

人対人の会話で自然な話題展開を 支援するための対話戦略の検討

前園 そよぎ^{1,a)} 原 直^{2,b)} 阿部 匡伸^{2,c)}

概要：本稿では、システムが人対人の会話に適度に介入して、自然な話題展開を促す対話戦略を検討する。提案する対話戦略では、システムが予め知っている話題の情報を、全て用いるのではなく、敢えて話題の情報を一部隠して発話することとし、会話している人同士に会話の展開を委ねる。本稿ではこれを誘発対話と呼ぶ。誘発対話により、自然に次の話題へと展開することを期待する。一方、話題の情報を全て提供する対話を誘導対話と呼ぶ。評価実験では写真に関する会話を話題と捉え、写真に関する情報を調整することで、話題展開を支援する実験を行った。システムの誘発対話と誘導対話は用意された写真を基に人手で作成した。主観評価実験の結果、誘発対話は話し手がその話題を鮮明に覚えている場合に有効であり、誘導対話は話し手がその話題をあまり覚えていない場合に有効であることが明らかとなった。

Investigation of dialogue strategies to support natural topic development in person-to-person conversations

1. はじめに

人と人が会話をする際、会話の継続が途切れる場合がある。会話が途切れた際、何らかの理由で、相手と会話を続けられない、又は続けたくない場合、会話そのものを終了すれば良い。相手と会話を続けたい場合、新たな話題を提案する、又は現在の話題を更に詳細に進めていく等を行う必要がある。しかし、話す話題が思いつかない、又は、話す話題が尽きたと感じている時、会話の継続は困難になる。このような場合に、会話支援システムが役に立つ。システムが、会話の継続のために取りうる戦略は、話題候補を提示する戦略 [1][2]、話題に対して積極的に提案する戦略 [3]、話題に対する無難な質問をする戦略 [4] 等が考えられる。

本報告では、自然な話題展開をして、1対1の会話を継続させる会話支援システムの対話戦略を検討する。ここで「自然な話題展開」とは、話し手に次の話題を迷うことなく展開させ、かつ、次の話題を盛り上げるような展開をすることである。

会話の継続のために取りうる戦略として、本報告では「誘発対話」という対話戦略を提案する。誘発対話は、システムは話題を示唆しながらも、話題の展開自体はユーザに委ねることで話題が盛り上がることを期待している。

本報告では、提案戦略の有効性を検証するために、話し手の旅行の写真を用いた実験システムを実装した。この実験システムでは、1つの写真を1つの話題と考えている。本システムの利用者は、旅行中の様々な出来事を反映した写真を様々に切り替えながら、その写真にまつわる話をする。ただし、写真の切り替え、つまり、話題の展開そのものは会話者らに委ねることとし、本システムは、写真のメタ情報から取捨選択したテキスト文を、次の話題を探す手がかりとして提示するのみとしている。

本報告では、誘発対話の有効性を検討するために、話題に対して積極的に提案する戦略(誘導対話)と話題に対する無難な質問をする戦略という2つの対話戦略との比較実験を行う。評価の観点としては、次の話題へと自然に話題展開できたか等の話題展開の支援に関する観点と、思い出振り返り、雑談などの状況を想定して、各システムが有効となる場面に関する観点の2つを評価する。

本報告は、以下の通りの構成である。第2章では、従来対話について述べる。第3章では、提案戦略である誘発対

¹ 岡山大学 大学院ヘルスシステム統合科学研究科

² 岡山大学 学術研究院ヘルスシステム統合科学学域

a) s_maezono@s.okayama-u.ac.jp

b) hara@okayama-u.ac.jp

c) abe-m@okayama-u.ac.jp

表 1: 従来の対話方法

	誘導対話	傾聴対話
概要	情報を用いてユーザを誘導する対話	ユーザの話の聞き役となる対話
利点	次の話題に迷わない	自由に話せる
欠点	ユーザの話を奪う恐れ	新しい話題を展開することが難しい

話について述べる。第4章では、写真を用いた実験システムについて述べる。第5章では、評価実験とその結果について述べる。最後に第6章で本論文のまとめを述べる。

2. 従来対話

誘導対話 [3] は、商品販売や勧誘、広告に使われており、ユーザをある目的に誘導する対話である。人対人の会話支援においても、誘導対話を用いると、ユーザの次の話題を迷うことなく誘導することができる。しかし、誘導対話は、ユーザの意図とは異なる話題に展開してしまう可能性や、誘導内容によっては、ユーザの話すことがなくなってしまう可能性がある。

また、傾聴対話 [4] は相槌などにより聞いていることを示す対話であり、話し相手としてユーザの話を聞いてほしいという欲求を満たすことが期待されている。人対人の雑談支援において、傾聴対話を用いると、話し手は話したいことを自由に話せる。しかし、システムはあくまで聞き役の為、次の話題に導くことは難しい。よって、ユーザ自身が次の話題を見つけなければいけないため、ユーザが次の話題に迷い、話題展開がうまくできず、気まずさが生まれる恐れがある。

2つの対話方法の概要と、その対話を会話支援に用いた時の利点と欠点を表1にまとめる。

以上を踏まえて、本報告では、会話支援システムのための対話方法として誘導対話と傾聴対話それぞれのメリットを両立しうる新たな対話戦略の研究を行う。この新しい対話戦略では、ユーザが迷うことなく話題を展開し、かつ、ユーザの話題を盛り上げる戦略となることが望ましい。

3. 誘発対話戦略

3.1 自然な話題展開のための対話戦略

会話において重要な要素の1つは、「自然な話題展開」であると考えられる。自然な話題展開とは、本人たちも気づかない間にいつの間にか話題が変わることである。例えば仲のいい人同士が話すときに、ある話題で出てきた話題から少しずつ話題が変化し、気づかない間に全く違う話に変わっていて「なんでこんな話になったっけ」と感じることもある。この現象は、話し手の情報がある程度知っている聞き手が、話し手と関連のある別の話題を、無意識に質問することによって徐々に話題が変わっていると推察される。

そこで、本報告では、話し手の情報を知っているシステムが話し手の話題と関連のある質問をする対話戦略を検討する。この対話戦略によって、話し手に次の話題へと迷う

写真			
時系列	10:30(写真1)	12:00(写真2)	14:00(写真3)
場所	車の中	アナゴ屋	姫路城
状況	祭りをみた	お昼ご飯を食べた	天守閣に登った

図 1: 写真とメタ情報の例

表 2: システムとメタ情報

システム	メタ情報	質問内容例
lead	「写真の時系列」 「写真の場所」 「写真の状況」 全て	その後、 <u>アナゴ屋</u> で <u>昼ご飯を食べた</u> んですね。
inspire	「写真の時系列」と「写真の場所」 「写真の時系列」と「写真の状況」	この後、どこで <u>昼ごはん</u> を <u>食べた</u> んですか？
generic	なし	他に何かしましたか？

ことなく展開させ、話を盛り上げることが期待できる。

具体的には、1つの話題に複数の情報が含まれていると仮定し、システムはその話題に含まれている情報から全て話すのではなく、一部を発話する方式である。このような対話戦略を、本報告では、誘発対話と呼ぶことにする。誘発対話により、ユーザが話したい話を全てシステムが話してしまうことなく、ユーザへの話題提供ができることが期待できる。

3.2 話題対象のメタ情報に応じた対話戦略の比較

本報告では、誘発対話が有効と考えられる状況として、本システムの利用者が、旅行中の思い出を写真とともに振り返りながら、友人と雑談をするような場面を想定して、検討を進める。また、1つの写真に関する会話を1つの話題と見立てて、システムは話題展開を行う。本システムでは、写真のメタ情報を活用しながら、提案戦略の対話方法や、比較対象の対話戦略を実現する。1つの写真につき、図1のような3つのメタ情報がある。このメタ情報の組み合わせと多寡により、誘導対話と誘発対話、戦略なしの3種類を作成する。本報告では、3つの対話戦略を人手で作成する。

それぞれの対話戦略とメタ情報、質問内容の例は表2に示す。質問例の下線部はメタ情報である。誘導対話には、メタ情報全てを用いた文章を発話する。誘発対話は、状況か場所のどちらかを除いた発話をする。戦略なし対話は、メタ情報を使わず発話する。

3.3 対話例

従来の対話方法と提案戦略を比較するために、誘導対話、誘発対話、戦略なしの3つのシステムを比較する。それぞれの対話方法の例を3.3.1項～3.3.3項に示す。いずれの例も冒頭では図1の写真1が表示されている。

3.3.1 誘導対話を用いたシステム

話し手 姫路に行く途中で祭りを見たんだ。

聞き手 へえー、とても楽しそう。

システム その後、レストランで昼ごはんをたべたんですよ。(写真2提案)

話し手 そうそう、とてもおいしかったよ。(写真2提示)

3.3.2 誘発対話を用いたシステム

話し手 姫路に行く途中で祭りを見たんだ。

聞き手 へえー、とても楽しそう。

システム その後、どこで昼ごはんをたべたんですか？
(写真2提案)

話し手 姫路城の近くにあるアナゴ屋さんに行ってきたよ。
とてもおいしかった！(写真2提示)

3.3.3 誘導や誘発を行わないシステム

話し手 姫路に行く途中で祭りを見たんだ。

聞き手 へえー、とても楽しそう。

システム 他には何をしましたか？

話し手 えっと、姫路城に行ったよ(写真3提示)

4. 実験システム

4.1 実験システムの概要

本報告では、話題を旅行に限定して会話支援システムを実装する。話題となる旅行には、話し手のみが参加しており、聞き手は参加していない状況を想定している。この会話支援システムは、話し手と聞き手の1対1の会話において、話し手の写真を材料に、話し手と聞き手の自然な話題展開を支援する。

そして、システムは適切なタイミングでテキストメッセージによる促しをすることで、話し手による写真の切り替え、すなわち話題の展開を促す。なお、本報告では、話題展開を促すタイミングは、話し手や聞き手とは別の、操作者が判断することとした。

システムを使った実験を行う前段階として、人対人での旅行の話題を話してもらい、複数写真をどの順番で話すのか調査した。実験参加者の話し手は全て時系列順に話していたことから、実験システムは、写真を時系列で提案する。

4.2 実験システムの構成

まず、話し手が操作する部分のシステム構成の説明を行う。開発にはPython (ver. 3.7.7)のWebフレームワークTornado (ver. 6.0.4)とJavaScriptを使った。システムは2つのウィンドウを持つ。写真閲覧用の簡易アルバムアプリ(図2)では、話し手はタッチ操作を行い、自分が話したい写真をタッチすると、画像が拡大して表示される。会話支援システムのウィンドウ(図3)では、システムが次の話題の提案を吹き出しの中にテキストメッセージで表示している。次の話題の提案を音声ではなく、文字で表示する理由は、システムが音声で発話することによって、ユーザの

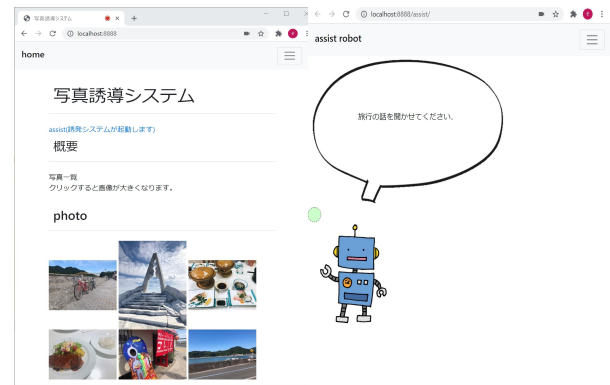


図2: 写真閲覧用の簡易アルバム アプリ

会話を中断させてしまうことを防ぐ為である。

続いて、会話支援システム用のウィンドウ(図3)について説明する。このウィンドウでは、システムが判断したタイミングで、次の話題を適切なテキスト形式で提供する。会話中、話し手や聞き手は、簡易アルバムアプリだけでなく、このウィンドウも見ている。ただし、前述のとおり、本報告では、その話題提供のタイミングは話し手や聞き手とは別の操作者が行うこととしている。音声認識には、Google ChromeのSpeech Web API[5]を使う。本来は、写真のメタ情報とは別に、ユーザの音声を認識し、その結果も使った会話支援を行うことが理想だが、本報告ではユーザの会話を記録するために音声認識を行っている。

最後に、本システムで記録する実験ログデータについて説明する。ユーザの音声が入力された時、音声認識を行い、会話内容、システムが誘導したい写真、システムの発話内容、話し手が選んだ写真、誘導の成功・失敗の5点をファイルへと記録している。

5. 評価実験

5.1 実験の概要

本実験では、誘導対話、誘発対話、戦略なしの3つの対話戦略を用いたシステムを比較し、評価する。実験参加者は、20歳代の大学生及び、大学院生で構成される6組12名である。実験は、話し手と聞き手の2人組で行い、主に話し手がシステムの操作を行う。評価は話し手と聞き手のそれぞれにアンケートに答えてもらい、その主観評価をもとに提案戦略の評価を行う。

5.2 実験手順

話し手にゼミ合宿の写真を15枚用意してもらい、写真のメタ情報を15枚すべてに付与してもらった。システムの発話文は、このメタ情報を用いて、あらかじめ人手で用意する。実験参加者は、誘導対話、誘発対話、戦略なしの合計3つのシステムを使用する。使用するシステムの順番による影響を避けるために、話し手と聞き手のペアごとに

表 3: 話し手事後アンケート

番号	項目
Q1	システムに誘導されているように感じた
Q2	自分で話したい内容をシステムに言われた
Q3	システムの発話で次にしたことを思い出した
Q4	自然に次の話題へと進めた
Q5	次にどの話をするか迷った
Q6	システムの質問で混乱した
Q7	システムの発話内容は自然だった
Q8	システムの質問により次の写真の話題が盛り上がった

使用するシステムの順番はランダムとする。最初に話し手と聞き手は、事前アンケートに答える。1つのシステムにつき、話し手は5枚の写真を話す。1つのシステムの使用が終了した後、話し手と聞き手はそれぞれ主観評価を行う。計3つのシステム、合計15枚の写真の話が終わったら、最後に全体アンケートに答える。

5.3 評価項目

主観評価の評価項目として、事前アンケート、事後アンケート、全体アンケートの3つのアンケートを行った。これらのアンケートでは、主にシステムの対話部分に着目して、話し手と聞き手がどう感じたかについての項目を用意した。以下に、3つのアンケートについて説明する。

事前アンケートは、評価実験を行う前のアンケートである。自分について、話す方が好きか、聞く方が好きかについて調査し、また、ペアとの親密度について調査を行った。

事後アンケートは、1つのシステムの実験が終わる毎に評価される。さらに、各システムの写真の切り替えごとに表3、表4の評価項目に答えてもらった。また、評価項目と対応している番号について、Q*は話し手側の評価項目でq*は聞き手側の評価項目である。話し手と聞き手は、表3、表4の各科目について「全くそうは思わない」、「そうは思わない」、「ややそうは思わない」、「どちらでもない」、「ややそう思う」、「そう思う」、「とてもそう思う」という7段階のリッカート尺度で評価する。結果の分析の際には、「全くそうは思わない」から「とてもそう思う」に評定値としてそれぞれ1から7の数字を割り当て、間隔尺度として分析する。また、話し手は、各システムで話した写真を覚えていたかを「鮮明に覚えていた」「覚えていた」「あまり覚えていなかった」「覚えていなかった」の4段階で評価する。

全体アンケートは、3つのシステムの実験全てが終了した後のアンケートである。表5、表6の評価項目において、話し手と聞き手は、各項目に対して最も当てはまるシステムを1つ選択する。

5.4 実験環境

表7に主な使用機器、図4に機器の接続とレイアウトを示す。実験中の話し手と聞き手の様子を記録するために個

表 4: 聞き手事後アンケート

番号	項目
q1	システムの発話の方が話し手より次の話題の情報量が多かった
q2	話し手の発話の方がシステムより次の話題の情報量が多かった
q3	システムではなく話し手からもっと話を聞きたかった
q4	話し手から十分に話を聞いた
q5	話し手が質問に対して答えづらそうだった
q6	話し手が質問に対して答えやすそうだった
q7	システムの発話内容は自然だった
q8	次の写真の話題は盛り上がったと思う
q9	会話がスムーズに展開されたと思う

表 5: 話し手全体アンケート

番号	項目
Q9	次の写真の話題へと自然に話題展開できたシステム
Q10	システムの質問内容の返答に困ったシステム
Q11	誘導されたと思うシステム
Q12	次の写真の話題が一番盛り上がったシステム
Q13	思い出を振り返るのに一番良いシステム
Q14	雑談するのに一番良いシステム
Q15	初対面の人に使うと一番良いシステム

表 6: 聞き手全体アンケート

番号	項目
q10	話し手が困惑していたシステム
q11	話し手の写真の話題が盛り上がったシステム
q12	話し手が答えやすそうだったシステム
q13	思い出を振り返るのに一番良いシステム
q14	雑談するのに一番良いシステム
q15	初対面の人に使うと一番良いシステム

表 7: 主な使用機器

名称	品名
個人用マイク	SHURE MX183
全体カメラ	SONY HDR-CX180,680
個人カメラ	Zoom Q2n
オーディオインタフェース	steinberg UR44C
システム	VAIO VJPB11C11N
タッチパネル	GeChic On-Lap1503

人カメラと全体カメラにより動画を撮影する。実験風景を図5に示す。個人カメラは話し手と聞き手の上半身、全体カメラは実験の様子全体を前から撮影する。また、二人が今話している写真が見えるように、後方の全体カメラで撮影する。また、操作者は話し手と聞き手からは見えないようにするために、仕切りを入れる。

5.5 実験結果

実験中の話し手と聞き手の会話とシステムの作動や発話は、概ねスムーズに進行した。ただし、6組計72回、次の写真の提案を行っているが、その内の2回はシステムのバグにより時系列通りに写真の提案ができなかった。この部

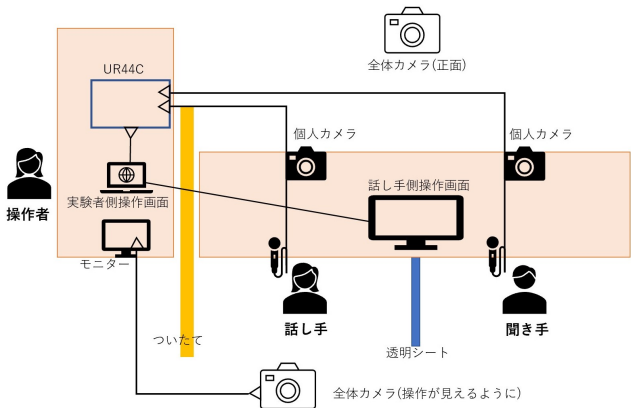


図 4: 使用機器のレイアウト

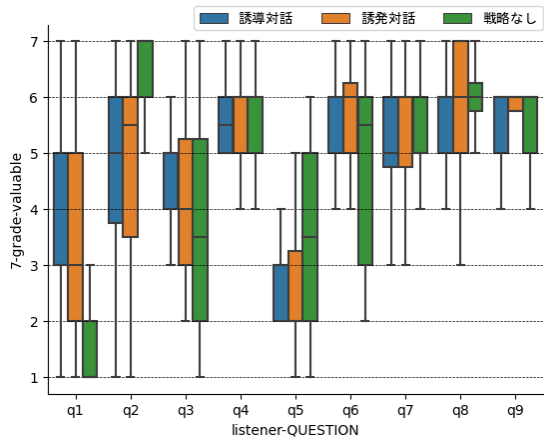


図 7: 聞き手の評価結果



(a) 前面

(b) 背面

図 5: 実験風景

表 8: 全体アンケート結果 (話し手)

項目	誘導対話	誘発対話	戦略なし
Q9	2	4	0
Q10	1	1	4
Q11	4	2	0
Q12	3	2	1
Q13	4	2	0
Q14	0	6	0
Q15	3	2	1

表 9: 写真の出来事に関する記憶の程度。回答は 1:鮮明に覚えていた, 2:覚えていた, 3:あまり覚えていなかった, 4:覚えていなかった, のいずれかである。

	A	B	C	D	E	F
誘導対話利用時の写真	2	3	2	1	3	2
誘発対話利用時の写真	1	2	2	1	2	2
戦略なし利用時の写真	1	3	3	2	3	2

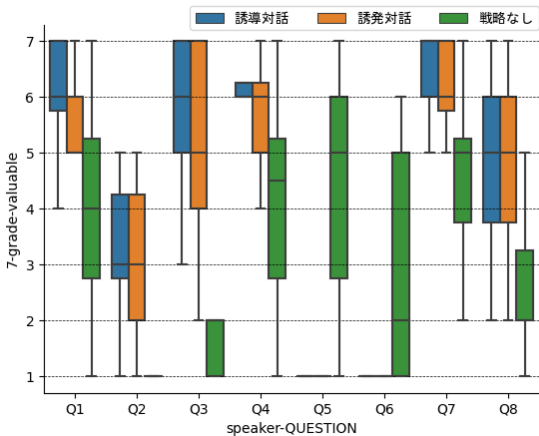


図 6: 話し手の評価結果

分については結果から除外している。

話し手と聞き手のペア 6 組 12 名による、誘導対話、誘発対話、戦略なしの結果を箱ひげ図に示す。まず、話し手による Q1~Q8 の評価を図 6 に、聞き手による q1~q9 の評価を図 7 に箱ひげ図として示す。それぞれの評価項目は、表 3、表 4 に示した項目である。縦軸は 7 段階のリッカート尺度となっている。次に、話し手の全体アンケートの結果を表 8 に示す。最後に、システムごとに話した写真についてその写真を覚えていたかどうかの結果を表 9 に示す。ここで、A、B、C、D、E、F は話し手である。

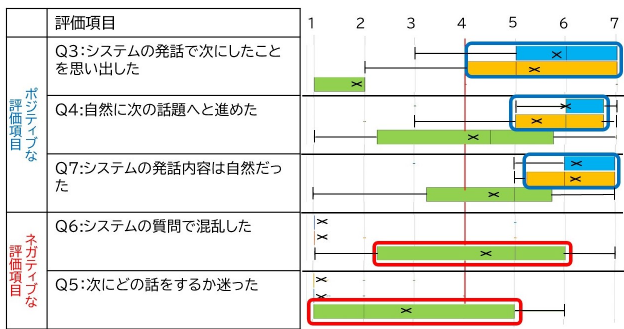


図 8: 各評価項目と箱ひげ図

5.6 考察

5.6.1 話し手による各システムの評価の比較

図 6 で示した評価結果のうち、特に Q3~Q7 の 5 つの評価項目に注目して、さらにポジティブな評価項目とネガティブな評価項目に分けて考察を行う。図 8 に図 6 から項目を抜粋した箱ひげ図を示す。青色は誘導対話、オレンジ色は誘発対話、緑色は戦略なしの対話である。Q3、Q4、

表 10: 評価項目×写真の記憶の程度

項目	システム	自然に話題展開できたシステム			次の話題が盛り上がったシステム		
		誘導対話	誘発対話	戦略なし	誘導対話	誘発対話	戦略なし
写真	鮮明に覚えている	0	2	0	0	2	0
	覚えている	0	2	0	1	0	1
	あまり覚えていない	2	0	0	2	0	0

Q7 のポジティブな評価項目に関しては、誘導対話と誘発対話が良い評価を示している。逆に、Q6, Q5 のネガティブな評価項目では、戦略なしが大きくばらついていることが分かる。よって、誘導対話と誘発対話は、いずれも戦略なしに比べて優れた対話戦略であるといえる。

5.6.2 話し手と聞き手の評価の違い

話し手と聞き手への共通の評価項目である「システムの発話内容は自然だった (Q7, q7)」と「(システムの質問により) 次の写真の話題が盛り上がった (Q8, q8)」に注目する。図 6 における Q7 と Q8 の回答の分布を見ると、話し手に関しては誘導対話、誘発対話に対して、戦略なしは大きく異なる評価値を回答している。一方、図 7 における q7 と q8 の回答の分布を見ると、聞き手に関しては 3 つの戦略は、いずれも同程度の評価になった。つまり、これら 3 つの対話戦略は、聞き手にとってはいずれも、発話の自然性や話題の盛り上がりの観点での、明確な違いを感じることはないが、話し手としての話しやすさに違いが表れている。

5.6.3 最も高い評価を得たシステム

写真の出来事の記憶に応じた、システムの有効性の違いについて、考察を行う。最も高い評価を得たシステム (表 8) と写真の内容に関する記憶の程度 (表 9) のクロス集計結果を表 10 に示す^{*1}。

自然に話題展開できたシステムでは、鮮明に覚えている人や覚えている人が誘発対話を選んだのに対し、写真をあまり覚えていない人が誘導対話を選んだことが分かる。次の話が盛り上がったシステムでは、誘発対話を選んだ人が 2 人とも、写真を鮮明に覚えていると回答したのに対し、誘導対話を選んだ人は写真をあまり覚えていない人が 2 名、覚えている人が 1 名という結果になり、鮮明に覚えている人はいないということが分かる。つまり、その話題が印象深いほど、提案戦略の誘発対話が有効であることがわかる。

5.6.4 状況別最も高い評価を得たシステム

さらに、全体アンケートとして回答を得た、ある状況を想定した場合の、最も高い評価を得たシステムについても、考察を述べる。

表 8 で示したように、思い出の振り返り用途 (Q13) では、誘導対話が高評価となった。これは、思い出の振り返りだけが目的であれば、誘導された方が、話し忘れることなく便利と感じられるからだと考える。雑談用途 (Q14) では、誘発対話が高評価となった。これは、メタ情報を一

^{*1} 写真を覚えていなかったと回答した人はいなかった為、表からは省略している。

表 11: Q12 と Q15 のユーザ別アンケート結果

	A	B	C	D	E	F
Q12:盛り上がったシステム	誘発対話	誘導対話	誘導対話	誘発対話	誘導対話	戦略なし
Q15:初対面に良いシステム	誘発対話	誘導対話	誘導対話	誘発対話	誘導対話	戦略なし

部隠すことにより、システムが第 3 者として会話に混ざっているように感じられるからだと考える。

最後に、初対面の人に使うと良いシステム (Q15) は人によって評価が分かれた。本実験においては、全ての話し手に対して、Q15 の選択と、Q12 の一番盛り上がったシステムの選択が、常に同じ結果であった。5.6.3 項で考察したように、盛り上がったシステムは写真をどれぐらい覚えているかによって誘導対話と誘発対話を選ぶ人が分かれた。つまり、写真をどのくらい覚えているかによって、初対面の人に使うと良いシステムの対話戦略を変える必要性が示唆される。しかし、話し手 F のように、戦略なしのシステムを選択する人もいたため、初対面の人に使うと良いシステムについては更なる検証が必要である。

6. おわりに

本稿では、会話支援システムの対話戦略として誘発対話を提案した。誘発対話は、システムは話題を示唆しながらも、話題の展開自体はユーザに委ねることで話題が盛り上がることを期待している。誘発対話を評価するために、6 組 12 名の実験参加者による主観評価実験を行った。各実験参加者は、誘発対話、誘導対話、戦略なし対話の 3 つの対話を使ったシステムを比較し、評価した。実験結果より、話題に対して強い思い入れがある程、誘発対話が有効であることが示された。また、思い出を振り返る場面では誘導対話、雑談する場面では誘発対話の評価が高いという結果が得られた。今後の課題として、メタ情報の抽出方法、メタ情報の追加、発話タイミングの自動化等があげられる。

参考文献

- [1] 天野健太, 西本一志: 六の膳:お皿に写真を投影するシステムによる食卓コミュニケーション支援, 情報処理学会論文誌, Vol. 31, pp. 103-108 (2004).
- [2] 木村有里, 丸山一貴: 会話中の単語に基づいた話題提供を行う雑談支援システムの提案, 電子情報通信学会技術研究報告, pp. 269-273 (2017).
- [3] 平岡拓也, 中村 哲, 戸田智基: 説得対話システムにおける話題誘導に基づく対話制御, 情報処理学会研究報告, Vol. 112, pp. 21-26 (2012).
- [4] 石田真也, 井上昂治, 中村 静, 高梨克也, 河原達也: 傾聴対話システムのための発話を促す聞き手応答の生成, 人工知能, Vol. 01, pp. 1-6 (2016).
- [5] Mozilla, M.: Speech Web API, MDN (online), available from <https://developer.mozilla.org/ja/docs/Web/API/SpeechRecognition> (accessed 2021-05-12).