TR-IT-0212

ジェスチャをともなう発話とともなわない発話の比較 Comparison between Utterances with Gestures and without Gestures 水梨豪 Suguru MIZUNASHI 1997.3

概要

ポインティング・ジェスチャをともなう発話とともなわない発話の差異を把握するために、マルチモーダル対話環境におけるジェスチャをともなう発話と電話対話における発話を、文の長さ、発話時間、ポーズ時間、発話のスピード、使用品詞、言い間違いなどに関して比較した。その結果、以下の結論を得た。

- 1) 1ターンの発話時間、発話のスピード、1ターン中のポーズ数、ひとつのポーズの長さ、発話時間に対してポーズ時間が占める割合、1ターン中の単語数と文字数、1ターン中の Communicative Act単位の数、構文の曖昧性、言い間違いの頻度・性質に関しては、差異が認められない。
- 2) 代名詞と連体詞の使用については、ジェスチャをともなう発話のほうが頻繁になる。そして、 そこでは、近接指示をあらわす「こ」がつく代名詞、連体詞が多用される。
- 3) 言い淀み (間投詞) は、ジェスチャをともなう発話よりも、ともなわない発話によく出現する。

ATR 音声翻訳通信研究所 ATR Interpreting Telecommunications Research Laboratories

(c) (株) ATR 音声翻訳通信研究所 1997

(c) 1997 by ATR Interpreting Telecommunications Research Laboratories

1 背景と目的

- 2 準備
 - 2.1 ジェスチャをともなう発話のCommunicative Actによる分類
 - 2.2 比較対象の決定
 - 2.3 調査した項目
- 3 比較
 - 3.1 発話時間
 - 3.2 発話のスピード
 - 3.2.1 ジェスチャあり発話とジェスチャなし発話
 - 3.2.2 ジェスチャあり発話内での差異
 - 3.3 ポーズ
 - 3.3.1 ポーズの数
 - 3.3.2 ポーズひとつあたりの時間
 - 3.3.3 発話時間に対してポーズ時間が占める割合
 - 3.4 ターンの長さ
 - 3.4.1 1ターン中の単語数、文字数
 - 3.4.2 1ターン中のCommunicative Act単位の数
 - 3.5 品詞の使用頻度
 - 3.6 構文の曖昧性
 - 3.7 言い淀み (間投詞)、言い間違い
 - 3.7.1 言い淀み
 - 3.7.2 言い間違い
- 4 結論

参考文献

付録A

付録B

付録C

付録D

ジェスチャをともなう発話とともなわない発話の比較

1997年3月 水梨 豪 ATR音声翻訳通信研究所 第4研究室

1 背景と目的

ATR音声翻訳通信研究所第4研究室では、マルチモーダル翻訳通信に関する研究を行っている。具体的には、マルチモーダル通信環境における対話の特性を調査するため、マルチモーダル翻訳通信環境シミュレータEMMI (ATR Environment for Multimodal Interaction) [1] を作成し、そのもとで異言語間の翻訳通信実験を行い、採取した実験データを分析している[2,3,4]。今回は、データ分析の一環として行った、ジェスチャをともなう発話とともなわない発話の比較について報告する。

2 準備

2.1 ジェスチャをともなう発話 のCommunicative Act による分類

ジェスチャをともなう発話とともなわない発話を比較するにあたって、まず、比較対象を決定しなければならない。表層文が同じ発話Aと発話Bがあり、Aにはジェスチャがともなっていて、Bにはともなっていないという例があれば、その両者を比較することが理想的であるが、実際に採取した発話データには、そのような例は皆無であった。そこで、表層は異なっていてもある程度意味が近いと考えられるジェスチャあり/なし発話同士を比較することにする。そのためには、まず、ジェスチャをともなう発話の意味を捉える必要がある。

ここでは、「発話の意味」を簡略化して発話意図とその命題内容と考え、まず、ジェスチャをともなう発話として設定する、WOZ実験[4]のマルチモーダル対話におけるAGENT(日本人)発話のうちのジェスチャをともなう発話をCommunicative Acts[5]を用いて分類した。表1は、それぞれのCommunicative Act (CA)にどのようなジェスチャが使用されているかを集計したものである。実際に使用されたジェスチャは、サークリング40回、ドラッギング34回、その他1回であったが、複数のCA単位で用いられているジェスチャは重複させて集計した。表中のNoneは、そのCAにはジェスチャが用いられなかったことを表わす。

表1より、ジェスチャをともなう発話の特徴として、以下の項目が挙げられる。

- ・Inform / BE-LOCATED 型の発話において、サークリング、ドラッギングがともに多用されている。 Inform / BE-LOCATED 型とは、具体的には、「A(もの、ひと)がB(場所)にあるということを聞き手に知らせる」という形である。
- · Action request型の発話において、ドラッギングが多用されている。Action request型とは、具体的には、 「ある動作(出る、行く、…)を聞き手が行うことを要求する」という形である。

表1 各Communicative Actで使用されたジェスチャ

	\$4. 3-7	頻度			
Communicative Act	動詞	Circling	Dragging	Other	None
Inform	BE-LOCATED	23	20	0	1
	COPULA	. 1	0	0	0
	降りる	0	0	0	1
	着く	1	0	0	1
	示す	0	1	0	0
	引く	0	1	0	0
	思う	0	0	0	1
	不明	1	1	0	0
Confirmation question	BE-LOCATED	2	0	0	. 0
	COPULA	1	0	0	0
YN question	BE-LOCATED	2	0	0	0
Suggest	いかがだ	2	0	. 0	0
Action request	出る	0	8	0	1
	行く	3	8	0	0
	乗る	0	2	1	1
	降りる	2	2	. 0	0
	乗り換える	0	4	0	0
	言う	2	0	0	0
	渡る	0	1	0	1
	通り過ぎる	0	1	0	0
	通る	0	2	0	0
	上がる	0	1	0	0
	曲がる	0	1	0	0
1	歩く	0	1	0	0
	進む	0	1	0	0
Acknowledgement		0	0	0	4
Yes		. 0	0	0	2

2.2 比較対象の決定

2.1の結果をふまえて、以降では、Inform/BE-LOCATED型、Action request型に特定して、WOZ実験のマルチモーダル対話におけるAGENT(日本人)発話のうちのジェスチャをともなう発話と、WOZ実験の電話対話におけるAGENT(日本人)発話の間で比較を行うこととする。つまり、上記ふたつの型のどちらかの発話が含まれているターンを、WOZ実験のマルチモーダル対話におけるAGENT発話とWOZ実験の電話対話におけるAGENT発話から抽出し、両者をトータルに比較することになる。

ふたつの型のどちらかの発話が含まれているターンを抽出した結果、WOZ実験のマルチモーダル対話におけるAGENT発話のうち40ターン(以後、ジェスチャあり発話と呼ぶ)、WOZ実験の電話対話におけるAGENT発話が83ターン(以後、ジェスチャなし発話と呼ぶ)が得られた。それぞれを、付録A「WOZ/MM/AGENT/ジェスチャつき発話一覧」と付録B「WOZ/TEL/AGENT/INFORM,ACTION-R発話一覧」に掲げる。

2.3 調査した項目

ジェスチャあり発話、ジェスチャなし発話のそれぞれのターンに対して、以下の項目を調査した。

a) 発話の時間

500msec以上の無音状態をポーズとし、ポーズで区切られた発話の開始・終了時刻を記録した。付録C 「発話とジェスチャの時刻(WOZ/MM/AGENT/ジェスチャつき発話)」と付録D「発話の時刻 (WOZ/TEL/AGENT発話)」を参照。

b) ジェスチャの時間 (ジェスチャあり発話のみ)

ジェスチャの開始・終了時刻とジェスチャの種類を記録した。「発話とジェスチャの時刻 (WOZ/MM/AGENT/ジェスチャつき発話)」では、ジェスチャが行われた時刻は各ターンごとに発話の時刻と対応させて記録してある。

c) 単語数、文字数

形態素解析結果から、句読点をのぞいた単語数と文字数を集計した。

d) Communicative Act単位の数

1ターン中の発話意図の数を知るため、Communicative Act単位の数を集計した。

e) 品詞の使用頻度

形態素解析結果から、品詞の使用頻度を集計した。品詞体系はATR基準[6,7]とした。

f) 構文の曖昧性

構文解析結果の構文木の数を調査した。解析にはボトムアップ探索を行うチャートパーザ[8]、229規 則からなる文脈自由文法、約2000語の辞書を用いた。

g) 言い淀み(間投詞)、言い間違い

各ターンに含まれる言い淀み(間投詞)と言い間違いを集計した。

3 比較

3.1 発話時間

2.3のa)の結果から、1ターンあたりの発話時間(ポーズ時間を含む)を集計した。平均と分散を表2に示す。

	3/2 17 V 3/7C 9 V 7/3C HILLY 141		
	ジェスチャあり発話	ジェスチャなし発話	
ターン数	40	83	
平均	8.63sec	7.73sec	
分散	46.3	22.6	

表2 1ターンあたりの発話時間

ジェスチャあり発話40ターンの発話時間と、ジェスチャなし発話83ターンの発話時間を2つの標本として、F検定とt検定を有意水準5%で行った結果、分散に有意差があり、平均には有意差がないことがわかった。なお、以降の検定もすべて同様に行った。

3.2 発話のスピード

3.2.1 ジェスチャあり発話とジェスチャなし発話の差異

2.3のa)とc)の結果から、単位時間(sec)に発話された単語数と文字数を発話スピードとして集計した。 ここでは、発話のスピードを「発話された単語数あるいは文字数/ポーズ時間を除いた発話時間」で定 義する。両者の平均と分散を表3、4に示す。

ジェスチャあり発話ジェスチャなし発話ターン数4083平均3.37単語/sec3.40単語/sec分散0.790.92

表3 発話スピード(単語数)

+ .	マッニナ ー	2.0	1.5	/ -l / : 161.\
₹4	金計ス	_	Ь.	(文字数)

	ジェスチャあり発話	ジェスチャなし発話
ターン数	40	83
平均	6.60文字/sec	6.83文字/sec
分散	1.59	1.67

検定の結果、単語数、文字数両者の発話スピードにおいては、分散にも平均にも有意差はないことが わかった。

3.2.2 ジェスチャあり発話内での差異

2.3のa)とb)の結果から、ジェスチャあり発話をジェスチャをともなう発話部分とともなわない発話部分に細分し、両者の発話スピードを調査した。その結果、ジェスチャをともなう発話部分が63、ともなわない発話部分が77抽出された。両者の平均と分散を表5、6に示す。

表5 ジェスチャあり発話内での発話スピードの差(単語数)

	ジェスチャあり	ジェスチャなし
標本数	63	77
平均	4.00単語/sec	3.59単語/sec
分散	4.77	9.12

表6 ジェスチャあり発話内での発話スピードの差(文字数)

	ジェスチャあり	ジェスチャなし
標本数	63	77
平均	7.39文字/sec	7.28文字/sec
分散	13.3	19.3

検定の結果、単語数、文字数の発話スピードに関しては、分散に有意差があり、平均には有意差はないことがわかった。

3.3 ポーズ

3.3.1 ポーズの数

2.3のa)を用いて、各ターンに含まれるポーズの数を集計した。ジェスチャあり発話、ジェスチャなし 発話のそれぞれの平均と分散を表7に示す。

 ジェスチャあり発話
 ジェスチャなし発話

 ターン数
 40
 83

 平均
 1.80個
 1.53個

 分散
 5.09
 2.13

表7 1ターン中のポーズの数

検定の結果、分散に有意差があり、平均には有意差がないことがわかった。

3.3.2 ポーズひとつあたりの時間

2.3のa)を用いて、ポーズひとつあたりの時間を集計した。ポーズはジェスチャあり発話には82、ジェスチャなし発話には147あった。表8にそれぞれの平均と分散を示す。

	ジェスチャあり発話	ジェスチャなし発話
ポーズ数	82	147
平均	1.59秒	1.22秒
分散	2.67	1.27

表8 ポーズひとつあたりの時間

検定の結果、分散に有意差があり、平均には有意差がないことがわかった。今回は、ポーズを500msec以上の無音状態と定義したので、それ以下の無音状態はすべてポーズ時間ゼロとして扱われている。したがって、ポーズ認定の基準をより厳しくして再検討する必要があるだろう。

3.3.3 発話時間に対してポーズ時間が占める割合

2.3のa)を用いて、発話時間に対してポーズ時間が占める割合、すなわち、「ポーズ時間/ポーズを含む発話時間」をターンごとに集計した。ジェスチャあり発話、ジェスチャなし発話のそれぞれの平均と分散を表9に示す。

表9 発話時間に対してポーズ時間が占める割合

	ジェスチャあり発話	ジェスチャなし発話
ターン数	40	83
平均	0.26	0.22
分散	0.05	0.03

検定の結果、分散に有意差があり、平均には有意差がないことがわかった。この件に関してもポーズ 認定の基準を厳しくして再検討する必要がある。

3.4 ターンの長さ

3.4.1 1ターン中の単語数、文字数

2.3のc)を用いて、1ターン中の単語数、文字数を集計した。表10、11に平均と分散を示す。

表10 1ターン中の単語数

	ジェスチャあり発話	ジェスチャなし発話
ターン数	40	83
平均	17.1単語	17.4単語
分散	86.7	51.2

表11 1ターン中の文字数

	ジェスチャあり発話	ジェスチャなし発話
ターン数	40	83
平均	34.4文字	36.6文字
分散	368.4	292.4

検定の結果、単語数に関しては分散に有意差があり、平均に有意差はなく、文字数に関しては分散に も平均にも有意差がないことがわかった。

3.4.2 1ターン中の Communicative Act 単位の数2.3のd)で得たデータを集計すると、表12のようになった。

表12 1ターン中のCommunicative Act単位の数

	ジェスチャあ	ジェスチャな
ターン数	40	83
平均	1.85個	2.13個
分散	1.00	0.77

検定の結果、分散にも平均にも有意差がないことがわかった。

3.5 品詞の使用頻度

2.3のe)で集計した品詞の使用頻度を表13に示す。

表13 品詞の使用頻度

	<u> </u>	TEL (83ターン)
固有名詞	33	-86
サ変名詞	1	7
形容名詞	0	2
普通名詞	102	227
数詞	7	26
代名詞	28	19
本動詞	71	159
補助動詞	1	5
語尾	127	282
形容詞	3	6
副詞	23	40
連体詞	8	3
接続詞	12	21
感動詞	5	22
間投詞	11	51
助動詞	83	186
格助詞	93	189
準体助詞	0	2
係助詞	11	13
副助詞	1.	5
並立助詞	. 0	10
接続助詞	26	57
終助詞	1	. 13
連体助詞	25	63
引用助詞	3	12
接頭辞	1	4
接尾辞	17	46
その他	2	4

各品詞の使用頻度は、ジェスチャあり発話のターン数 (40) とジェスチャなし発話のターン数 (83)

の比率に準じているが、代名詞と連体詞に関してはジェスチャあり発話の方が多くなっている。使用されている代名詞と連体詞を具体的にみると、表14のようになっている。ジェスチャあり発話では、「ここ」や「こちら」のような直示表現が多く使用されているが、ジェスチャなし発話ではそれらは全く用いられていないことがわかる。

表14 使用された代名詞と連体詞

	ジェスチャあり発話	ジェスチャなし発話
代名詞	ここ(18)、こちら(5)、 そこ(2)、あなた(2)、 わたし(1)	そこ(18)
連体詞	この(8)	その(3)

3.6 構文の曖昧性

2.3のf)で得られた構文の曖昧性を集計した。ジェスチャあり発話40ターン、ジェスチャなし発話83ターンのうち、それぞれ11ターン、24ターンが解析不能となった。解析不能となったのは、主に、名詞、数詞の連接がある場合、連体詞句と普通名詞の間に副詞が挿入されている場合(ホテルの<u>すぐ</u>近くに…)、文中に間投詞がある場合などであった。

ここでは、解析不能のターンを除いた、ジェスチャあり発話29ターン、ジェスチャなし発話59ターン について平均・分散の検定を行う。表15に構文の曖昧性(解析構文木の数)平均と分散を示す。

表15 構文の曖昧性

	ジェスチャあり発話	ジェスチャなし発話
ターン数	29	59
平均	3.21	3.71
分散	13.7	19.8

検定の結果、分散にも平均にも有意差はないことがわかった。

3.7 言い淀み (間投詞) 、言い間違い

2.3のg)の調査の結果を表16に示す。

表16 言い淀み(間投詞)、言い間違いの頻度

	ジェスチャあり発話	ジェスチャなし発話
ターン数	40	83
言い淀み	9	42
言い間違い	6	13

371 宣い淀み

観察された言い淀みを表17に具体的に示す。また、各ターンごとに言い淀みの頻度を集計して平均と 分散を求めたものを表18に示す。これを用いて検定をおこなったところ、分散にも平均にも有意差がみ とめられた。よって、ジェスチャをともなわない発話のほうがジェスチャをともなう発話よりも言い淀 みが生じやすいと言える。ジェスチャあり発話において発声された言い淀み9回のうち、7回の言い淀み についてはジェスチャが行われていない時に発声されているという事実と考え併せると、ジェスチャは 言い淀みを抑制する機能を持つと考えられる。

表17 言い淀み(間投詞)の種類

ジェスチャあり発話	ジェスチャなし発話
え(6)、あ(2)、えと(1)	え (29) 、あ (4) 、えと (3) 、 お (2) 、あっと (1) 、う (1) 、 はあ (1) 、ふん (1)

表18 言い淀み回数の平均と分散

	ジェスチャあり発話	ジェスチャなし発話
ターン数	40	83
平均	0.23回	0.51回
分散	0.28	1.01

3.7.2 言い間違い

各ターンごとに言い間違いの頻度を集計して平均と分散を求めたものを表18に示す。

表19 言い間違い回数の平均と分散

		V = 20 10 1
	ジェスチャあり発話	ジェスチャなし発話
ターン数	40	83
平均	0.15回	0.16回
分散	0.13	0.16

検定の結果、分散にも平均にも有意差がないことがわかった。言い間違いの内容については、両発話とも「繰り返し」「言い直し」「挿入」「削除」などの例がひととおり出現しており、目だった差異はみられなかった。

4 結論

以上の分析結果をまとめると、ジェスチャをともなう発話とともなわない発話の差異は、以下のようになる。

- ・1ターンの発話時間、発話のスピード、1ターン中のポーズ数、ひとつのポーズの長さ、発話時間に対してポーズ時間が占める割合、1ターン中の単語数と文字数、1ターン中のCommunicative Act単位の数、構文の曖昧性、言い間違いの頻度・性質に関しては、差異が認められない。
- ・代名詞と連体詞の使用については、ジェスチャをともなう発話のほうが頻繁になる。そして、そこでは、近接指示をあらわす「こ」がつく代名詞、連体詞が多用される。
- ・言い淀み(間投詞)は、ジェスチャをともなう発話よりも、ともなわない発話によく出現する。

参考文献

- [1] Kyung-ho Loken-Kim, Fumihiro Yato, Kazuhiko Kurihara, Laurel Fais, Ryo Furukawa, EMMI ATR Environment for Multi-Modal Interactions, ATR Technical Report TR-IT-0018 (1993)
- [2] 谷戸 文広, キュンホ ローケンキム, ローレル ファイス, 森元 逞, 道案内タスクにおけるマルチモーダル対話の会話文の特徴分析, 電子情報通信学会論文誌Vol.J77-D-「No.8(1994)
- [3] Young-Duk Park, Kyung-Ho Loken-Kim, Laurel Fais, Suguru Mizunashi, Analysis of Gesture Behaviour in a Multimedia/multimodal Interpreting Experiment; Human vs Wizard of Oz Interpretation Method, ATR Technical Report TR-IT-0091 (1995)
- [4] 水梨 豪, ローケン-キム キュンホ, パク ヨンドク, 友清 睦子, 森元 逞, マルチモーダル翻訳対話における発話とジェスチャの関連性, 情報処理学会研究報告95-SLP-6(1995)
- [5] 友清 睦子, 対話行為ラベルとその自動付与, ATR Technical Report TR-IT-0069 (1994)
- [6] 浦谷 則好, 田代 敏久, 山田 久子, 松本 香, 音声言語データベースにおける日本語形態素解析マニュアル, ATR Technical Report TR-IT-0009 (1993)
- [7] 浦谷 則好, 田代 敏久, 森田 千帆, 音声言語データベースにおける日本語形態素解析マニュアルの補遺, ATR Technical Report TR-IT-0053 (1994)
- [8] 田代 敏久, 森元 逞, 音声言語処理のための構文解析ツールキット, 情報処理学会研究報告 95-NL-106(1995)

付録B WOZ/TEL/AGENT/INFORM,ACTION-R発話一覧

WOZ実験の電話対話におけるAGENT(日本人)発話のうち、Inform / BE-LOCATED 型、Action request型を持つものを被験者ごとに集計した。発話中の/は500msec以上のポーズを表わす。

******AC.TEL.WOZ.TR

acl A:左のほうに、/曲がっていただきまして。

ac2 A:はい、橋を渡っていただきまして。

ac3 A:橋を渡っていただきまして。

ac4 A: [あ] 電車では行けませんので、/バスか、タクシーを、使ってください。

ac5 A:電車では、行けませんので、/バスか、タクシーを、使ってください。

ac6 A:六番のバスに乗ってください。

ac7 A:六番、のバスに、乗ってください。

ac8 A:烏丸口を、通ってください。

******AN.TEL.WOZ.TR

anl A:はい、/近くに、都ホテルがあります。/そこでいいですか。

an2 A:はい、/京都ロイヤルホテルがあります。

an3 A:鳥丸口を出てください。/目の前にバス乗り場があります。/そこで三番のバスに、乗ってください。

an4 A:鳥丸口から出てください。/目の前にバス乗り場があります。/三番のバスに乗ってください。

an5 A:京都駅、烏丸口から、/出てください。/目の前に、バス乗り場があります。/そこで、三番のバスに乗ってください。

an6 A:京都駅の、/北にある、/鳥丸口、から出てください。

dp1 A:はい、/ [え] それでは、タクシー乗り場で、/タクシーをひらって。

dp2 A:駅の前の、タクシー乗り場で、タクシーに乗ってください。/ [え] 行き先<yukisaki>は、/国際交流センターです。

dp3 A:駅の前の、タクシー乗り場で、/タクシーを、/ひらって、/ [え] 行き先<yukisaki>は、国際交流センターと、おっしゃってください。

dp4 A:はい、[えー]、/近くのホテルでは、都ホテル、というのがありますが。

dp5 A: [え] 交流センターからは歩いて十分ぐらいです。/すぐ近くです。/ [え] 都ホテルは、/ [え] シティーホテルです。

dp6 A: [え] 都ホテルは、 [え] 一流ホテルです。/ [え] そして、交流センターのすぐ近くにあります。

erl A:分かりました。/ [え] それでは、 [えー] タクシーの運転手に、 [えー] 国際交流センターへ、行ってくだ

さい、と、 [え一] おっしゃってください。/料金は、/だいたい、三千円です。

er2 A:はい、/それでは、/タクシーの、運転手に、/国際交流センターへ、行ってくださいと、言ってください。/料金は、だいたい、/三千円です。

er3 A:国際交流センターへと、/タクシーの、運転手に、伝えてください。

er4 A:もう少し、安い、ホテルが、/少し、離れたところに、あります。/京都、ロイヤルホテル、/一泊、八千円です。

**************************LF.TEL.WOZ.TR

lfl A: [え] それではですね、/新幹線(を)、のコンコースを、/ [え] 出てください。/(あ、ちゃうな。)

lf2 A:はい、 [えと] 新幹線のコンコースの、 [え] / 西のほうへ出てください。

lf3 A:そしてそこを出たら、/[えーと] (きたて、きた) 北出口へ真っ直ぐ進んでください。

lf4 A:コンコースを出たら、北出口へ、真っ直ぐ直進してください。

lf5 A:北出口を出て、そのまま直進していくと、バス乗り場があります。/そしてそこを右へ、(いっ)行ってください。

lf6 A:北出口を出て、真っ直ぐ行くと、バス乗り場があります。

lf7 A: そしてそこを右へ(曲がっ)真っ直ぐ行ってください。

lf8 A: そしてそこを右へ真っ直ぐ行ってください。

lf9 A:バス乗り場が見えたら右へ、曲がってください。/(あ、ちゃうか。)

If10 A: {I;そ;I} /すると [あ] 左手にタクシー乗り場がありますから。

lf11 A:少し行くと、左手にタクシー乗り場がありますから。

If12 A: [あっと] でそのタクシー乗り場、 [あ] (た) そこのタクシーに乗って。

If13 A: (そこのタクシーに乗って、/国際会議事務、あ、ちゃうわ、そこ) そこのタクシーに乗って。

lf14 A: [あっ] そこのタクシーに乗って、国際交流センター、まで行ってください。

If15 A: それでは、国際交流センターのすぐ近くに、/都ホテルというところがありますから、そこでよろしいでしょうか。

If16 A:国際交流センターの近くに、/都ホテルというところがありますが、そこでよろしいでしょうか。

************************NO.TEL.WOZ.TR

nol A: それでは、/中央出口を出てもらって、/左側に、バス乗り場があります。

no2 A:中央出口を、出てもらって、左側に、行きますと、バス乗り場があります。

no3 A:はい、/そこから、/五番のバスに乗って、/三条京阪の駅まで行ってください。

no4 A:はい、/そこから五番のバスに乗って、三条京阪まで行ってください。

no5 A:はい、/[う]三条京阪から、京津<keishin>線の、(浜お)浜大津行き<iki>の電車に乗ってもらって、二つ目の駅、/蹴上<keage>駅で、降りてください。

no6 A:三条京阪から、京津<keishin>線、浜大津行き<iki>に乗ってもらって、蹴上<keage>駅で、降りてください。

no7 A:はい、/三条京阪から、/浜大津行き<iki>の電車に乗ってもらって、蹴上<keage>駅で、降りてください。

no8 A:京阪の、蹴上<keage>駅から、/歩いて、一分のところ、すぐにあります。

rdl A:タクシーでしたら、/駅の、/北側と、/南側と、両方に乗り場があります。

rd2 A: タクシーに乗ったら、/ (蹴上まで行って国際交流センターと言ってください。)/蹴上の、/国際交流センターまで、と言ってください。

rd3 A:タクシーに乗って、/蹴上の、/国際交流センターと言ってください。

rgl A:ではそこから国際交流センターまで乗って来てください。/時間は、およそ十分ぐらいです。

rg2 A:ではその場所から、/タクシーで、国際交流センターまで来てください。/時間はおよそ十分ぐらいです。

rg3 A:はい、一番近くに、都ホテルがあります。

m1 A: [えーと] すぐ正面の、北口、の階段を上がっていただい {II; て; II} 。

rp2 A: [えー] 京都北口の改札を出ていただいて、/タクシー乗り場か、バス乗り場に行ってください。/バスならば、/ [え] 一番のバスに乗ってください。/ [え] タクシーならば、直接、/ [え] 国際交流センターと、指示してください。

rp4 A:階段を上がって、/正面の出口 |I;に、まず出てください;I |。

rp5 A: (|I;北口の階段を;I| まず上がってください。) / [え] 北口の階段を、まず上がって、ください。/正面に出ますので。

rp6 A: [え] タクシーか、 |I1;バス乗り場か;I1 | があります。

rp7 A:タクシーか、バス乗り場があります。/よろしい {I2;か;I2}。

rp8 A:タクシーか、/バス乗り場があります。/分かりましたか。

rp9 A:京阪の、三条駅で、/(お)まず、降りてください。

rp10 A:京阪駅から、/京阪、京津<keishin>線に、 |I;乗り換えてください;I|。

rp11 A: [| I2,おー;I2 | 、/え] 京阪、三条駅から、/京阪、/京津<keishin>線、/ [え] ケー、/イー、アイ、エスエイチアイエヌ、ライン、/京阪京津<keishin>線、/に、乗り換えてください。

rp12 A:京阪、三条駅から、/京阪、京津<keishin>線に、乗り換えてください。/そして、蹴上駅で、/降りてください。/そうすると、目の前に、国際、交流センターがあります。/分かりましたか。

rp13*A:はい、京阪三条 {I1;駅;I1} から。/

I: {A1:はい:A1} 、お願いします。

A:京阪 | I2;京津<keishin>線に、乗り換えてください;I2 | 。

rp14 A: [おー、え] 京阪三条駅から、/ [え、え] 蹴上の駅まで、/すぐですから、/そこからもうタクシーに乗ってください。/わずか五、六分で到着します。/分かりましたか。

rp15 A: [え] 北口を、出たところに、すぐにあります。

rp16 A: そうです、北口を、/出て、すぐのところに、/乗り場が、あります。

rp17 A: { [はあ] } / [ふん] / はい、/すぐ近くに、/都ホテルが、あります。/どうでしょう。

m18 A:すぐ、近くに、/都、ホテルが、ありますが、どうでしょうか。

rp19 A:もう一つ、/京都ホテルがあります。/一万円です。

**********************************SS.TEL.WOZ.TR

ssl A:はい、/ (バスで) バスでしたらば、/三番乗り場、/から乗っていただきまして、/京阪の、三条駅で降りてください。

ss2 A:はい、/三番乗り場に乗っていただきまして、/京阪の三条駅で降りてください。

ss3 A:三番乗り場から乗っていただきまして、/京阪の、三条駅で降りてください。

ss4 A:はい、京阪三条駅に着きましたら、/電車、京津<keishin>線に乗ってください。

ss5 A:はい、京津<keishin>線の、蹴上駅で降りてください。

ss6 A:はい、蹴上駅の改札を出られましたら、左に、歩きます。

ss7 A:蹴上駅の、/駅を出られまして、まず、左に歩いてください。

ss8 A:二十メートルぐらい、歩きますと、右側に、ガソリンスタンドがあります。/そこを、右に曲がってください。

ss9 A:いえ、違います。/ガソリンスタンドを、右に曲がります。

ss10 A:はい、/そのまま、歩いていただきますと、右手側に国際交流センターと、ありますので、/そのまま歩いてください。

ss11 A:はい、/一番近いホテルは、都ホテルになりますが。

付録A WOZ/MM/AGENT/ジェスチャつき発話一覧

WOZ実験のマルチモーダル対話におけるAGENT(日本人)発話のうちのジェスチャをともなうターンを、被験者ごとに集計した。発話中の<>D,<>Cなどは、<>で囲まれた発話の部分がドラッギング・ジェスチャ、サークリング・ジェスチャにともなわれていることを表わす。また、発話中の/は、500msec以上のポーズを表わす。なお、各ターンID(ac2、an3など)の先頭に*がついたものは、Inform/BE-LOCATED型、 $Action\ request$ 型以外の型を持つもので、比較対象とはしていない。

**********AC.MM.WOZ.TR*******

*acl A:<手で [あの] 示しますので、線を引きまして>D。

ac2 A:はい、こちらのほうから、出て<いただきまして、/東地下通路を通っていただきまして。>D

ac3 A:<[え]このように出口を出られまして>D。

*ac4 A:<このように、/ひがし、つ>D

ac5 A:<東地下通路を真っ直ぐ上がっていただきまして、/左手に、/折れていただきましたところが、出口になります>D。

ac6 A: < {左手に曲がられまして}、/(東つか)東地下通路を真っ直ぐ、進んでいただきます>D。

ac7 <>C

A:ここになります。

ac8 <>C

A:こちらが<都ホテルになります>C。

anl A:<あなたは今ここにいます>C。

*an2 A:<あなたは今ここですね>C。

an3 A:分かりました。/lそれでは、<(出口を、)/八条口に出てください。/lタクシー乗り場は、>D<ここにあります>C。

an4 A:分かりました。|それでは、/<八条口に出てください。/|タクシー乗り場は>D、<ここにあります>C。

an5 A:<ここが国際交流センターです>C。

an6 A:タクシーの、運転手に、/南禅寺<の>C、/近くの、<国際交流センターに行って>Cください、と言ってください。

dp1 A:[えー]その前に、タクシー乗り場がありますから、/そこで、タクシーに乗ってください。<>D

*er1 A:<(京都駅を)/>C京都駅を/降りて、/[えー]今、この、丸のところに着きます。

er2 A:もし、タクシーを、お使いでしたらば、/国際交流センターへ、行ってください、と言えば、通じます。//タクシー乗り場は、/<ここです>C。<>D<>D

er3 A: <こちらが>C国際交流センターですが、/すぐ、<近くに、/都ホテル>Cがあります。

lfl A:(それでは、/[え]、今いる、ところ、を真っ直ぐ)/[あ]今いるところから、タクシー乗り場がすぐ見えると思いますから、/<ここです>C。

If2 A:すぐ前に、タクシー乗り場がありますから、/[え]<ここから、>S/タクシーに乗ってください。

************NO.MM.WOZ.TR********

*nol A:<このあたりですか>C。

no2 A:<バス乗り場に行かれたら、そのまま>D [えー]/五番系統のバスで、(す)三条京阪の、停留所まで行ってください。

*no3 A: (すー、すー) <三条京阪に>C 着いたら。

no4 A:はい、/三条京阪に着いたら、/京阪、/京津線に乗ってもらって、/二つ目の駅、/蹴上駅で、降りてください。<>D <>C

no6 A:[あ]それから、[えー]京津線の、浜大津駅行きの電車に乗ってもらって、/二つ目の、/<蹴上駅で、降りて>Cください。

no7 A: そこから、/<北へ>D、歩いて、五分ぐらいのところに、国際交流センターがあります。

*no8 A:[うーん]それでは、/<京都ロイヤルホテルはいかがでしょうか。>C

*no9 A: それでは、<京都ロイヤルホテルは>C、いかがでしょうか。

*rd1 <>C

A:ここですか。

rd2 <>C

A:すぐ前の、ここに、タクシー乗り場があります。

rd3 A:京都駅は、<>D/<こちらです>D。

rd4 A:<蹴上はここです>C。

rd5 <>C

A:国際交流センターはそのそばです。

rg1 A:今、あなたは、/<この場所に、いらっしゃいます>C。/<ここが出口になっていますので、>D/とりあえず、/この、 $\int R$ のチケット売場のほうまで行ってください。

rg2 A:はい、タクシーでしたら、/<近鉄線から>C<真っ直ぐ行っていただいて、JRのチケット売場がここに>D。

rg3 A:JRのチケット売場の<ところを、/通り過ぎて。>D

rg4 A:はい。/<近鉄の乗り場から>C、/<>Dわたしが今書いた線のとおりに歩いてください。ここにタクシー乗り場があります。<>D

rg5 <>C

A:この丸を書いたところがタクシー乗り場です。

rg6 A:この場所<がタクシー乗り場です>C。

rg7 A: (こくれ) 国際交流センター<は、/この丸を書いたところにあります>C。

*rg8 <>C

A:都ホテルですね。

rpl A:では、近鉄線から、/正面<の改札を>D出て、そして、/<橋を渡って>D、/<正面の、/出口、から、/真っ直ぐ、>D行ってください。/lそうすると、/<右手に>D、/タクシー乗り場、/あるいは、/<左にバス乗り場が、あります。>D/はろしいか。

rp2 A:まず、/<京阪>D、/三条駅まで、/<>Cバスで、/行ってください。

rp3 A:ここ<です。>C

rp4 A:そして、/< (京津線) >D、/京津線に、/乗り換え<て、/今度は、/ここ>D<まで、行ってください>C。/ようすると、/<目の前に>C、/あります。

rp5 A:ここで乗り換えて、/<そして、/-つ、/二つ目の、<math>>D駅で、降りて、ください。/!すぐ<前に>D、あります。

rp6 A:すぐ前に、/<>C都ホテルが、あります。

*ss1 A:こちらですね。<>C

ss2 A:今、<こちらにおられますので>D、/<ここから>D、/この一番端が、三番です。<>D

ss3 A:一番近いホテルは、/<>C 蹴上駅の、駅前の、都ホテルになりますが。

付録C 発話とジェスチャの時刻(WOZ/MM/AGENT/ジェスチャつき発話)

WOZ実験のマルチモーダル対話におけるAGENT発話それぞれに対して、ポーズ(500msec以上の無音状態とした)で区切られた発話の開始・終了時刻、ジェスチャの開始・終了時刻を記録した。発話IDの先頭に○がついていないものは、、Inform / BE-LOCATED 型、Action request型以外の発話で、比較対象とはしていない。また、表中の記号と数字は以下の意味である。

Start, End 発話またはジェスチャの開始・終了時刻(単位は0.01sec)

S, G 発話とジェスチャ

Tsp ポーズを含めた全発話時間(単位は0.01sec)

ΣTp 総ポーズ時間(単位は0.01sec)

Np ポーズの個数

なお、G(ジェスチャ)の欄の開始・終了時刻のあとに記されているD, C, Sは、それぞれそのジェスチャがドラッギング、サークリング、スクランブリングであることをあらわす。

Turn ID		Start	End		Start	End		Start	End	Start	End	Start	End	Start	End	T	sp	ΣΤρ	Np	ΣTp/Tsp	ΣTp/Np
ac1	S	0	284														284	0	0	0	0
	G	0	589	D																	
ac2O	S	0	365		1084	1371										1	371	719	1	0.52	719
	G	264	1459	D																	
ac3O	S	254	566														312	0	0	0	0
	G	0	1049	D																	
ac4	S	98	191		384	529											431	193	1	0.45	193
	G	0	548	D																	
ac5O	S	622	960		1174	1211		1291	1563								941	394	2	0.42	197
	G	0	1728	D												-					
ac6O	S	464	647		1028	1416											952	381	1	0.40	381
	G	0	1631	D																	
ac7O	S	186	296														110	0	0	0	0
	G	0	142	С																	
ac8O	S	666	859														193	0	0	0	0
	G	0	188	С	746	935	С														
an l 🔾	S	0	126														126	0	0	0	0
	G	0	195	С																-	

Turn ID		Start	End		Start	End		Start	End		Start	End		Start	End	Start	End	Tsp	ΣΤρ	Np	ΣTp/Tsp	ΣTp/Np
an2	S	0	190															190	0	0	0	0
	G	75	412	С							, , , , , , , , , , , , , , , , , , , ,											
an3〇	S	0	64		143	301		530	703		772	1026						1026	298	3	0.29	99
	G	253	939	D	939	1105	С															
an4O	S	0	220		412	569		682	919									919	305	2	0.33	153
	G	236	806	D	806	919	С															
an5O	S	0	217			, , , , , , , , , , , , , , , , , , , ,												217	0	0	0	0
	G	0	226	С																		
an6O	S	0	186		321	471		557	1026									1026	221	2	0.22	111
	G	393	552	С	689	866	С															
dp1	S	0	346		412	663												663	66	1	0.10	66
	G	678	771	D																		
er1	S	109	187		413	511		710	783		1162	1429						1320	804	3	0.61	268
	G	0	273	С									-									
er2O	S	0	293		423	1033		1139	1243		1488	1549						1549	481	3	0.31	160
	G	1324	1559	С	4074	4184	D矢	4403	4483	D矢												
er3〇	S	153	417		492	583		666	783									630	158	2	0.25	79
	G	0	181	С	509	692	С															

Turn ID		Start	End		Start	End		Start	End		Start	End	Start	End	Start	End		Tsp	ΣΤρ	Np	ΣTp/Tsp	ΣTp/Np
lf1O	S	0	71		238	511		730	1115		1255	1288	A. I.					1288	526	3	0.41	175
	G	1183	1289	С																		
lf2O	S	0	267		429	466		625	776								-	776	321	2	0.41	161
	G	312	617	S																		
nol	S	0	117	-														117	0	0	0	0
	G	0	122	С																		
no2O	S	96	438		552	1134				. /								1038	114	1	0.10	114
	G	0	442	D																		
no3	S	, 0	392															392	0	0	0	0
	G	197	298	С																		
no4O	S	0	66		165	272		346	393		476	633	722	793	885	1109		1109	437	5	0.39	87
	G	1219	1328	D	1328	1432	С															
no5O	S	. 0	312															312	0	0	0	0
	G	198	329	С	403	519	С															
no6O	S	0	638		698	749		828	1031									1031	139	2	0.13	70
	G	859	983	С																		
no7O	S	0	68		241	800												800	173	1	0.22	173
	G	213	386	D																		

Turn ID		Start	End		Start	End		Start	End	Start	End	Start	End	Start	End	Tsp	ΣΤρ	Np	ΣTp/Tsp	ΣTp/Np
no8	S	0	423		696	964										964	273	1	0.28	273
	G	793	932	С		****														
no9	S	0	333													333	0	0	0	0
	G	115	243	С																
rd1	S	243	345													102	0	0	0	0
	G	0	223	С																
rd2O	S	270	599													329	0	0	0	0
	G	0	255	С																
rd3O	S	0	134		283	368										368	149	1	0.40	149
	G	169	259	D矢	283	564	D矢													
rd4O	S	63	219									<u></u>				156	0	0	0	0
	G	0	231	С																
rd5O	S	302	612													310	0	0	0	0
	G	0	229	С																
rg1O	S	0	140		226	401		718	980	1273	1529	1630	1943			1943	977	4	0.50	244
	G	155	421	С	616	1262	D													
rg2O	S	0	165		304	818										818	139	1	0.17	139
	G	319	441	С	441	964	D													

Turn ID		Start	End	Tsp	ΣΤρ	Np	ΣTp/Tsp	ΣTp/Np															
rg3〇	S	0	256		379	491									*				491	123	1	0.25	123
	G	203	552	D																			
rg4O	S	0	25		366	543		1703	2214										2214	1160	2	0.52	580
	G	428	557	С	583	1389	D	2256	2325	D													
rg5O	S	318	597																279	0	0	0	0
	G	0	170	С											,			-					
rg6O	S	0	220																220	0	0	0	0
	G	54	231	С																			
rg7O	S	0	228		297	454													454	69	1	0.15	69
	G	189	521	С																			
rg8	S	114	247																133	0	0	0	0
	G	0	151	С																			
rpl	S	0	207		441	726		835	1005		1225	1323		1436	1573		1719	1906	3309	1682	11	0.51	153
	G	- 466	576	D	835	1183	D	1225	1823	D	2223	2395	D	2804	3238	D							
	S	2029	2105		2223	2305		2428	2502		2641	2655		2804	3031		3242	3309					
	G																						
rp2O	S	0	32		423	470		559	685		975	1017		1109	1175				1175	862	4	0.73	216
	G	62	510	D	797	973	С																400000000000000000000000000000000000000
rp3O	S	0	83															-	83	0	0	0	0
	G	65	224	С																			

. ^

Turn ID		Start	End	Tsp	ΣΤρ	Np	Σ Tp/Tsp	ΣTp/Np															
rp4O	S	0	55		345	410		584	699		790	854		1003	1026		1252	1446	2276	1684	8	0.74	211
	G	345	490	D	832	1302	D	1302	1599	С	1811	2207	С										
	S	1611	1683		1836	1886		2222	2276														
	G																						
rp5O	S	0	116		256	286		376	419		502	812		1045	1200				1200	546	4	0.46	137
	G	243	650	D	1096	1183	D																
rp6O	S	0	109		463	627									-				627	354	1	0.56	354
	G	149	453	С																			
ss1	S	0	103																103	0	0	0	0
	G	240	558	С														-					
ss2O	S	0	209		328	408		736	957										957	447	. 2	0.47	224
	G	119	271	D	337	625	D	1059	1248	D													
ss3〇	S	0	163		517	954													954	354	1	0.37	354
	G	269	506	С																			
																·							

付録D 発話の時刻 (WOZ/TEL/AGENT発話)

WOZ実験の電話対話におけるAGENT発話それぞれに対して、ポーズ(500msec以上の無音状態とした)で区切られた発話の開始・終了時刻を記録した。表中の記号と数字は以下の意味である。

Start, End 発話の開始・終了時刻(単位は0.01sec)

S 発話

Tsp ポーズを含めた全発話時間(単位は0.01sec)

ΣΤρ 総ポーズ時間 (単位は0.01sec)

Np ポーズの個数

Turn ID		Start	End	Start	End	Start	End	Start	End	Start	End	Start	End	Tsp	ΣΤρ	Np	ΣTp/Tsp	ΣTp/Np
ac1	S	0	105	213	334					 **********				334	108	1	0.32	108
ac2	S	0	225											225	Ó	0	0	0
ac3	S	0	191											191	0	0	0	0
ac4	S	0	216	362	602									602	146	1	0.25	146
ac5	S	0	196	285	548									548	89	l	0.16	89
ac6	S	0	185											185	0	0	0	0
ac7	S	0	236											236	0	0	0	0
ac8	S	0	211											211	0	0	. 0	0
an1	S	0	35	 104	313	426	561							561	182	2	0.32	91
an2	S	0	37	115	281									281	78	1	0.28	78
an3	S	0	188	 284	443	500	729							729	153	2	0.21	77
an4	S	0	212	312	513	649	853							853	236	2	0.28	118
an5	S	0	262	360	423	497	769	865	1195					1195	268	3	0.22	89
an6	S	0	118	250	359	437	698							698	196	2	0.28	98
dp1	S	0	49	120	321	404	639					····		639	154	2	0.24	. 77
dp2	S	0	466	 547	663	729	1014							1014	147	2	0.14	74
dp3	S	0	254	 343	426	530	644	735	1218					1218	180	3	0.15	60
dp4	S	. 0	166	402	858									858	236	1	0.28	236

න .

Turn ID		Start	End	5	Start	End	Start	End	Start	End	Start	End	Start	End	Tsp	ΣΤρ	Np	ΣTp/Tsp	ΣTp/Np
dp5	S	0	362		426	519	645	743	852	986					986	299	3	0.30	100
dp6	S	0	318		391	746			-						746	73	1	0.10	73
er1	s	0	73		192	1101	 1220	1272	1372	1546					1546	338	3	0.22	113
er2	S	0	30		105	143	226	467	590	1056	1168	1396	1500	1589	1589	497	5	0.31	99
er3	S	0	196		298	617									617	102	1	0.17	102
er4	S	0	269		388	646	765	963	1046	1233					1233	321	3	0.26	107
lf1	S	0	103		188	366	504	588	954	1003					1003	589	3	0.59	196
lf2	S	0	344		433	511									511	89	1	0.17	89
lf3	S	0	98		182	443									443	84	. 1	0.19	84
lf4	S	0	381												381	0	0	0	0
lf5	S	0	344		413	653									653	69	1	0.11	69
lf6	S	0	296												296	0	0	0	0
1f7	S	0	222			-									222	0	0	0	0
1f8	S	0	191												191	0	0	0	0
1f9	S	0	243		604	658									658	361	1	0.55	361
lf10	S	0	24		347	603									603	323	1	0.54	323
lf11	S	0	305												305	0	0	0	0
If12	S	0	351												351	0	0	, 0	0

Turn ID		Start	End	Start	End	Start	End	Start	End		Start	End	Start	End		Tsp	ΣΤρ	Np	Σ Tp/Tsp	Σ Tp/Np
1f13	S	0	150	263	636											636	113	1	0.18	113
lf14	S	0	603												-	603	0	0	0	0
lf15	S	0	277	413	746											746	136	ı	0.18	136
lf16	S	0	208	306	694					·						694	98	1	0.14	98
nol	S	0	51	137	347	508	743									743	247	2	0.33	124
no2	S	0	616													616	0	0	0	0
no3	S	0	36.	115	184	289	427	500	773							773	178	3	0.23	59
no4	S	0	24	137	568											568	113	1	0.20	113
no5	S	0	45	181	1022	1108	1313									1313	222	2	0.17	111
поб	S	0	824													824	0	0	0	0
no7	S	- 0	53	233	343	436	995									995	273	2	0.27	137
no8	S	0	209	389	603											603	180	1	0.30	180
rdI	S	0	115	223	269	350	441	526	818							818	274	3	0.33	91
rd2	S	0	118	215	654	773	846	1169	1450							1450	539	3	0.41	180
rd3	S	0	109	196	258	326	589									589	155	2	0.26	78
rgl	S	0	371	464	728											728	93	1	0.13	93
rg2	S	0	126	215	574	653	851									851	168	2	0.20	84
rg3	S	0	335													335	0	0	0	0

Turn ID		Start	End	Start	End	Start	End		Start	End	Start	End		Start	End	Tsp	ΣТр	Np	ΣTp/Tsp	ΣTp/Np
гр19	S	0	66	266	463	 576	697					·	· · · · · · · · · · · · · · · · · · ·			697	313	2	0.45	157
ss1	S	0	44	279	459	580	678		824	949	1062	1337				1337	615	4	0.46	144
ss2	S	0	33	126	350	419	704									704	162	2	0.23	81
ss3	S	0	259	371	629											629	112	1	0.18	112
ss4	S	0	282	 377,	654		:	Y-2-1								654	95	1	0.15	95
ss5	S	0	396													396	0	0	0	0
ss6	S	0	495													495	0	0	0	0
ss7	S	- 0	90	170	623											623	80	1	0.13	80
ss8	S	0	567	730	997											997	163	1	0.16	163
ss9	S	0	117	190	499											499	73	1	0.15	73
ss10	S	0	57	148	739	 862	998									998	214	2	0.21	107
ss11	S	0	57	151	483											483	94	1	0.19	94
								<u> </u>												
																			,,	
				1																