

実作業を伴う音声模擬対話システムのユーザ発話現象

望主雅子 佐藤奈穂子 酒寄哲也 小島裕一

(株) リコー 情報通信研究所

1 実作業での発話

近年、音声対話現象の収集、分析がなされているが、対話の場や役割の設定が現実的でない場合も多く、リアルな対話現象がとれるかどうかには疑問が残る。これに対し、ゲーム的な要素を入れ、強い動機付けによる対話実験を行なう試みもある。

また、音声対話システムが使われる場面を考えた場合にキーボードが使用可能な状況よりも、使えない状況に音声の必然性があると考えている。

我々は、音声使用の必然性があり、実際の作業を伴うことでリアルな課題設定ができる、料理を実験タスクとし、データ収集を行なった。本稿では対話システムを使った際になされる発話を収集し、調査した結果を報告する。

2 先行研究

人間と人間の対話については、自己修復などの話しことば特有の現象は対話者が共有信念や知識を形成するためになされるという指摘 [1] や、日本語の話しことばについて話し手が文を完成させずに、聞き手が相手の文を補ったり、確認したりして対話を行なうことが指摘されている [2]。

最近では音声対話システムや音声模擬対話システムを使って人間と機械の対話現象を収集する試みも盛んに行なわれている。機械に対する人間の発話は、人間どうしでなされる発話と異なり、不適格現象が抑えられる傾向があることが指摘されている [3][4]。

3 対話データ

3.1 音声模擬対話システム

音声模擬対話システムを使ってデータ収集を行なった。音声模擬対話システム (Wizaro of OZ システム) は音声対話システムの音声認識の部分を人間が肩代わりして行なう実験システムである。技術的に完成度の低い部分を人間が補うことで、将来、技術が向上したときに予想される問題の把握や、完成度の低い部分の能力を向上させるための基礎データを収集することを目的とした実験方法である。

この方法では被験者 (以下ユーザ) に人間がシステムの一部になっていることをさとられずに、機械と話していると思い込ませて行なう。

3.2 実験タスク

実験では 20 名のユーザに音声対話システムで情報をやりとりしながら料理を作ってもらった。料理メニューは一般になじみのない、作り方の知られていない料理である (表 1)。

表 1: 料理タスクの実験

メニュー	アボカドの温製、半熟卵添え、ベアルネーズソース (アボカドを温めたものに、半熟卵のをせ、卵黄、エストラゴン、ピネガー、バターをベースにしたソースをかけたもの)
被験者	20 名 (30 才~50 才代の主婦)
設定	非対面、音声のみのやりとり
条件	発話スイッチの有無

実験システムの能力は以下である。

- タスクに関係する単語、句については認識 (スポッティング)、理解できる
- ユーザの質問には予め設定された知識だけで答える
- システム発話はタスク知識をもとに自動生成する
- 出力タイミングはユーザ発話の終了後 2 秒以内

システムは料理の作り方をユーザの質問や作業終了の発話に応じて音声で提示していく。対話の一例を表 2 に示す。

ユーザの作業状況は、ユーザからの発話のみによって検知する。システム側からユーザへの質問 (例: できましたか) や、状況を察しての発話は行なわない。

表 2: 対話例

s	: 卵を茹でます
u	: はい
s	: 6 分アボカドと一緒に茹でます
u	: アボカドの、お鍋の中に卵を入れるんですか
s	: アボカドと一緒に茹でます
u	: はい
u	: 卵はいくつですか
s	: 卵は 2 個です

(「s」はシステムの発話、「u」はユーザの発話)

実験によってユーザ発話 2185 を収集した。システム発話 (出力) は 1448 であった。

4 対機械発話と非対機械発話

4.1 発話の意図

ユーザの発話を意図で分類したのが表 3 である。

表 3: 発話意図と発話数 (全発話数:2185)

意図	意味	発話数
質問	機械に対して質問する	638
作業終了	現在の作業が終わったことを知らせる	415
もう一度	機械に再発話を要求する	26
返事	機械の発話に対して返事をする	163
待って	機械に発話を待つように要求する	5
あいづち	機械の発話にあいづちをうつ	93
独り言	独り言をつぶやく	518
復唱	機械の発話内容を復唱する	327

発話のうちの質問、作業終了、もう一度、返事、待って、あいづちが機械に対して話しかける意図の発話であるが、残りの、独り言、復唱は機械に対して話しかける意図を持たない発話と考えられる¹。

これらを、対機械、非対機械発話としてまとめたのが表 4 である。

表 4: 対機械、非対機械発話の割合

機械に対する発話	1340(61.3%)
機械に向けられない発話 内、独り言	845(38.7%) 518(23.7%)

作業を伴うユーザ発話には、機械に対する発話以外に機械に向けられない発話もあることがわかった。

4.2 対機械、非対機械の不適合性

それぞれにどのくらいの不適合現象が含まれているかを調べた。不適合表現として以下の現象を調査した。表 5 に発話の意図ごとの出現数をあげる。

- 口語表現 (例: 切るんですか)
書き言葉の語が話し言葉特有の語に置き換わった
- タスク外の語 (例: 疲れちゃった)
料理とは関係ない語の発話
- 言い淀み (例: えっと、次は)
- 言い直し (例: アボ、アボカドは)
- 言い替え (例: 分量は適宜、分量はお好みですか)

¹復唱はあいづちの一種ともとらえられるが、本データでは自分の確認のための独り言に近かった

表 5: 不適格現象の数

意図 (数)	口語	タスク	淀み	直し	替え
質問 (638)	146	5	55	21	15
作業終了 (415)	4	0	165	8	4
もう一度 (26)	1	0	7	0	1
返事 (163)	4	2	10	0	0
待って (5)	1	0	2	0	0
あいづち (93)	0	3	4	0	0
独り言 (518)	159	139	250	4	0
復唱 (327)	38	20	75	6	0
総数 (2185)	353	169	568	39	0

意図	不完全	省略	倒置	不整合	継続
質問	55	33	17	2	20
作業終了	3	9	2	0	0
もう一度	0	0	0	0	0
返事	0	1	0	0	0
待って	0	0	0	0	0
あいづち	0	0	0	0	0
独り言	91	19	23	3	2
復唱	45	6	6	0	2
総数	194	68	48	5	24

- 文末が不完全な表現 (例: 卵はいくつ)
述部部分が省略されている表現
- 助詞などの省略 (例: 分量教えてください)
- 倒置 (例: 入れました、卵)
述部が格より前に位置するもの
- 係り受けの不整合
「アボカド自身が、どの程度温めるんですか」など、助詞の誤り、挿入句によって係り受けの整合がとれないもの
- 継続
「大さじ1杯、でしょうか」など、一度言い終わったあとにつづけて発話する

これらに対機械発話、非対機械発話で1発話あたりの違反数で比べたのが表 6 である。機械に対する発話に比べて、そうでない発話には違反数が多く、1発話に少なくとも一つの不適合現象が含まれることになる。ユーザは機械に対して言語的な規則違反の少ない発話

表 6: 対機械、非対機械発話の1発話あたりの違反数

機械に対する発話	0.45(600/1340)
機械に向けられない発話	1.05(888/ 845)

をしていたことがわかる。機械に理解させるように発話していたことが予想できる。

4.3 発話スイッチの有無による差

実験条件に機械に対して発話する前にスイッチを押すかどうかを設定した。その場合の不適合現象の現われ方を比較したのが表 7 である。

表 7: 発話スイッチの有無による差 (1 発話あたりの違反数)

スイッチ	違反数
有り	0.37(133/363)
無し	0.62(432/694)

スイッチを押す方が押さない場合より違反数が少なく、スイッチによって機械に発話することが強く意識され、違反が抑えられたことがうかがえる。

5 対機械発話の現象

対機械発話の不適合表現には、ユーザが機械が受理できると想定した表現や発話上自然に現われるものが含まれていると考えられる。以下にいくつかの不適合現象について述べる。

5.1 相手にわかってもらうための発話現象

不適合表現の中でも相手に理解させるために発話している現象がある。一度発話したものに、さらに文法的に正しい形で情報を付け加える言い替え現象は、対機械発話のみに現われた。

u: ん、こしょうはじゃあ、黒こしょうは使いますか
u: はい。まだ、お、まだお湯になってないけど、 まだお湯が沸いてないんですけどいいんですか

また、一度不完全な形で言い終わったのち、句をつなげる現象も対機械発話に多くみられた。

u: 卵黄は割って崩して入れるの、かしら
u: ちょっとアボカドが崩れかけましたが、 よろしいですか

対人間の発話では、最初の言い終わりで発話終了するが、機械に相手の発話を補う能力がないことと、人間よりテンポが悪いこと、また機械への受理を考慮して文として完全な形に自分で補ったことが考えられる。

5.2 文末が不完全な発話

文末を完全にしないで、相手に委ねる、あるいは相手が補う発話現象が人対人では多くみられることを料理をつくるタスクで以前に確認した [5]。この実験の設定を表 8 に示す。

対人間と今回の実験の対機械の結果を比較したのが表 9 である。対人間に比べ、対機械では文末の不完全な発話は少なかった。

表 8: 人対人の料理タスクの対話実験の設定

設定	音声のみ、非対面
メニュー	4 品を同時につくる (鶏肉ハーブ焼き、ターメリックライス ほうれん草と菜の花のミルクチーズ炒め ビクルス) お菓子 (小豆入り蒸しケーキ)
被験者	4 名、親近性のある話者
発話数	ナビゲータ:446, ユーザ:499

表 9: 対人間、対機械での文末が不完全な表現の数

発話意図	割合 (現象数/発話数)
対人間	27.7(138/ 499)
対機械	8.9(194/2185)

5.2.1 質問の意図と表現

これらの不完全な発話とその意図との関係をみるために全発話中の質問・応答のパターンに着目した。

人対人では、質問文中、文末に質問を意図する言語表現 (「ですか」「でしたっけ」など) が現われていた発話は 46%(79/170) であった。

残りは文末が不完全で、名詞+「は」や格助詞 (例: 「フライパンに鶏肉が」「塩 1/4 と」) が一番多く、他に、終助詞 (「これ塩だったのかな」)、動詞、形容詞の終止形、接続助詞 (「で」「まだお米いれてないけど」) など、多くの形が位置した。

人対人では、質問の意図を表わす表現を用いない、あるいは、発話者に全く質問の意図がなくても相手がそれに答えるという対話現象があった。意図が限定されていたことも関係するが、言語表現や明確な意図に反応して発話しているわけではないことが予想される。

これに対し、機械に質問した場合の表現パターンを調べると、「ですか」「ですね」「ますか」といった通常の疑問文が 91%(583/638) と多く、文末が欠けた発話文は全体の 9% 程度であった。人に対するときと違い、はっきりと質問の意図を表わす形がとられたのがわかる。機械の能力も関係するが、対人間と対機械とでは発話すること、意図そのものに違いがありそうだ。

対機械発話中の文末が欠けた表現には、量や次の手順に関する質問が多く、対象物、疑問詞、項目と「～は～だ」(「ワインビネガーはどのくらい」「分量は?」) を組み合わせたパターンが多かった。しかし、中には人対人で現われた、動詞が省略された表現 (「どれくらいの量を」) もあり、機械の能力向上とともにこれらが増えていく可能性もある。

6 非対機械発話の現象

6.1 非対機械発話に特徴的な語

非対機械発話には不適格現象が多かったが、構成語も異なっていた。言語化できない現象も多い。非対機械発話に特徴的な語を調べたのが表 10 である。

表 10: 非対機械発話で特徴的な語

品詞	語句	対機械	非対機械
終助詞系	かな	2	66
	かしら	2	13
助動詞系	ちゃった	5	26
	あれ	1	15
感動詞	あ	14	52
	よいしょ	0	41
	わからない	3	22
動詞系	困る	0	8
	ちよっと	3	41
副詞	難しい	0	4
	熱い	0	12

発話者の意外な気持ち、不安が口にされた。

非対機械発話では口語的な表現のバラエティも多く、対機械発話の質問の発話が 3 パタンの表現(「一緒に茹でるんですか」「ベアルネーズソースって何」「汁けが無くなっちゃったんですけど」)で 93%(136/146)をカバーできるのに対し、独り言では 10 パタンで 62%(99/159)をカバーできる程度であった。

6.2 非対機械発話の出現箇所、意図

非対機械発話は、作業中や、システムへの質問に至る過程に現われた。

u: 全然出ない。どうしよう。困ったな。
u: 全然でない
u: ソースの水分が無くなっちゃったんですけど どうしよう
s: アボカドの温製を拭いて下さい。
u: 何って言ってる(ゆってる)んだろう
u: えー、何(なん)で聞こうかな
u: もう一回言って(ゆって)ください

内容は作業の感想、作業結果の評価、システムへの不満であった。ユーザの作業に関する感想や評価は、将来、タスクを円滑に進めるために、ユーザの状況を検知する手がかりとなる可能性もある。

システムへの不満はシステムの情報提示の内容が理解できない場合や、提示の順序が望ましくない場合に現われた。

作業の感想、評価

s: 半熟卵の殻を剥いて下さい
u: 半熟卵。<息>ちゃんとなってるかな
なってるか、ちゃんとなってる。よかった
u: きつ。全然剥けません。

システムへの不満

s: こしよは白いこしよです。
u: はい。先に言ってください、それを。<笑い>
s: 香草のスープをボウルで冷まします
u: ボウルで冷まします。ボウルで冷まします
u: ボウルで冷まします
わからないな、出来ないじゃないの、これじゃ

これが、対人間の場合は、同じような発話でも以下のように相手が何かしらの反応をする場合もある [5]。

(以下「N:」はナビゲータ、「U:」はユーザ)

N: お鍋はにたりましたか、煮立ったでしょうか
U: まだ、です
U: あ、蓋してないや
N: 蓋はしない

人対人では、情報伝達や情報要求の明確な意図がない場合も、発話することで、作業や対話の場を形成していくことが考えられる。

また、対人間では「わからない」「先に言ってください」などの明確な不満は口にだされることはなかった。機械であるがゆえにこういった不満などが遠慮なく言え、機械との対話の一つのメリットだともいえる。不満以外にも、同じことを何回も聞いたり、話題や対話を途中で打切るといったこともできる。

7 まとめ

音声対話を使った、リアルな課題でのユーザ発話には、機械に対する発話とそうでない発話があり、対機械発話では不適格現象が抑えられることを確認した。これらはユーザが機械への受容性を意識したこと、システムの能力(受理できる表現、知識、タイミングの悪さ)が影響していると考えられる。機械の能力向上に伴い、多くの不適格現象や独り言に近い発話が機械に向けられる可能性もある。また、機械と対話することと人間と対話することとの本質的な違いとそのメリットを考慮する必要がある。今後はこれらをふまえてシステムを構築していく。

参考文献

- [1] H.Clark and D.Wilkes-Gibbs. 1986 「Referring as a collaborative process」 *Cognition* 22
- [2] 水谷信子 1988 「あいづち論」 *日本語学* 7
- [3] S.Oviatt. 1993 「Predicting Spoken Disfluencies during Human-Computer Interaction」 *Proceedings of International Symposium on Spoken Dialogue*
- [4] 上條俊一、秋葉友良、伊藤克互、田中穂積 1994 「音声対話データの分析と発話理解への応用」 人工知能学会研究会資料, SIG-SLUD-9402-6
- [5] 望主雅子、酒寄哲也、小島裕一、佐藤奈穂子 1996 「ナビゲーション対話における省略文の分析」 情報処理学会研究報告 音声言語情報処理, SLP-13-5