

第3章 状態提示行為制御のモデル化

3.1 はじめに

音声対話インタフェースにおいては、対話の様々な状況を理解し、状況の推移を制御する対話制御技術の重要性が唱えられている [6]。対話制御によって話題を限定することで音声認識対象を絞り込んだり、状況に応じて対話の主導権を移行したりすることでより円滑な対話が実現できる。

対話制御技術の機能目標は、アプリケーションに依存するところが多い [26]。例えば、実用的な公共システムへの応用を考えた場合には正確さや効率の良さが優先的に要求されるであろうし、娯楽システムへの応用の場合には親和性や協調性が要求されるであろう。従って、対話制御モデルの構築においては、各アプリケーションに要求される機能目標をあらかじめ明らかにし、それに依拠してモデル化することが適切である。ここでは、一般性が高くなおかつ実現可能性が高いと考えられるデータベース検索タスクを取り上げ、有効な対話制御モデルを構築することを検討する。

以下、3.2では、主観的な評価指標の特性と対話制御の機能目標を明らかにすべく、人間同士の模擬的な対話を収集して評価を行った実験について説明する。3.3では、実験結果に基づいて評価指標の特性と機能目標の分析を行う。分析結果から導かれる重要な結論として、被験者によらず効率が重視されることから、3.4では、効率に関する収集対話に対する主観評価結果と客観量との関係を分析し、効率の主観評価向上に影響を与える客観量を明らかにする。次に3.5では、効率向上を機能目標とする対話制御モデルについて説明する。特に、状態提示発話の最適な出力タイミングを決定するための状態エントロピーのモデルについて具体的に示す。最後に、提案した対話制御モデルの一部を実際の対話システムに実装、評価した結果を3.6に示す。

3.2 対話評価実験

これまでに音声対話システムの構築を目的として、人間同士の対話が多数収録されているが [27]、当事者あるいは第三者の視点のいずれにおいてもその対話を評価した例はない。人間対システムの対話においては、システムの応答やふるまいについて評価した例が多数見られるが、複数の評価指標間の相関関係を明らかにした例はない。

ここでは、まずタスクに依存した機能目標を明らかにするとともに人間同士の対話における評価指標の特性を調べることを目的として、人間同士の対話を収録・評価する実験を行なう。なお、計算機で人間との対話を実現する際には、人間対システムの対話における評価指標の特性を調べるべきとする立場もあり得る。しかし、本研究はあくまでも、ユーザインタフェースとしてコミュニケーション効率の向上、発想の支援、心理的負担の軽減などを目的として人間同士に近い対話を計算機上で実現しようとするものであるため、人間同士の対話における評価指標の特性のみに焦点をあてる。

3.2.1 評価指標についての仮説

まずはじめに、対話の評価指標について幾つかの仮説をたてる。実験はこの仮説を検証する形で行なう。なお、指標とは、評価における観点を意味する。

仮説 H_1

対話聴取後の評価において、評価者本人が参加した対話とそうでない対話の評価で重視する指標は変わらない。

仮説 H_2

人間が対話において重視する指標には個人差がある。

仮説 H_3

指標間の相関関係は様々であるが、個人差は少ない。

仮説 H_4

対話後と対話聴取後とでは重視する指標が異なる (ものがある)。

仮説 H_1 については、対話聴取後の評価結果について、総合的評価と各指標との相関関係について、評価者自身が参加した対話とそうでない対話とを比較することで検証する。

この仮説が正しいことが証明されれば、後の仮説検証においては対話聴取後評価として、必ずしも各被験者が参加した対話に対する評価結果だけでなくよりサンプル数の多い全対話に対する評価結果を用いることが可能となる。

仮説 H_2 については、対話評価における総合評価と各指標の相関関係を被験者毎に調べ、相関の大小から指標としての影響力の違いを明らかにすることで検証する。

仮説 H_3 については、指標間の相関関係を調べ、相関の大小を被験者毎に比較することで検証する。

仮説 H_4 については、同一被験者の対話後と対話聴取後の評価結果の相関を、総合評価および各指標について調べる。また上述の仮説 H_2 および H_3 の検証結果が対話後と対話聴取後でどのように異なるかを調べる。なお、対話後と対話聴取後の評価には、前者は当事者、後者は第三者となる視点の違いによる差が生じることが予想される。対話制御の評価を簡便にするために聴取評価が望まれ、そのためにあらかじめ対話後の評価と聴取評価の関係を明らかにしておく必要がある。

次節にこれらの仮説を検証するために行なう実験の概要を示す。

3.2.2 対話評価実験の概要

本実験では人間同士の対話における評価指標の特性を調べることを目的としているため、人間同士の生きた対話の評価する必要がある。ここでは、音声対話検索システムを想定して、検索オペレータとユーザの対話を行なうことにする。ユーザには、自然な検索要求を持たせるために、学術文献を読ませその文献を理解するために必要な論文を探すように教示する。教示文を図 3.1 に示す。

なお、与えた文献は以下の3件である。

- 田窪他、"岩波講座・言語の科学 2. 音声", 岩波書店、pp.218-220, 1998
- 辻野他、"イントラネット・エクストラネット上での利用者相互の情報流通に関する一考察," 電子情報通信学会総合大会講演論文集, pp.200, 1999.3.
- 植垣他、"発話速度適応型音声再生速度変換手法の検討," 電子情報通信学会総合大会講演論文集, pp.222, 1999.3.

手順

1. 与えられた文献を一通りよく読んで理解して下さい。
2. 理解するために必要な論文を、オペレータとの対話によって探して下さい。
3. 探した論文のタイトルや内容などを聞いて、欲しいと思っていた論文がありそうだったらその旨をオペレータに伝えて対話を終えて下さい。
4. 対話によって論文を探し終えたら、アンケートに答えて下さい。

注意事項

- 文献によっては理解が難しいものもあります。その場合は理解が不完全でも構わないですから手順通り理解に必要な論文を探して下さい。
- オペレータはあなたの代わりにデータベース検索を行いません。
- オペレータとは電話で会話をするようなつもりで話して下さい。
- オペレータはあなたに渡された文献を持っていません。

図 3.1: 対話収録時の教示文

(1) 対話収録環境

対話収録は、早稲田大学理工学部 55 号館地下一階マルチメディアスタジオにて行なった。被験者はスタジオ内、オペレータは操作室に入り、互いが見えない非対面の状態で対話を行なった。オペレータは、ネットワーク接続されたパーソナルコンピュータを用いて学術情報センターの論文検索サービス*によって、被験者の検索要求に対して随時検索を行ない、その結果に基づいて対話を行なった。

なお、実際のシステム利用時の状況を想定して、対話中に文献を見ても構わないことにした。ただし、被験者の検索要求が拘束されないようにするために、文献の中の参考文献欄は伏せた。また、システム役は文献を見ず、その旨を被験者にあらかじめ伝えた。

表 3.1 に対話収録環境の概要を示す。

以上の環境で、被験者は 5 名（いずれもドメインにあまり精通していない学部 4 年生）、オペレータは 1 名（ドメインに精通した博士課程の学生）文献は上述した 3 件で、計 15 対話を行ない収録した。

*<http://www.nacsis.ac.jp/els/els-j.html>

表 3.1: 対話収録環境の概要

対話環境	人間=人間
対話様式	自由対話
対話形態	非対面
非言語情報	文献参照可能
タスク	論文検索
録音環境	スタジオ
マイクروفオン	接話型 (SENNHEISER HMD25-1)

なお、本実験では対話評価における指標間の相関を調べるのが主な目的であるため、確認の仕方や間の置き方などのオペレータの対話方法が理想的である必然性はない。ただし、一人の被験者が3件の文献を理解するために行なう検索対話について、1件の検索が終了する毎に評価を行なうために、オペレータの対話方法は一つの対話を通じてはある程度一貫している必要がある。しかし全ての対話で対話方法が同じであれば、相関関係の分析結果の信頼性が低くなる可能性があるため、対話毎に対話方法を変えるようにオペレータに意識させた。

(2) 対話評価手順

被験者には対話終了毎にアンケートに答えてもらう。アンケートは、表 3.2 に示した全項目について5段階評価を記入する形にした。

表中の (a) は対話についての指標、(b) は相手の参加者についての指標である。なお、表 3.2 においては、暫定的に総合評価、効率、透過性、リズム、主体性、多様性と分類しているが、それらの分類、およびその下位分類となっている指標およびその質問文が分類として妥当かどうかは実験結果の分析により判明する。この時点では著者自身の主観的な分類である。

また、3.2.1 の仮説 H_4 に述べたように、対話後の評価と聴取評価の比較を行なうために、同様のアンケート評価を数日後に行なわせた。その際には、各対話音声データをオーディオファイルに変換し、被験者には繰り返し聞くことを許した。また、被験者には自分が参加した対話以外の全ての対話を聴取評価させた。

表 3.2: 実験に用いる対話評価指標

評価指標	質問文	記号
総合評価 (a)	「対話ほうまくできましたか」	Xa
総合評価 (b)	「相手の話し方は総合的に良かったですか」	Xb
効率 (a)	「効率的な対話ことができましたか」	Aa
効率 (b)	「相手の話し方は効率的でしたか」	Ab
論文検索の効率 (a)	「論文を効率良く探せましたか」	A1a
無駄時間 (a)	「対話に無駄な時間はなかったですか」	A2a
無駄時間 (b)	「相手の話し方に無駄な時間はなかったですか」	A2b
発話量 (a)	「対話の量は適切でしたか」	A3a
発話量 (b)	「相手の話の量は適切でしたか」	A3b
透過性 (b)	「相手の状態はわかりやすかったですか」	Bb
処理状態 (b)	「相手が何をしているかわかりましたか」	B1b
意思伝達 (b)	「自分の言葉が伝わっているかわかりましたか」	B2b
安心感 (b)	「話し方に安心感を感じましたか」	B3b
リズム (a)	「対話のリズムは良かったですか」	Ca
リズム (b)	「相手の話し方のリズムは良かったですか」	Cb
間 (a)	「対話の間は良かったですか」	C1a
間 (b)	「相手の間は良かったですか」	C1b
スピード (b)	「相手の話のスピードは良かったですか」	C2b
バランス (a)	「お互いの話の量のバランスは適切でしたか」	C3a
主体性 (主導権) (a)	「主導権のバランスは良かったですか」	Da
表現の多様性 (b)	「相手の言葉は多様でしたか」	E1b
話し方の多様性 (b)	「相手の話し方は多様でしたか」	E2b
リズム (b)	「相手のリズムは多様でしたか」	E3b

3.3 評価指標と機能目標の分析

ここでは、3.2に示した実験によって得られた対話評価結果について、3.2.1に示した仮説を検証する形で分析する。

3.3.1 仮説 H_1 の検証：聴取評価における対話参加・不参加の影響

(1) 対話についての総合指標: Xa との相関係数

評価時期：対話聴取後 (HE)

評価対象：対話についての総合指標 (a)

まず、対話についての総合的評価と大まかな指標との相関について、被験者によらず相関係数を調べた結果および各被験者毎に相関係数を調べた結果を表 3.3 に示す。なお (a)

が評価者自身が参加した対話について、(b) が評価者自身が参加していない対話についての分析結果である。また、以降の表において太字は無相関の検定によって1%の危険率で有意差が認められたものである。

表 3.3: 対話についての総合指標: X_a との相関係数

(a) 評価者自身が参加した対話

被験者	全体	A	B	C	D	E
Aa (効率)	0.7353	0.8660	0.8386	—	1.0000	0.5000
Ca (リズム)	0.7320	0.5000	0.5000	—	0.5000	0.5000
Da (主体性)	0.6130	0.8660	0.1890	1.0000	0.9449	-0.5000

(b) 評価者自身が参加していない対話

被験者	全体	A	B	C	D	E
Aa (効率)	0.8081	0.7998	0.8300	0.7467	0.7866	0.9389
Ca (リズム)	0.7479	0.7870	0.6355	0.6869	0.9087	0.9532
Da (主体性)	0.4209	0.2791	0.0251	0.6401	0.7255	0.9047

この表において、(a) に示した各被験者毎の相関係数は各 3 対話のみの評価結果から求めているためやや信頼性に欠けるが、それを考慮した上で (b) と比較すると、まず、Aa (効率) については、(a)(b) とともにほぼ全員の評価結果が強い相関を示している。(a) の被験者 C は Aa の評価が全て等しかったために相関係数の算出が不可能)。このことから、聴取評価において Aa が重視される度合は本人の参加・不参加に依存しないと言える。

一方、Ca (リズム) については、全体では (a)(b) いずれも高い相関を示しているのが、被験者毎に見ると (a)(b) で相関の強さが多少異なっている。ただ、いずれも正の比較的高い相関を示していることから、評価対象件数を増やせば (a) での結果が (b) の結果に近づく可能性も考えられる。ここでは、聴取評価における Ca の重要性が本人参加・不参加に依存するとは言えないとしておく。

ところが、Da (主体性) については、全体でも被験者毎でも相関の強弱の傾向が (a) と (b) で明らかに異なる。例えば、被験者 E などは (a) では負の高い相関を示しているのが (b) では正の高い相関を示している。このことから、こと Da に限っては、その重要性に評価の対象となる対話への本人の参加・不参加が影響を及ぼしていると言える。これは、Da が主体性を示す指標であり、主体性を評価する際には主体が本人であるか否かが大きく係わるためと考えられる。

(2) 対話についての詳細指標: Xa との相関係数

評価時期: 対話聴取後 (HE)

評価対象: 対話についての詳細指標 (a)

ここでは前項に示した総合指標の下位分類とした詳細指標と総合的評価との相関関係を分析する。

まず、対話についての総合的評価と詳細な指標との相関について、被験者によらず相関係数を調べた結果および各被験者毎に相関係数を調べた結果を表 3.4 に示す。なお (a) が評価者自身が参加した対話について、(b) が評価者自身が参加していない対話についての分析結果である。

表 3.4: 対話についての詳細指標: Xa との相関係数

(a) 評価者自身が参加した対話

被験者	全体	A	B	C	D	E
A1a (論文検索効率)	0.7525	1.0000	0.6547	-0.5000	—	1.0000
A2a (無駄時間)	0.6886	0.7559	0.6547	-0.5000	1.0000	0.5000
A3a (発話量)	0.6442	0.8660	0.9449	0.5000	—	—
C1a (間)	0.6986	0.5000	0.7857	-0.5000	1.0000	1.0000
C3a (バランス)	0.6773	1.0000	0.1890	1.0000	1.0000	—

(b) 評価者自身が参加していない対話

被験者	全体	A	B	C	D	E
A1a (論文検索効率)	0.6982	0.8227	0.6678	0.6262	0.8647	0.9447
A2a (無駄時間)	0.7786	0.4563	0.8008	0.7095	0.8883	0.8744
A3a (発話量)	0.7437	0.8181	0.7242	0.8716	0.5292	0.8816
C1a (間)	0.7458	0.7104	0.5653	0.7441	0.9315	0.9449
C3a (バランス)	0.6231	0.1367	0.5136	0.7233	0.8696	0.8629

前項で示した様に Aa (効率) と総合的評価の相関は、評価者の対話への参加・不参加に係わらず強いが、表 3.4 において Aa の詳細指標である A1a、A2a、A3a についてもほぼ全ての被験者において本人の対話への参加・不参加に係わらず総合的評価との相関が強いことが示されている。唯一被験者 C のみ、本人が参加した (a) において A1a (論文検索効率) と A2a (無駄時間) の総合的評価との相関が負になっている。

Ca (リズム) については、総合的評価との相関は、評価者の対話への参加・不参加に依存しているとは言えないという程度であったが、その詳細指標である C1a (間)、C3a (バランス) についてもほぼ全ての被験者において (a)(b) とともに正の強い相関を示してお

り、評価者の対話への参加・不参加に依存しているとは言えない。

従ってここまでの分析から、対話聴取後の評価における対話への参加・不参加の影響に対して、対話についての指標に関しては以下の様にまとめられる。

- 効率に関する指標評価の重要性は、評価者の対話への参加・不参加に依存しない。
- リズムに関する指標評価の重要性は、評価者の対話への参加・不参加に依存するとは言えない。
- 主体性に関する指標評価の重要性は、評価者の対話への参加・不参加に依存する。

(3) 相手についての総合指標: Xb との相関係数

評価時期：対話聴取後 (HE)

評価対象：相手の参加者についての総合指標 (b)

次に、相手の参加者の話し方についての総合的評価と大まかな指標との相関について、被験者によらず相関係数を調べた結果および各被験者毎に相関係数を調べた結果を表 3.5 に示す。なお (a) が評価者自身が参加した対話について、(b) が評価者自身が参加していない対話についての分析結果である。

表 3.5: 相手についての総合指標: Xb との相関係数

(a) 評価者自身が参加している対話

被験者	全体	A	B	C	D	E
Ab (効率)	0.8438	0.8660	1.0000	—	1.0000	1.0000
Bb (透過性)	0.7509	0.5000	0.8660	—	1.0000	-0.5000
Cb (リズム)	0.7175	—	0.8660	—	0.9707	-1.0000

(b) 評価者自身が参加していない対話

被験者	全体	A	B	C	D	E
Ab (効率)	0.6767	0.6364	0.4848	0.6295	0.9159	0.7600
Bb (透過性)	0.5790	0.3133	0.4078	0.2650	0.8941	0.7184
Cb (リズム)	0.5649	0.3015	0.3391	0.5755	0.8237	0.3422

この表において、まず Ab (効率) を見ると、いずれの被験者の評価結果も本人の参加・不参加に係わらず総合的評価と強い相関を示している。

一方、Bb (透過性) については、被験者 A と B の評価結果において本人が参加した対話 (a) よりも本人が参加していない対話 (b) の方が総合的評価との相関が低くなっており、被験者 E の評価結果に至っては本人が参加した対話 (a) では負の相関だったのが本人が参加していない対話 (b) では正の強い相関を示している。このことから、Bb の重要性については評価者の対話への参加・不参加が影響を与えていると言える。

Cb (リズム) についても Bb と同様に、(a) と (b) で総合的評価との相関関係が変わっている被験者が多いことから、評価者の対話への参加・不参加が影響を与えていると言える。

(4) 相手についての詳細指標:Xb との相関係数

評価時期：対話聴取後 (HE)

評価対象：相手の参加者についての詳細指標 (b)

相手の参加者の話し方についての総合的評価と詳細な指標との相関について、被験者によらず相関係数を調べた結果および各被験者毎に相関係数を調べた結果を表 3.10 に示す。なお (a) が評価者自身が参加した対話について、(b) が評価者自身が参加していない対話についての分析結果である。

この表において、A2b (無駄時間) については被験者 A と計測できない C を除いて、評価者自身の参加・不参加に関わらず総合的評価との相関が高い。A3b (発話量) については、被験者 A と被験者 D において評価者自身の参加・不参加により相関係数が大きく異なっている。これは、発話量を考慮する際に、本人の発話量を客観的に判断することが難しいためと考える。

B1b (処理状態) は被験者毎に比較できるサンプルが 2 名分しかないため、結論は出せない。B2b (意思伝達) については、被験者 D を除いて全ての被験者の評価結果において、総合的評価との相関が本人が参加していない対話では弱いのにに対して本人が参加している対話では比較的強い。これは評価対象の対話における意思の状態が、対話の当事者である場合の方が推定しやすいことによると思われる。また、B3b (安心感) についても、若干 (b) の方が (a) より全体に相関が弱くなっている。これも B3b と同様の理由によると考えられる。前項に示した Bb (透過性) の (a)(b) による特性の違いもこの B2b およ

表 3.6: 相手についての総合指標:被験者毎の Xb との相関係数

(a) 評価者自身が参加している対話

被験者	全体	A	B	C	D	E
A2b (無駄時間)	0.6093	0.8660	0.8660	—	0.7559	0.5000
A3b (発話量)	0.5604	1.0000	0.8660	—	0.9449	-0.5000
B1b (処理状態)	0.7596	-0.5000	0.8660	—	1.0000	—
B2b (意思伝達)	0.6062	0.8660	0.8660	—	0.5000	0.8660
B3b (安心感)	0.7911	0.8660	0.8660	—	0.9707	—
C1b (間)	0.8438	0.8660	0.9820	—	0.9707	1.0000
C2b (スピード)	0.4782	0.5000	-0.8660	—	1.0000	0.5000
E1b (表現の多様性)	0.6541	—	0.5000	—	1.0000	1.0000
E2b (話し方の多様性)	0.6912	-0.5000	1.0000	—	1.0000	1.0000
E3b (リズムの多様性)	0.6008	—	0.5000	—	1.0000	1.0000

(b) 評価者自身が参加していない対話

被験者	全体	A	B	C	D	E
A2b (無駄時間)	0.5984	0.0909	0.4915	0.6766	0.8307	0.5023
A3b (発話量)	0.4716	0.0909	0.5497	0.2557	0.5384	0.4961
B1b (処理状態)	0.5561	0.3223	—	0.3105	0.8941	0.6501
B2b (意思伝達)	0.4107	0.3656	0.2651	0.0491	0.8410	0.3133
B3b (安心感)	0.7000	0.6364	0.5513	0.6295	0.8627	0.8008
C1b (間)	0.5343	0.2132	0.4207	0.5704	0.8353	0.1414
C2b (スピード)	0.2155	0.2548	-0.2447	-0.0331	0.4256	0.6000
E1b (表現の多様性)	0.3423	0.1741	0.3052	0.4436	0.4274	0.3333
E2b (話し方の多様性)	0.3295	0.1612	0.3874	0.1299	0.4274	0.3333
E3b (リズムの多様性)	0.3771	0.3015	0.5709	-0.0776	0.4869	0.4648

び B3b の影響による所が大きいと考えられる。

C1b (間) に関しても、総合的評価との相関が、本人が参加している対話よりも本人が参加していない対話では弱くなっている。これは間が対話の参加者両方によって成されるものであるために対話の当事者としての視点がその評価に影響を与えるためと考えられる。一方、C2b (スピード) に関しては、被験者 B を除く全ての被験者の評価結果において (a)(b) とともに総合的評価との弱い相関を示しており、被験者 B の評価結果においては (a)(b) とともに負の相関を示している。これらから C2b の評価に評価者本人の対話への参加・不参加は影響を与えないと言える。

E1b (表現の多様性)、E2b (話し方の多様性)、E3b (リズムの多様性) については、(a) の場合の被験者 A と C の評価結果からの相関係数算出が不能であるため比較対象となるサンプル数が少ないが、全体に (a)(b) とともに同じ傾向を示していると言える。若干 (b)

の方が (a) より相関が弱くなる傾向が見られる。

従ってここまでの分析から、対話聴取後の評価における対話への参加・不参加の影響に対して、相手の話し方についての指標に関しては以下の様にまとめられる。

- 効率の評価には、評価者本人の対話への参加・不参加が影響しないと言える。
- 効率の詳細指標のうち、無駄時間には評価者本人の対話への参加・不参加が影響しないが、発話量については評価者本人の対話への参加・不参加が影響すると言える。
- 透過性評価には、評価者本人の対話への参加・不参加が影響すると言える。
- 透過性詳細指標のうち、処理状態の透過性評価には評価者本人の対話への参加・不参加が影響しないが、意思伝達の透過性および安心感の評価には評価者本人の対話への参加・不参加が影響する。
- リズムの評価には、評価者本人の対話への参加・不参加が影響すると言える。
- リズムの詳細指標のうち、間の評価には評価者本人の対話への参加・不参加が影響し、スピードの評価には評価者本人の対話への参加・不参加が影響しない。
- 多様性評価には、いずれも評価者本人の対話への参加・不参加が影響しない。

(5) 対話聴取評価における評価者の参加・不参加の影響

以上の4項の分析結果から、対話聴取評価において対話の当事者であるか否かが評価に影響を与えない指標は以下のようにまとめられる。

- 対話についての指標
 - Aa (効率) とその詳細指標である A1a (論文検索効率)、A2a (無駄時間)、A3a (発話量)
 - Ca (リズム) とその詳細指標である C1a (間)、C3a (バランス)
- 相手の話し方についての指標
 - Ab (効率)

- A2b (無駄時間)
- C2b (スピード)
- E1b (表現の多様性)、E2b (話し方の多様性)、E3b (リズムの多様性)

一方、対話聴取評価において対話の当事者であるか否かが評価に影響を与える指標は以下の通りである。

- 対話についての指標
 - 主体性 (Da)
- 相手の話し方についての指標
 - A3b (発話量)
 - Bb (透過性) とその詳細指標である B1b (意思伝達)、B3b (安心感)
 - Cb (リズム) とその詳細指標である C1b (間)

なお、B1b (処理状態) に関しては比較できるサンプル数が不足しており上記のような結論を導けない。

以上から、対話についての指標の評価結果を対話後と対話聴取後で比較する場合には、Da (主体性) を除けば全ての指標で全サンプルを用いても良いが、相手の話し方についての指標の半数は、評価者が対話に参加しているか否かを考慮する必要がある。従って以降の対話聴取評価の分析を以下の様に行なう。

1. 対話についての指標に関しては、Da (主体性) を除いて全サンプルを用いる。
2. Da に関しては、評価者が参加していない対話への評価結果のみを用いる。
3. 相手の話し方についての指標に関しては、評価者が参加していない対話への評価結果のみを用いる。

3.3.2 仮説 H_2 の検証：重視する指標の個人差(1) 対話についての総合指標: X_a との相関係数（対話後）

評価時期：対話後 (RE)

評価対象：対話についての総合指標 (a)

まず、対話についての総合的評価と大まかな指標との相関について、被験者によらず相関係数を調べた結果および各被験者毎に相関係数を調べた結果を表 3.7 に示す。なお以降の表において太字は無相関の検定によって 1% の危険率で有意差が認められたものである。

表 3.7: 対話についての総合指標: X_a との相関係数（対話後）

被験者	全体	A	B	C	D	E
Aa (効率)	0.7489	1.0000	0.8660	0.8660	0.0000	0.8660
Ca (リズム)	0.7260	1.0000	0.9449	-0.5000	1.0000	0.5000
Da (主体性)	0.8039	—	0.8660	0.5000	0.8660	1.0000

表 3.7 の全体として示した結果から、いずれの指標も総合的評価との相関が高い指標であると言える。

表 3.7 の被験者ごとの結果は各被験者が各 3 対話のみについて評価した結果から求めているためやや信頼性に欠けるため、以降の分析結果はあくまでも目安という位置にとどめておきたい。

まず、被験者 A, B, C, E においては総合評価 X_a との相関が高い Aa の指標（効率）が、被験者 D においては無相関である。また、指標 Ca（リズム）については被験者 A, B, D においては総合評価 X_a との相関が高いのが、被験者 E ではそれほど相関が高くなく、被験者 C に至っては負の相関が見られている。指標 Da（主体性）についても、被験者 B, D, E においては相関が高いが、被験者 C においてはそれほどでもない（被験者 A は Da の評価が全て等しかったために相関係数の算出が不可能）。

以上のことから、対話における効率、リズム、主体性についての評価は総合的評価との相関が全体に高いが、必ずしも全ての被験者において全ての指標に同様の傾向が見られるわけではなく、3 つの指標のうちのいずれかは総合的評価との相関が低いものも存在すると言える。

(2) 対話についての詳細指標: X_a との相関係数 (対話後)

評価時期: 対話後 (RE)

評価対象: 対話についての詳細指標 (a)

ここでは前項に示した総合指標の下位分類とした詳細指標と総合的評価との相関関係を分析する。

まず、対話についての総合的評価と詳細な指標との相関について、被験者によらず相関係数を調べた結果および各被験者毎に相関係数を調べた結果を表 3.8 に示す。

表 3.8: 対話についての詳細指標: X_a との相関係数 (対話後)

被験者	全体	A	B	C	D	E
A1a (論文検索効率)	0.5026	1.0000	0.7559	0.5000	-0.3273	—
A2a (無駄時間)	0.6516	0.9449	0.7559	1.0000	0.5000	0.7559
A3a (発話量)	0.4987	0.5000	0.8660	-0.8660	0.3273	—
C1a (間)	0.8577	1.0000	1.0000	0.5000	0.8660	0.5000
C3a (バランス)	0.4287	0.5000	0.5000	0.5000	0.8660	-1.0000

前節では総合評価との相関が高いとした A_a (効率) の詳細指標である A1a (論文検索効率)、A2a (無駄時間)、A3a (発話量) について、ほとんどの被験者は総合評価との相関が高いのだが、A1a において被験者 D が負の低い相関、A3a において被験者 A と D が低い相関、被験者 C が負の相関を示している様に、総合評価に与える各指標の影響は被験者ごとに若干異なる。

また、前項の分析では、総合評価 (X_a) との高い相関が見られた指標 C_a (リズム) の詳細指標である C1a (間)、C3a (バランス) について、表 3.8 においては、C1a は被験者ごとにも被験者によらなくても総合評価との相関が高いが、C3a はいずれも相関が低い (被験者 E に至っては負の相関が高い)。これは、C1a (間) は C_a (リズム) の詳細指標として正しく機能し得るが、C3a (バランス) はそうではないことを示すと考える。これら指標の関係について詳しくは次節の仮説 H_2 の検証において述べる。

これらをまとめると、詳細指標についてさらに分類を見直す必要があるものの、対話の総合評価との相関関係について以下の結論が導ける。

- 一般に重要とされているタスク効率について、その主観的評価の貢献度は被験者により異なる。

- 無駄時間の量についての主観的評価は比較的どの被験者においても貢献度が高い。
- 発話量についての主観的評価はどの被験者においても貢献度が低い。
- 間についての主観的評価はリズムの詳細指標として正しく機能し、被験者によらず総合的評価への貢献度が高い。
- バランスについての主観的評価はリズムの詳細指標として正しく機能しない。

(3) 相手についての総合指標: Xb との相関係数 (対話後)

評価時期：対話後 (RE)

評価対象：相手の参加者についての総合指標 (b)

次に、相手の参加者の話し方についての総合的評価と大まかな指標との相関について、被験者によらず相関係数を調べた結果および各被験者毎に相関係数を調べた結果を表 3.9 に示す。

表 3.9: 相手についての総合指標: Xb との相関係数 (対話後)

被験者	全体	A	B	C	D	E
Ab (効率)	0.5608	1.0000	0.5000	—	0.7559	-1.0000
Bb (透過性)	0.4372	1.0000	-0.5000	—	0.8660	—
Cb (リズム)	0.6736	1.0000	0.7559	—	1.0000	—

表 3.9 に示した結果から、3.3.2 項で示した対話についての評価の場合とは異なり、いずれの指標も総合的評価との相関が低いと言える。

しかし、表 3.9 の被験者毎の結果について見れば、被験者 A や被験者 D の様に全ての指標において総合的評価 (Xa) との相関が高いケースがある様に、被験者によっては相手の参加者の話し方についても各指標と総合的評価の相関が高くなる。この違いは、個人ごとに重視する指標が異なるという理由以上に、対話についての評価においては各指標あるいはその質問文の解釈において被験者ごとの差が大きくないのに対して、相手の話し方の評価においては各指標あるいはその質問文の解釈が被験者ごとに異なることが主な原因と考えられる。このことは、各指標を詳細にした詳細指標について同様の相関分析を行な

うことで明らかにできる。次項に相手の話し方についての詳細指標と総合的評価の相関関係について分析する。

なお、Cb (リズム) については被験者によらず総合的評価との相関が強い。

(4) 相手についての詳細指標:Xb との相関係数 (対話後)

評価時期：対話後 (RE)

評価対象：相手の参加者についての詳細指標 (b)

相手の参加者の話し方についての総合的評価と詳細な指標との相関について、被験者によらず相関係数を調べた結果および各被験者毎に相関係数を調べた結果を表 3.10 に示す。

表 3.10: 相手についての総合指標:Xb との相関係数 (対話後)

被験者	全体	A	B	C	D	E
A2b (無駄時間)	0.6591	0.5000	0.7559	—	0.7559	—
A3b (発話量)	0.6622	0.5000	0.8660	—	0.7559	—
B1b (処理状態)	0.3866	—	-0.5000	—	1.0000	0.8660
B2b (意思伝達)	0.4422	1.0000	—	—	0.5000	0.5000
B3b (安心感)	0.7399	0.8660	1.0000	—	1.0000	—
C1b (間)	0.8925	1.0000	1.0000	—	1.0000	—
C2b (スピード)	0.4959	—	—	—	1.0000	0.5000
E1b (表現の多様性)	0.2774	-0.8660	—	—	—	0.5000
E2b (話し方の多様性)	0.2735	-1.0000	—	—	—	—
E3b (リズムの多様性)	0.3942	-1.0000	1.0000	—	1.0000	—

いずれの詳細指標も被験者によらず総合的評価と正の相関を持つが、特に B3b (安心感)、C1b (間) の影響力が大きいことがわかる。

ここで、前節に述べた様に、Cb (リズム) と総合的評価の相関が被験者によらず大きい (表 3.9 参照)、その詳細指標である C1b と総合的評価の相関についても同様のことが言える。それに対して、総合指標である Bb (透過性) については総合的評価とそれほど強い相関が見られなかったのが、その下位分類とした詳細指標 B3b においては強い相関が見られることから、B3b が Bb の下位分類として不適切である可能性がある。次節において、指標間の相関を調べることによってさらにこれを明らかにする。

なお、表 3.10 に示した結果より、相手の参加者についての詳細な指標と総合的評価の相関について、以下のことが言える。

- 効率の詳細指標と総合的評価の相関は、被験者によらず同じ程度に高い。
- 透過性の詳細指標について、処理状態の透過性評価と総合的評価との相関関係には大きな個人差が見られる。
- 意思伝達の透過性評価と総合的評価の相関については、被験者間において同じ程度に相関が見られる。
- 安心感の主観的評価と総合的評価の相関は、被験者によらず透過性の総合的評価以上に強い。
- 間の主観的評価についても、被験者によらず総合的評価との相関が強い。
- スピードの主観的評価と総合的評価の相関には個人差が見られる。
- 多様性の詳細な指標と総合的評価の相関に関してはいずれも大きな個人差が見られる。

以下には、これまでの 4 項と同様の分析を対話聴取後の評価結果について行なう。

(5) 対話についての総合指標: X_a との相関係数 (対話聴取後)

評価時期：対話聴取後 (HE)

評価対象：対話についての総合指標 (a)

まず、対話聴取後に行なった対話についての総合的評価と大まかな指標評価との相関について、被験者によらず相関係数を調べた結果および各被験者毎に相関係数を調べた結果を表 3.11 に示す。なお、3.3.1 節に示した様に、対話聴取後の評価結果として、評価者自身の対話への参加の影響を考慮して、 A_a 、 C_a は全対話への評価結果、 D_a は評価者自身が参加していない対話への評価結果を用いた。

表 3.11 の全体として示した結果から、いずれの指標も総合的評価との相関が高い指標であると言える。

被験者ごとに見ると、 A_a (効率)、 C_a (リズム) とともに全被験者において総合的評価 X_a との相関が強い。一方、 D_a (主体性) については被験者毎に総合的評価 X_a との相関関係がかなり異なる。

表 3.11: 対話についての総合指標: Xa との相関係数 (対話聴取後)

被験者	全体	A	B	C	D	E
Aa (効率)	0.7898	0.8362	0.8028	0.7302	0.7225	0.9419
Ca (リズム)	0.7483	0.7828	0.6244	0.6661	0.9026	0.9511
Da (主体性)	0.4209	0.2791	0.0251	0.6401	0.7255	0.9047

これらから、聴取評価においては、対話における効率、リズムについての評価は常に重視されるが、主体性については重視する被験者とそうでない被験者がいると言える。

(6) 対話についての詳細指標: Xa との相関係数 (対話聴取後)

評価時期：対話後 (HE)

評価対象：対話についての詳細指標 (a)

ここでは前項に示した総合指標の下位分類とした詳細指標と総合的評価との相関関係を分析する。

まず、対話についての総合的評価と詳細な指標との相関について、被験者によらず相関係数を調べた結果および各被験者毎に相関係数を調べた結果を表 3.12 に示す。なお、3.3.1 節に示した様に、対話聴取後の評価結果として、これらの指標はいずれも評価者自身の対話への参加の影響を受けないため、全対話への評価結果を分析に用いた。

表 3.12: 対話についての詳細指標: Xa との相関係数 (対話聴取後)

被験者	全体	A	B	C	D	E
A1a (論文検索効率)	0.7131	0.8497	0.6619	0.5739	0.8736	0.9542
A2a (無駄時間)	0.7711	0.4919	0.7750	0.6752	0.9066	0.8892
A3a (発話量)	0.7359	0.8361	0.7144	0.8463	0.4630	0.9010
C1a (間)	0.7447	0.6841	0.5979	0.6953	0.9398	0.9480
C3a (バランス)	0.6449	0.4945	0.4181	0.7627	0.8405	0.8834

この結果において、いずれの指標も被験者によらず正の比較的強い相関を示していると言える。しかし A2a (無駄時間) の指標において被験者 A が、A3a (発話量) の指標において被験者 D が、若干弱い相関を評価結果において示していることから、これらの被験者はそれぞれの指標を重視していないと言える。また、C3a (バランス) については、被験者 C、D、E の結果が比較的強い相関を示しているのに対して、被験者 A、B の結果は

弱い相関を示している。この指標を重視するか否かも被験者により異なると考えられる。

以上の考察をまとめると、対話聴取評価時に対話についての指標について評価者が重視する指標の個人差に関して以下の結論が導ける。

- 聴取評価においては、対話における効率、リズムについての評価は重視されるが、主体性については重視する被験者とそうでない被験者がいると言える。
- 効率の詳細指標に関して、無駄時間、発話量については重視する被験者とそうでない被験者がいる。
- リズムの詳細指標に関して、バランスについては重視する被験者とそうでない被験者がいると言える。

(7) 相手についての総合指標: Xb との相関係数 (対話聴取後)

評価時期：対話後 (HE)

評価対象：相手の参加者についての総合指標 (b)

次に、相手の参加者の話し方についての総合的評価と大まかな指標との相関について、被験者によらず相関係数を調べた結果および各被験者毎に相関係数を調べた結果を表 3.13 に示す。なお、3.3.1 節に示した様に、対話聴取後の評価指標の多くは評価者自身の対話への参加の影響を受けるため、評価者自身が参加していない対話への評価結果を分析の対象とした。

表 3.13: 相手についての総合指標: Xb との相関係数 (対話聴取後)

被験者	全体	A	B	C	D	E
Ab (効率)	0.6767	0.6364	0.4848	0.6295	0.9159	0.7600
Bb (透過性)	0.5790	0.3133	0.4078	0.2650	0.8941	0.7184
Cb (リズム)	0.5049	0.3015	0.3391	0.5755	0.8237	0.3422

表 3.13 に示した結果において、Ab (効率) はいずれの被験者においても総合的評価との強い相関を示しているが、他の指標に関しては全体としては総合的評価との相関が強いが、各被験者を見ると被験者 C の Bb (透過性) や被験者 A の Cb (リズム) の様に相関の低い結果も見られる。