

話し合いを評価するための評価パラメータの 検討(1)

水上悦雄^a, 森本郁代^b, 大塚裕子^c, 鈴木佳奈^a
竹内和広^d, 奥村学^e, 柏岡秀紀^a

^a ATR, ^b 関西学院大, ^c IBS, ^d 大阪電通大, ^e 東工大

<http://lssl.jp/>

1. 背景・問題意識

- よい議論のためのコミュニケーションプロセス評価指標の探求
- 7つの評価指標およびその下位項目によるコミュニケーションおよびディスカッションのチェック項目の抽出(水上他, 2008; 鈴木他, 2008など)



- ・印象評定＋因子分析
- ・場面相対評価
- ・話し合い過程の言語的やり取りの分析

7つの評価指標

評価指標	評価の観点
誠実な参加態度	自分の意見をしっかり伝え、人の発言をしっかり聞いたか
対等な関係性	全員が対等に議論に参加したか
議論の活発さ	議論は活発だったか
意見の多様さ	いろいろな意見が出ていたか
議論の深まり	一つ一つの意見が十分に検討されていたか
議論の管理	議論の流れがしっかりコントロールされていたか
意見の積み上げ	結論に向かって一つ一つの意見が積み上げられていたか

- 評価時に、評価者の視点、基準、受け止め方等の主観によりゆれが生じる
- 客観的に測定可能なものなのか？
(内的評価基準に左右されない評価は可能か？)

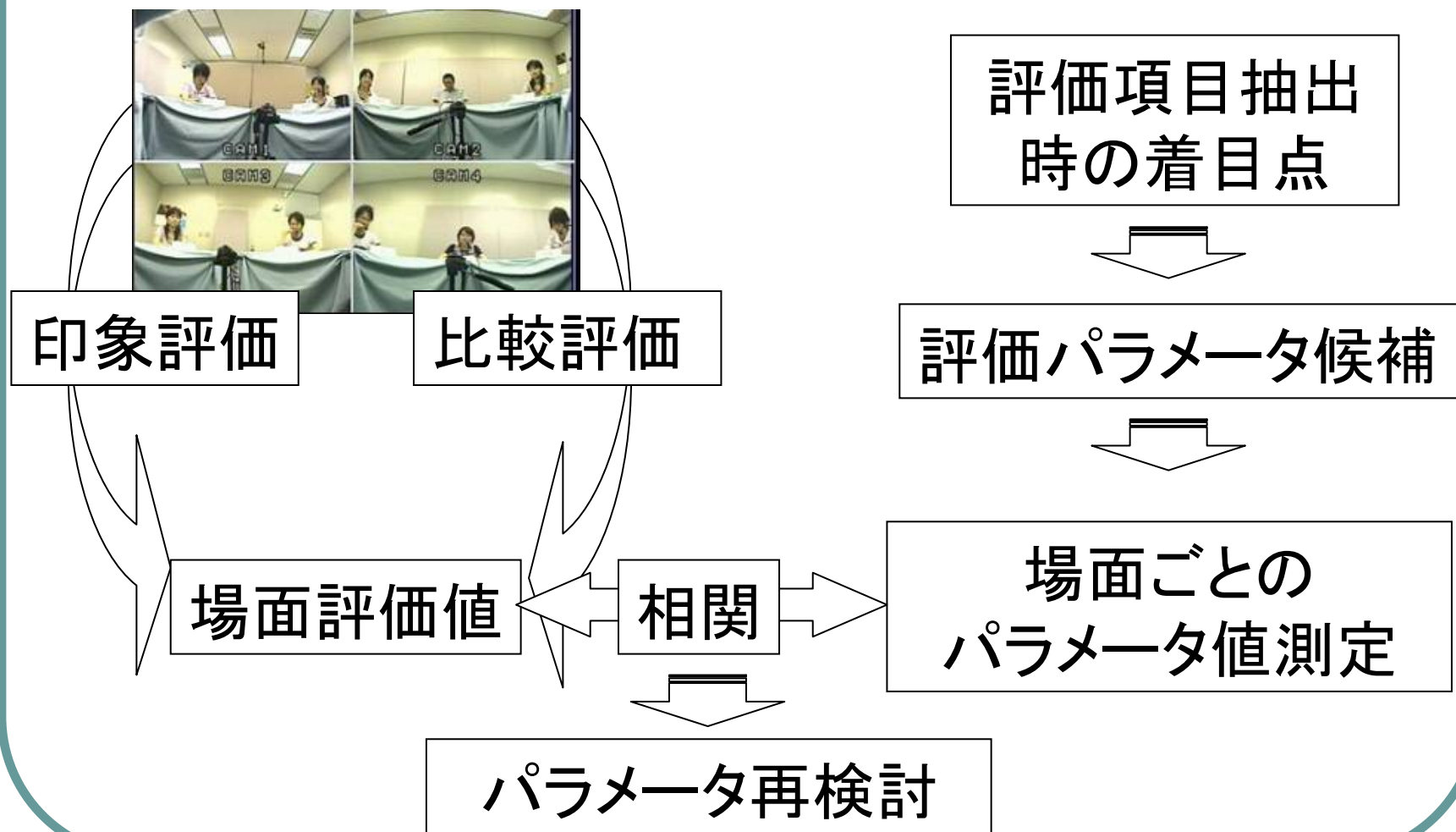
2. 本研究の目的

- 基準のゆれの少ない可能な限り客観的な評価ができないか？
- 抽出された評価項目は、具体的にどのような言語的・非言語的要素(評価パラメータ)によって、近似・測定できるのか？



- 各評価項目に対応する評価パラメータを検討
- 今回は、主に「2. 対等な関係性」に関わるパラメータを検討

3-1. 手法のアウトライン



3-2. 対象データ

- 27 場面

- 9グループ(大学2~4回)
- 6名 / グループ
 - 男性:女性=理系:文系=1:1
- 9グループ × 3回 → 27回
- 各回の10分経過後の8分間



- このうち、後述する4場面を分析対象

- 順位得点と因子得点(関係性因子のみ)の正負の組みあわせ

3-3. 二つの評価法：印象評定と順位付け

- 印象評定 → 因子分析 → 因子得点
 - 一つのダミーを評定後, 三つの場面を評定
 - 23対の対義語, 7段階評定
 - 印象得点
 - 各因子の平均因子得点から寄与率を加味して算出
- 順位付け → 順位得点
 - 三つの場面を比較し, いい議論だと思ふ順に並べる
 - $R(i) = N_1(i)*20 + N_2(i)*5 - N_3(i)*5$ (6名の評者の総合得点)
- 被験者
 - ディスカッションに参加した54名
 - 自分が参加していないグループを評価

3-4. パラメータ候補

2. 対等な関係性に関わる

下位項目	関連するパラメータ候補
個人攻撃になっていないか	不同意数, 言語情報 (含・感情情報)
一部の参加者が議論の輪から外れていないか	各人のフロア数, 発話交換 (発言・応答ペア)の成立数
専門用語などわかりにくい表現を使っていないか	言語情報, 問い返し(意味確認)数 (-)
対立意見を頭から否定し, 感情的に反応していないか	不同意数, 言語情報 (含・感情情報)
反論された際の反応として, 沈黙や安易な同意を示していないか	(反論の)連鎖数, 言語情報 (含・主張の整合性)

本発表では, 上記2. に関する
パラメータ候補を検討

1. 誠実な参加態度に関わる

下位項目	関連するパラメータ候補
議論に集中しているか (聴く態度を示しているか)	(話し手以外の)視線, あいづち等の有無, (しぐさ)
発言の根拠を提示するよう ところがけているか	言語情報 (修辞構造)
明確な発言をところがけて いるか	言語情報 (統語構造), 音声 のパワー

3. 議論の活発さに関わる

下位項目	関連するパラメータ候補
誰かの発言中, 傍観者とな っていないか	応答者バランス, 音声的同意 表現の有無
誰かの発言に対して, 応答 しようとしているか	発話量, ポーズ量, 発言- 応答ペアの成立数
発言者が偏っていないか	各人のフロア数
発言の受け止め役やまとめ 役が固定されていないか	各人のフロア数, メタ議論 発言者バランス

4. 以降は省略

(参考) 各因子と順位点の相関

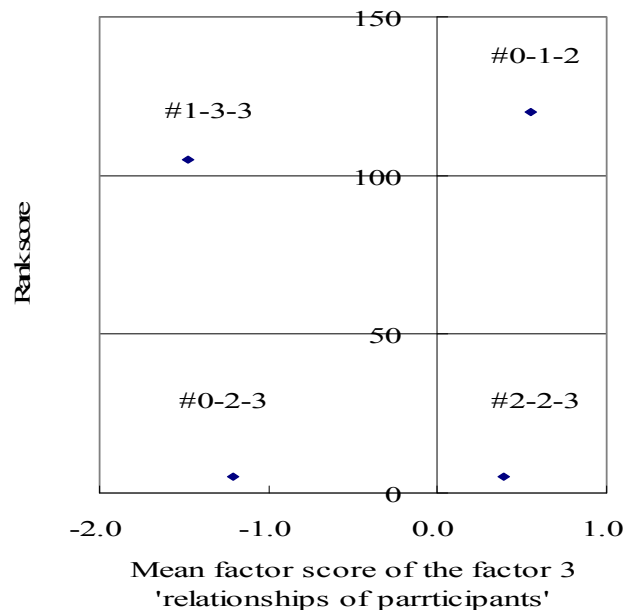
	Kendall's τ	P (** $p < 0.01$)
活発さ	0.527	**
多角・統合	0.452	**
関係性	0.118	n.s.
展開・洗練	0.591	**
誠実さ	0.464	**

7つの評価項目と、因子との関係性

1. 誠実な参加態度 ← 第5因子 誠実さ
2. 対等な関係性 ← 第3因子 関係性
3. 議論の活発さ ← 第1因子 活発さ
4. 多様な意見 ← 第2因子 多角
5. 議論の深まり ← 第4因子 展開
6. 議論の管理 ← 第2因子 統合
7. 意見の積み上げ ← 第4因子 洗練

関係性因子のみ順位得点とは相関がなく、注意が必要

3-5. 分析対象の 選択



	#0-2-3	#2-2-3	#2-2-3	#0-1-2
順位得点	低	低	高	高
因子得点 (関係性)	低	高	低	高

場面 番号	Factor 1 活発さ	Factor 2 多角 統合	Factor 3 関係性	Factor 4 展開 洗練	Factor 5 誠実さ	印象 得点	ランク 得点
0-1-2	0.795	0.568	0.552	0.307	0.437	0.272	120
<u>1-3-3</u>	<u>0.055</u>	<u>-0.657</u>	<u>-1.474</u>	<u>-0.570</u>	<u>-0.592</u>	<u>-0.210</u>	<u>105</u>
2-1-2	-0.016	0.518	0.332	0.689	0.192	0.119	90
0-1-1	0.323	0.213	0.169	0.011	0.168	0.097	80
1-3-1	0.209	1.049	-0.185	0.931	0.141	0.185	75
1-3-2	0.806	-0.226	-1.659	-0.277	0.173	0.000	75
2-1-1	-0.468	0.400	-0.769	-0.226	-0.503	-0.139	75
2-1-3	0.385	-0.138	-0.017	0.021	-0.099	0.057	60
0-1-3	0.279	0.643	0.676	0.188	0.390	0.182	55
1-2-3	0.555	-0.530	0.535	-0.135	0.168	0.092	55
0-3-1	0.502	-0.029	0.053	0.089	0.457	0.116	50
2-3-3	-0.767	-0.258	-0.238	-0.216	-0.055	-0.199	45
0-3-2	0.054	-0.235	0.084	0.070	-0.223	-0.007	40
2-2-2	-1.159	-0.572	0.047	-0.790	-0.845	-0.345	35
2-3-1	-1.138	-0.031	0.435	-0.568	-0.252	-0.231	35
0-3-3	0.310	-0.954	-0.151	-0.203	-0.254	-0.058	30
1-1-1	-1.196	-0.533	0.225	-0.544	-0.574	-0.309	25
2-3-2	-1.107	-0.880	0.480	-0.655	-0.481	-0.310	25
1-2-2	-0.273	-0.302	0.187	-0.582	-0.298	-0.113	20
0-2-2	-0.602	-0.447	-0.167	-0.970	-0.421	-0.243	15
1-1-2	-0.893	-0.702	-0.308	-0.733	-0.568	-0.318	10
0-2-3	-0.556	-1.309	-1.207	-1.365	-0.296	-0.405	5
2-2-3	-0.292	-0.996	0.391	-0.528	-0.267	-0.158	5
1-2-1	-1.258	0.230	-0.276	-0.787	-1.024	-0.322	0
1-1-3	-0.993	0.067	0.048	-0.921	-0.139	-0.243	-10
2-2-1	-1.205	-1.276	-0.028	-1.059	-0.587	-0.429	-10
0-2-1	-1.401	-0.917	-0.471	-0.926	-0.902	-0.467	-30

3-6. 同意／不同意タグの付与手順

- 発話単位化
 - 150msecのポーズに区切られた発話区間
 - 統語的, 意味的つながりを加味し結合
 - 節単位強境界末(接続詞「けど」「て」「し」+語尾ストレスはつなげる, 節単位強境界末+下降調, 上昇調はつなげないなどのメタルールを設ける
- 話し合いのための談話タグ(pDTD)の付与
 - 実際には, その前処理としてDialogue Act(DAMSLを一部改変)を付与
 - 文脈を考慮したうえで, その発言が先行発言の同意となっているのか, 不同意となっているのかを判別
 - 試行的に著者が付与

pDTD	説明
propose	新規のあるいは改変された提案・主張
reason	提案・主張の理由付け
question	他者の発言に対する質問
answer	質問に対する応答
downgrade	主張の取り下げあるいは軟化
exemplify	主張に関する例示
substream	主題から離れたやり取り
agreement	同意応答・発言
agree	明示的な同意
agree/propose	同意+提案・主張
agree/reason	同意+理由付け
weak-agree	非明示的な同意
weak-agree/propose	非明示的な同意+提案・主張
weak-agree/reason	非明示的な同意+理由付け
disagreement	不同意応答・発言
disagree	明示的な不同意
disagree/propose	不同意+提案・主張
disagree/reason	不同意+理由付け
weak-disagree	非明示的な不同意
weak-disagree/propose	非明示的な不同意+提案・主張
weak-disagree/reason	非明示的な不同意+理由付け
reserve-agree	同意の保留

4-1.結果1:同意／不同意の頻度

- 因子得点の正負と、不同意頻度が、関わる？
- 特に、順位得点も因子得点も低い#0-2-3は、不同意／同意比が大
- #0-1-2では明示的な否定表現は一切なかった

pDAD	1-3-3(NG)	0-2-3(NB)	0-1-2(PG)	2-2-3(PB)
propose	31	16	25	27
reason	24	14	6	21
downgrade	2	0	2	0
exemplify	1	4	8	2
question	0	4	3	6
answer	0	4	4	7
agree	44	16	64	97
agree/propose	0	2	4	3
agree/reason	3	2	2	7
weak-agree	1	2	8	1
weak-agree/reason	1	0	0	0
weak-agree/propose	1	0	0	0
Total agreement	50	22	78	108
disagree	0	2	0	1
disagree/reason	2	1	0	0
disagree/propose	1	3	0	1
weak-disagree	1	1	1	1
weak-disagree/reason	3	0	0	0
weak-disagree/propose	0	0	1	1
reserve-agree	2	2	1	0
Total disagreement	9	9	3	4
Total utterances	117	73	129	175
Rate of Disagree	7.7%	12.3%	2.3%	2.3%
Disagree/Agree	18.0%	40.9%	3.8%	3.7%

4-2.結果2:同意／不同意応答の傾向

	1-3-3(NG)			0-2-3(NB)			0-1-2(PG)			2-2-3(PB)		
	count	rate	recipient	count	rate	recipient	count	rate	recipient	count	rate	recipient
propose-agree	11	50.0%	1.6	8	50.0%	1.1	21	80.8%	1.9	29	90.6%	2.3
propose-disagree	5	22.7%	1.0	2	12.5%	1.5	2	7.7%	1.0	2	6.3%	1.0
propose-φ	6	27.3%	-	6	37.5%	-	3	11.5%	-	1	3.1%	-
reason-agree	21	77.8%	1.3	9	52.9%	1.4	7	87.5%	1.9	17	85.0%	1.8
reason-disagree	4	14.8%	1.0	4	23.5%	1.8	1	12.5%	0.0	0	0.0%	0.0
reason-φ	2	7.4%	-	4	23.5%	-	0	0.0%	-	3	15.0%	-
question-answer	0	0.0%	0.0	4	100.0%	1.0	4	100.0%	2.0	6	100.0%	1.3
question-φ	0	0.0%	-	0	0.0%	-	0	0.0%	-	0	0.0%	-
answer-agree	0	0.0%	0.0	0	0.0%	0.0	3	75.0%	2.3	5	71.4%	1.6
answer-disagree	0	0.0%	0.0	1	25.0%	1.0	0	0.0%	0.0	1	14.3%	1.0
answer-φ	0	0.0%	-	3	75.0%	-	1	25.0%	-	1	14.3%	-

- #0-2-3は、応答部がない(φ)場面が目立つ
 - 提案や、理由付け、質問の答え、に対する反応(評価)がない

5. 結論

- 「2.対等な関係性」に関わる評価パラメータとして、以下のパラメータ候補の利用可能性が示唆
 - 不同意／同意比率
 - 応答ペアの成立／不成立

6.課題

- 同意／不同意タグを含む，話し合いのための談話タグ自体の検討の必要性
 - 様々なレベルの同意／不同意の混在
 - 提案に対する同意
 - 「~ということでもいいですか?」「はい」
 - 理由説明への了解
 - 「~があったりするでしょう?」「ええ」
 - 質問という形をとった反論
 - 「ていうことは, ~なんじゃないの?(矛盾の指摘)」
 - など, 同意／不同意の多様性をどう整理するか
 - 同意／不同意と, 発話単位との関係性
 - 一つの発話単位で一つの同意／不同意
 - 複数の発話単位で一つの同意／不同意
 - 今回は, 各発話単位にpDTDを付与
- その他のパラメータの検討
 - 単純な発話統計量だけでは, 評価困難

(参考)その他のパラメータ候補との相関

● 単純統計量

- ポーズ量(誰も発話していない総時間率)
 - 活発さ因子と高い負の相関
 - $\tau = -0.425$ $P < 0.01$
 - 活発さの指数として, 利用可能性あり
- 発話バランス(各話者発話時間の偏差)
 - 関係性因子や活発さの因子との相関なし
 - 発話量の比率は, そのまま評価パラメータとして利用はできない←役割によって, 発話量は変わる
- 発話バランス(各話者発話単位数の偏差)
 - 関係性因子との負の相関
 - $\tau = -0.276$ $P < 0.05$
 - 見かけ上の発話回数=議論に貢献する発言の数とは限らない⇒発言内容の加味の必要性