみであった。とりわけ、オーバーラップ (発話の時間的重複) の時間やポーズ (無音区間) の時間を正確に記録することは不可能であった。

2. 収録条件の不統一性

予備実験では、地図上の目標物のイラストを日本語環境に合うように一部描き変えて使用した。このため、認識しやすさという点ではむしろ好ましいともいえるが、こうした配慮は、言うまでもなく、エジンバラコーパスとの相互利用を妨げることになる。また、性別、視認可能条件、親近性条件などの実験デザインは厳密に適用されておらず、エジンバラコーパスとの統一性を欠いている。

3. 転記テキスト記述上の不正確さ

1で指摘されたように、オーバーラップの時間が不正確であるだけでなく、オーバーラップ中の発話内容も聞き取りにくさから、文字化する際に不正確さを免れなかった。転記テキストに「音声のもつ音韻情報を正確に、かつ全てを反映させることは非常に困難である」（小磯他：1995）ことは容易に想像がつくが、やはり必要最低限の厳密な音声情報の記述は欠かせない。直接的に音声的現象とはかかわりがないと思われる言語現象でさえ、対話理解のための重要な手がかりとなりえることがあるからである。

この他にも議論すべき問題点は多くあり、改善が求められていることは言うまでもない。この3点の問題点をふまえて、より信頼できうる対話収録とコーパス作成のために改善すべき点は以下のようにまとめられるだろう。

1. 音声情報の最低限の正確さを保つために、2チャンネル方式の音声収録設備を利用して、2チャンネル音声データを確保することが第一である。
2. 地図上の目標物のイラストはエジンバラコーパス（名称なし）と同一のものを使用する。また、性別、視認可能条件、親近性条件などの実験デザインはエジンバラコーパスとの相互利用を考慮して、これと統一させる。
3. 転記テキストの記述は、「読みやすく、かつ直観的に対話の把握が可能な形式」「対話現象を包括的に表現可能な形式」⁹を優先する一方で、必要最低限の音声情報を文字化させることに努める。部分的に必要と考えられる言語現象に関しては音声的情報（たとえば、指示詞の強勢をタグづけするなど）を用いることも検討する。¹⁰ エジンバラコーパス（名前なし）との相互利用のため、ローマ字表記を一部使用することも検討する。さらに当面の研究目的に応じた転記テキスト記述のための基準を整備することを目指す。

5 おわりに

本稿では日英語対照研究のための小規模対話コーパスの作成を目指して、その予備実験の概要と転記テキストを掲載した。到達目標と対話のコンテキストが明らかな小規模コーパスの利点は、その扱いやすさにある。実際、既存の大規模コーパスを言語資料として対話の分析をするためには経験と技術の蓄積が必要とされる。対話をとりまくコンテキストを特定する難しさに加えて、分析の方法も十分確立されていないことや対話というジャンル独特の用法を既存の言語的アプローチで説明できないという問題に常に直面するからである。対話という複雑なコミュニケーション手段を正確に記述することは非常に困難な作業であるが、より自発的で自然な対話コーパスなしにはありのままの対話を記述することさえままならないだろう。したがって、信頼できる対話データに基づき、既存の指示システムを検証することを今後の目標とし、談話における指示表現のメカニズムを説明できる柔軟な談話モデルを提示する方向性を探りたい。

註

* 日本語名称なしコーパスの構想は、エジンバラ大学 HCRC で収録、転記されたエジンバラコーパス

の未公開部分の一部利用から得られた。コーパスのデザインに関しては、エディンバラ大学理論応用言語学科の Dr. Robin Lickley と Prof. Jim Miller に貴重な情報提供と助言をいただいた。日本語名称なしコーパス作成の準備には土屋俊氏、近藤真理子氏、および岩敦子氏に多くの実際的助言をいただいている。ここに謝意と共に銘記する。最後になるが、予備実験に参加してくれた三重大学人文学部の学生および対話の収録と転記作業を手伝ってくれた 1997 年度人文学部授業「アメリカの言語」受講生にはこの場を借りて深く感謝したい。なお、本研究の一部は 1997 年度三重大学教育研究学内特別経費の助成を受けた。

- 1 コーパスの詳細については Anderson et al. (1991) を参照。また、HCRC のホームページ上でも公開されている。http://hcrc.ed.ac.uk/dialogue/maptask.html
- 2 千葉大コーパスの対話収録と転記作業は 1994 年から 1997 年にかけておこなわれた。コーパスの詳細については青野ほか (1994) 堀内ほか (1997) を参照。
- 3 対話中に仕切り越しに相手の顔を見ることができる。
- 4 対話相手に対して親近性があるか否か。親近性のある関係とはその発話の参加者が互によく知り合っているという関係であり、親近性がない関係とはそれまでに面識がない関係である。(堀内ほか：1994)
- 5 談話研究のための音声言語転記の方法は Du Bois et al. (1993) が参考になる。
- 6 音声情報の不足のため、今回は「重複」と「準重複」を区別せず、単に「重複」として扱った。(堀内ほか：1994)
- 7 Du Bois et al. (1993) を参照。
- 8 発話単位は「自己発話内の 400[ms]以上の無音区間によって区切られた音声的連続」と定義される。(堀内ほか：1997)
- 9 小磯ほか (1995) 参照。
- 10 土屋俊氏との私信による。

参考文献

- Anderson, A. H., M. Bader, E. G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtdo, J. McAllister, J. Miller, C. Sotillo, H. Thompson, and R. Weinert (1991) 'The HCRC Map Task Corpus', *Language and Speech*, 34 (4): 351 - 366.
- 青野元子, 市川熹, 小磯花絵, 佐藤伸二, 仲真紀子, 土屋俊, 八木健司, 渡部直也, 石崎雅人, 岡田美智男, 鈴木浩之, 中野有紀子, 野中慶子 (1994) 「地図課題コーパス (中間報告)」『情報処理学会音声言語情報処理研究会資料』94 - SLP - 3 - 5, 25 - 30.
- Carletta, J., R. Caley, and S. Isard (1993) 'A collection of self-repairs from the Map Task Corpus', *Research Paper HCRC/TR-47* Edinburgh: University of Edinburgh (HCRC Publications)
- Du Bois, J. W., S. Schuetze-Coburn, S. Cumming, D. Paolino (1993) 'Outline of Discourse Transcription', In Edwards, J. A. and M. D. Lampert, *Talking data : transcription and coding in discourse research*. Hillsdale, New Jersey : Lawrence Erlbaum Associates
- Deverell, J. (1994) 'Referring expressions in the Map Task Corpus' *unpublished paper* (Centre for Cognitive Science, U of Edinburgh)
- 堀内靖雄, 吉野文, 仲真紀子, 土屋俊, 市川熹 (1997) 「千葉大学地図課題対話コーパスプロジェクト」『千葉大学工学部研究報告』Vol 48. No 2. 33 - 60.
- 小磯花絵, 堀内靖雄, 佐々木聡, 仲真紀子, 土屋俊, 市川熹, 石崎雅人, 岡田美智男, 鈴木浩之, 中野有紀子 (1995) 「千葉大学課題コーパス作成・利用環境について - 1995 年度活動報告 -」『人工知能学会研究会資料』SIG-SLUD - 9503 - 5.
- Miller, J. and R. Weinert (1998) *Spontaneous Spoken Language*. Oxford: Clarendon Press