

コンピュータと人間の会話: 現状と課題

Conversation between a Human and a Machine: the Present and Problems by Akira SHIMAZU (Japan Advanced Institute of Science and Technology, School of Information Science).

島津 明¹

¹ 北陸先端科学技術大学院大学情報科学研究科

1. はじめに

コンピュータと人間は自然言語でどれくらい会話ができるようになっているのであろうか。SF映画に登場するようなコンピュータはまだないが、話題をかざると、ある程度の会話ができるシステムが実験的に作られている。話題を限定せずに会話できるプログラムとしてはELIZAが有名であるが、人間の意図通りに答えてはくれない。会話システムは、自然言語処理としては、どのような言語表現を受け付けることができるかが問題であるが、コミュニケーションとしては、どのような会話が扱えるかが問題である。これまでの会話システムは文単位の発話交替を基本とし、ほとんどの場合、会話の主導権がコンピュータかユーザかのどちらか一方である。

音声言語により会話できれば、コンピュータはより身近なものになると期待される。このためには、人間同士の会話の分析も重要である。人間の話し言葉による会話を見ると、これまでの会話システムにはない種々の特徴が見られる。会話システムと対比すると、人間同士の会話では必ずしも文単位の発話交替ではなく、会話の主導権は可変である。

本稿は、自然言語によるコンピュータと人間との会話について、これまでの会話システムの会話の特徴、人間同士の会話の特徴、それらの比較、今後の課題について概観する。

2. 会話システムと人間の会話

これまでの技術レベルを会話システムが人間と交わす会話の特徴から概観する。会話を特徴付ける要因は多く、言語学、心理学、工学などの分野で研究が続けられているが、ここでは、主だった特徴といえる以下の3点を考える。なお、主なシステムの概要については文献17)などに譲る。

- (1) 形態：どんな形態で会話が進められるか。
- (2) 理解：相手の発話がどの程度理解できるか。
- (3) 発話：どんな発話が生成できるか。

これらの各面について見ると、音声で扱えるシステム

も含めて、これまでのシステムは限定されたものである。

3つの面について見ていく前に、会話システムがどのような観点から作られているかについて補足しておく。会話システムには実際的な自然言語インタフェース (Natural Language Interface) から言語理解の基礎研究によるものまである。これらの区別は必ずしも明確であるとはかぎらないが、すべてが同じ目的で作られているわけではない。

自然言語インタフェースの場合は、コンピュータへのアクセスを自然言語で行う実用システムの開発に主眼がある。その結果、効率的な構文解析、データベース・コマンドへの変換などの技術が具体化され、商品も出ている¹⁴⁾。

言語理解の研究はシステム化そのものより人間の会話の分析やモデル化に主眼がある⁷⁾。たとえば、会話参加者が話者と聴者の役割を交替しながら会話を進め、その過程で相互信念 (mutual belief) が形成されていくようなモデルを考える。話者は意図を発話により聴者に伝え、聴者はその発話から話者の意図を理解しようとする。意図を理解するということは、話者の目的および心的状態を推定することと扱われる。ここでは発話は1つの行為として扱う言語行為の考え方がとられる。

2.1 会話の形態・談話構造

発話の系列 会話システムSとユーザUの発話の交替 (turn) は次のように進められる。

U:s₁ S:s₂ U:s₃ S:s₄ U:s₅...

ユーザ発話 si ($i = 1, 3, \dots$) は文の場合が多い。システムの発話 si ($i = 2, 4, \dots$) は文のほかにも名詞句や「はい」「いいえ」の場合もある。ユーザが名詞句入力できるシステムもあるが、内容は限定されたものである。たとえば、「先月の〇〇社の売上額は」などのように、後に「いくらですか」とか「何ですか」などが続く入力を仮定している。あるいは、図-1のように前の質問と同様のパターンを繰り返す場合に動詞が省略可能というものである。

文単位というのは、文単位でひとまとまりの情報が

U1: カナダの首都はどこ。
S2: オタワです。
U3: ドイツは。
S4: ベルリンです。

図-1 会話例1

U5: 1ドルは何円ですか?
S6: 現在, 123円35銭です。
U7: アポロが月に初めて着陸した年は?

図-2 会話例2

S: Where do you go?
U: I want to go to Milano.
S: Where do you leave from?
U: I will leave from Torino.

図-3 会話例3⁹⁾

S: Automatic Railway Information System. Please speak after the tone. Tell me from where you want to leave and where you want to go.
U: ... leaving from Pisa at eight [NOISE] this morning [NOISE] Firenze-
S: From Pisa Aeroporto. Where are you going to?
U: To Firenze.
S: From Pisa Aeroporto to Firenze?
U: From Pisa ... Stazione Centrale to Firenze
S: From Pisa Centrale?
U: Yes

図-4 会話例4 (音声言語)³⁾

表現されているということである。文が単文であれば、述語とその目的語、主語などの格要素が表現される。省略語があれば、それは文脈から補完できることが仮定される。1発話でひとまとまりの情報を伝える点は、演劇などの台詞と似ている。このような点は日常の人間同士の会話とは異なる。

会話は参加者の発話の系列からなり、普通は連続する発話の間には関連性がある。発話の中で言及される概念や事象につながりがあるということである。コンピュータと人間の会話の場合、これには幅がある。データベースアクセスのための自然言語インタフェースであれば、連続するユーザ発話の間に関連がある必要は特にない。前の問合せの結果を受けて質問する場合は関連性を持つが、システムはユーザの入力を独立して解釈できるから、図-2のように、突然、何の前ぶれもなく話題を変える発話も可能である。このような会話は人間同士の会話では普通はない。

会話システムは発話の間隔も任意である。図-1において、システムがS2の出力をした後、かなりの時間をおいて、ユーザはU3の入力をすることも可能である。人間同士の会話において、「オタワです」と言った後、1時間も経って、相手が「ドイツは」と言った後、「ドイツがどうしました」などとなるであろう。

会話の構造 会話における文の並びには構造がある。

単純なものからより複雑なものまでである。

自然言語インタフェースでは文系列の構造は単純である。それはユーザ入力とシステムの応答からなる文の隣接対の系列である。ユーザ入力を受けては、それに答えるというサイクルの繰り返しである。隣接対と次の隣接対との間に関係がある必要はない。

航空便予約システムや交通経路案内など、特定の目的達成に向けてシステムとユーザとの間の会話が進められるような場合には、図-3のように一連の発話系列には関連が出てくる。

航空便予約などのシステムでは、予約という目標に向かった会話が進められ、そこに談話構造が形成される^{7),12)}。談話構造というのは、上位レベルでは、会話の開始、ユーザの要望の明確化、要望の解決、会話の終了という流れがあり、下位レベルでは、部分タスクに関する部分会話、不明確な点を説明する部分会話、発話やプランなどを訂正する部分会話などがある。

このような会話では、個々の発話は目的達成に必要な情報伝達あるいはそれを得るのに必要な問合せとして捉えられる。目的達成までの行動はプラン木として表現される⁷⁾。このようなプラン木に沿って、目的指向の会話をするシステムは、領域知識やユーザモデルなども使い、ユーザのプランを推定し、会話を進め、会話の談話構造が形成されていく。また、エキスパートシステムにおいて、ユーザの「なぜそうするのか」といった質問に対し、文脈に応じてユーザプランを推定し、システムがリアクティブに説明する技術が研究されている¹⁵⁾。このような問題はまだ研究途上であり、総合的な会話システムには十分に反映されていない¹⁾。

音声言語システムでは、図-4のように音声認識に起因する語認識を会話により修復する試みがある³⁾。

会話の主導権 会話の特徴付ける要素として、会話の主導権 (initiative) を誰が持っているか、主導権を持つことに自由度があるかという点がある。この主導権は談話構造の形成にもかかわる。CAIシステムなどでは特に教師のシステムが生徒のユーザに質問しつつも、ときにユーザからも質問ができる柔軟性が必要である。しかし、たいていの会話システムでは、会話の主導権はユーザか会話システムの一方が持つのが普通である。これはシステムの能力や頑健性などの弱点をカバーする意味合いがある。この制約のために、柔軟な会話が難しくなる。自然言語インタフェースの問題点として、しばらく使用していると煩わしくなることがあるともいわれているが、これはシステムと決まった形でしか会話できないのも理由と考えられる。自然な会話のために、昔から主導権を固定しない (mixed initiative, variable initiative) システムが指向されている¹⁷⁾ が、技術が確立されたとはいいがたい。

図-1と図-2はユーザが主導権を持っている例であ

る。例のように、ユーザが一方向的にシステムに質問あるいは指示をし、システムがそれに答えていく。このようなシステムがユーザに向かって、質問とか指示とかをしないわけではない。ただ、システムの質問は特別の会話モードでなされる。積木世界を対象にするシステム SHRDLU¹⁸⁾ の例を図-5に示す。

このようにシステムの質問にユーザは番号とか Yes/No で応えることとします。システムの質問にユーザが自然言語で応えた場合、ユーザの応答が解析できなかったり、曖昧であったりすることがある。そうすると、またシステムが質問するようなことが続き、際限のない深みに落ちる恐れがあるということである。これは Lisp listener で、error が生じたときに、ユーザの入力が不適切であると、error モードが深まるのに似ている。

このような問題は、目的達成に向けた会話の場合、会話内容が限定されるために、多少は扱いやすくなる。たとえば、航空便予約を対象とする会話システム GUS⁶⁾ の場合、予約に必要な項目を複数のフレームに表現しておいて、システムはそれらの項目を埋めるように会話を進める。ユーザの発話はフレームの項目に即して解釈される。その結果、図-6の(1)と(2)のようにシステムの質問にユーザが直接に答えなくても、システムは再度質問することなく会話を進めることができる。

音声言語システムで主導権が可変のシステムも作られている¹⁷⁾。図-7はユーザが主導権を持っている場合で、ユーザの発話内容に関連することをシステムが発話し、ユーザが問題点を自分で見つける(例では(1)の箇所)という会話である。

2.2 発話理解

システムが受け付けるユーザ発話はたいてい単文である。文型は Yes/No 型疑問文、WH 型疑問文、陳述文、命令文である。疑問文は検索式に変換され、データベースに照らして解釈され、その結果が応答となる。命令文はデータベースへのデータの登録や出力への指示などとして解釈される。米国では、1970年代から1980年代にかけて大規模な英語文法が開発され、自然言語インタフェースに適用された。その1990年代に商品が発表されている。

システムの頑健性を上げるために、入力文に余分な語、語の脱落、語の置換などの誤りがあってもよいようにする研究が進められている^{9), 16)}。音声言語に関しては、言い直し (repair) を扱う研究も進められている。交通経路案内の会話 (58,000語) における言い直し発話の95%を解析できる文法の報告もある¹³⁾。航空便予約をタスクにする音声言語システムの研究では、図-8のような言い直しを含む文の認識が報告されている²⁰⁾。ただし、例にも見るように音声言語システムといっても発話されるものは書き言葉の文である。

- U: How many things are on top of green cubes?
 S: I'm not sure what you mean by "on top of" in the phrase "on top of green cubes".
 Do you mean:
 1-directly on the surface
 2-anywhere on top of?
 U: 2
 S: Three of them.

図-5 会話例5¹⁵⁾

- U: I want to go to San Diego on May 28.
 S: What time do you want to leave?
 U: I must be in San Diego before 10 am. (1)
 S: Would you like P.S.A. flight 102 that arrives at 9.15 am?
 U: What is the next flight? (2)
 S: Air California flight 310 that leaves at 8.30 am and arrives at 10.00 am.
 U: I'll take the first one.

図-6 会話例6⁶⁾

- S: This is the circuit fix it shop. How may I help you?
 U: I want to fix this rs111. (i want fix this rs111)
 S: I am familiar with that circuit. The LED is supposed to be displaying alternately flashing one seven.
 U: The LED is displaying an alternately flashing one and the top corner of a seven. (be led it displaying alternately flashing one in top corner a seven)
 S: There is supposed to be a wire between connector nine nine and connector one zero zero.
 U: It is there. (it in there)
 S: There is supposed to be a wire between connector nine eight and connector one zero two.
 U: There is no wire on connector one zero four. (stays know wire i connector one zero for)
 S: There is supposed to be a wire between connector nine four and connector one zero two.
 U: The circuit is working. (the circuit is working) (1)
 S: Glad to have been of assistance. Good-bye.

図-7 会話例7. 括弧内は音声認識結果¹⁷⁾

- U: Give me flights from Minneapolis to I'm sorry Milwaukee to Seattle
 U: Could I have the airline and flight number to from Dulles to Tampa

図-8 発話例1 (言い直し発話例)²⁰⁾

構文解析の後、データベースアクセスの場合は、入力文に対する検索式でデータベースを照会し、文型に即して応答する。ユーザに親和的な情報システムのためには、ユーザ意図の理解⁷⁾が望まれる。たとえば、「図書館の本を検索できますか?」に対して、「はい」だけでなく「〇〇〇とタイプしてください」などと具体的に答えてくれたほうがよい。このような間接言語行為と呼ばれる発話も理解するための技術として、航空便予約システム GUS のような対象領域に関する知識の利用がある。また、発話も行為の1つと見なす言語行為の考え方に基づいて、言語行為と意図に関する

知識によりユーザの行動プランを推定し、間接言語行為などの発話を理解する方法も研究されている。

2.3 発話生成

言語生成技術の研究は地味だが着実に進められている⁹⁾。簡単なものから人間の発話のシミュレーションまで幅がある。

多くの会話システムはテンプレートを用いる簡便な方法によりシステムの発話を生成している。データベースアクセスのための自然言語インタフェースの場合、文以外にも表などの提示もでき、文の生成はテンプレートで足りるといえる。航空便予約のような特定タスクのインタフェースの場合も同様である。

意図を適切に伝えるために、意図表現からプランニング技術により発話を生成する方法も提案されている。ユーザとシステムとが何を共通に了解している、どのような点に了解（信念）のずれがあるかを推論しながら、意図表現から言語表現を生成する。また、適切な表現を生成するために、文法を用いる方法も研究されている。これらの方法は計算量も多く会話システムに組み込まれるほどにはなっていない。

人間の発話生成をモデルにする研究もある（文献15）のSmedtほかの論文）。発話内容が与えられて文の統語構造や言語表現を文頭から文末に向かって漸進的（incremental）に作るというものである。この研究は、人間の発話生成を模すものであるが、生成される文は書き言葉の文である。

漸進的生成法で、話し言葉の特徴を持つ表現を生成する方法も提案されている⁴⁾。この研究の会話システムは図-9に示すようにユーザの反応を受けながら交通経路を漸進的に案内する。ユーザは、システムの発話にあいづちを打ったり、駅名を問い返したりする。このユーザの反応によってシステムは発話の続け方を変更する。この変更の仕方は、人間同士の会話の統計分析に基づいている。システムは話し言葉の特徴として、「愛甲石田で降りてください」、「そこで降りて」などのように細かい情報単位で発話を進める。その結果、ユーザがあいづちや問返しがしやすくなっている。

3. 人間の会話の特徴

前章で着目した会話形態、理解、発話の3つの面について、会話システムとの対比で見てみる。

3.1 会話形態

発話系列 人間同士の会話における発話の系列は、書き言葉のような文の系列ではない、会話システムのように整然と発話交替がされない¹²⁾。図-10に見るように同時発話や発話の重なりが生じるという特徴がある。図-10で、<はい>は挿入発話を、=は間が空かずに相手発話が続くことを、[3] はそこから同時発話が始まることをそれぞれ示す。

情報は図-10に見るように細かい単位により伝達さ

- S: 小田急線に乗っくはい>で愛甲石田で降りてください愛甲石田が厚木センタの最寄り駅になりますので
U: はい
S: そこで降りて北口バス乗り場というバス乗り場が
U: 北口バス乗り場
S: はいありますので
S: そこで森の里青山行きというバスに
U: 森の里
S: 青山行き<はい>それに乗って

図-9 会話例8（音声言語）⁴⁾

- D: 今、どこにいます?=
K: =あ、あの研究所の前の [3] 正門のところですよ。
D: [3] あ、あ、そうですか。<はい>えっとですなえ、えっとー、具体的に経路ってというのは、<はい>最寄りの駅から愛甲石田まで、...

図-10 会話例9

- R: えーと、僕の机の上にありますね、
A: はい
R: 本あるでしょう、4冊。
A: あっ、はい
R: それをですね
A: はい
R: ちょっと持って来てくれますか。
A: はい

図-11 会話例10⁸⁾

- R: 僕の机の上に本が4冊ありますか?
A: はい
R: それを持って来てください。
A: はい

図-12 会話例11⁸⁾

れる。簡単のために図-11により説明する⁸⁾。図-11の会話をこれまでの会話システムとの会話にすると図-12のようになるであろう。例に見るように、書き言葉で1つの文で表現されるものが、話し言葉の会話では、「僕の机の上にありますね」、「本あるでしょう、4冊」というふうに細かい情報単位に分れて、コミュニケーションが進んでいることが分かる。この点が会話システムと人間の会話の違う点である。発話単位当たりの格要素数は平均1.01個という分析⁴⁾がある(表-1)。参考として、科学技術解説文献で述語に対する格要素が平均1.83個という分析がある。書き言葉の場合でも、文脈から分かる要素は表現されないが、話し言葉の場合は、細かい単位で情報伝達が進められることから、述語当たりの格要素数が少くなっている。

会話の構造 談話理解の研究として長く研究がされており、入れ子構造、隣接対の木構造（contribution tree）²⁾などの分析が報告されている。たとえば、「A：質問₁，B：質問₂，A：質問₂に対する答₂，B：質問₁に対する答₁」といった構造である。

会話は参加者のさまざまな調整（coordination）により進められる。この調整には発話、談話、意図、

表-1 格要素の出現傾向

対象	平均	分散	最大
交通経路案内（発話単位当たり）	1.01	0.29	3
科学技術解説文献（述語当たり）	1.83	0.85	5

表-2 会話における調整の階層

階層	内容
発話	あいづち、間合いなどにより形式的に会話が進められる。
談話	前提、意味などが整合的に用いられる。
意図	字義通りの意味だけでなく、言外の意味も含めた意図が理解される。
協力	相手の意図に協力的に会話が進められる。

協力の各レベルがある⁸⁾ (表-2)。たとえば、喧嘩をしていても少くとも発話レベルの調整はなされている。“A：おー、B：ばーろー、A：てめえー...”という具合に発話レベルでは協調的な同期がある。ここでは、相手発話に呼応して発話するなど会話を成立させる制約が満たされている。明らかにどのレベルもが満たされるわけではない。協力的でない会話もあるし、質問されているのに答の発話をしない会話もある。

会話の主導権 話し言葉による会話は、例にも見るように、あいづち、問返しなどにより会話進行の調整がされている。そういう点や細かい単位による発話のために、話者の交替が会話システムに比べて多い。しかし、これは会話の主導権も同様に交替しているわけではない。交通経路案内のような会話では、説明する側が主導権を持っている。図-13において、「... 駅になるんですよ」といったところで、案内者は、いったん、主導権を保留する。例では相手がそこで質問をするが、もし、相手の発話が「はい」だけであったら、また、主導権をとって案内が続くことになる⁸⁾。

3.2 発話理解

前述したように、話し言葉は漸進的に細かい単位で表出され、そこには省略、倒置、言い直し、言い替え、誤りなどもあるが、人間はそのような特徴を持つ発話を実時間で次々に理解していく。発話には間接言語行為もある。あいづちの半数以上が相手発話の終了前に始まっているという分析もある。同時発話もある。これらのことから、相手発話が終わるより前にかんがりの理解が進み、発話の生成処理が進められていると想像される。このようなことができるためには、発話の音声面、表現面、内容面などに会話を予測させるような特徴があるとみられ、このような面につながると思われる研究もある。たとえば、会話の調整にかかわる談話標識語（あいづち、フィラー、接続助詞など）の韻律特徴、あいづちとポーズの関連などの分析がある。

書き言葉で1個の文で表現されるものが、話し言葉では細かい情報単位に分けて表現されるから、話し言葉の理解は複数の発話からなる談話が対象となる。細かい単位からなる談話では、まず、何に注目するかが示される。着目対象を領域と呼ぶことにすると、次に、

D: えー、そこが、あの一、厚木研究センターの最寄りの、え、駅になるんですよ。
 T: はい。愛甲石田は急行は止まるんですか?
 D: えっと、愛甲石田、急行一、止まります。

図-13 会話例12

A: 前売券を買う所があるらしいんですけどもくはい>どちらかご存知でしょうか?
 B: はい、あの前売券は、〇〇〇じゃなくてもいいんですけど、旅行代理店や、
 A: はい
 B: あと××の窓口があるんですよ。

 A: 18歳以上が4,800円です。
 B: 4,800円ですわね。

図-14 会話例13

領域がどのような状態か、あるいは、領域に関し、どのような行為をするかが示される⁸⁾。図-11でいうと、「僕の机の上にありますね」で領域が示され、次に「本あるでしょう、4冊」で、領域における状況が示される。このような構造は再帰的になっていて、領域と行為が合わさって1つの領域となり、それに対して次の行為が示される。このように、書き言葉の1文に相当する情報内容が、日常会話では、相手とのやりとりによる情報も含め、複数の発話系列で表現される。

3.3 発話生成

情報伝達の観点からは、上述したような細かな単位による発話が話し言葉として特徴的である。話し言葉の特徴として、言い淀み、言い直しなどがよく上げられるが、言い直しがある発話は、交通経路案内80対話の分析では、全発話の5%程度である¹³⁾。言い淀みや言い直しがなくても話し言葉らしくしているものがあるということである。

言い直しにもさまざまな場合がある^{11),13)}。Levelt¹¹⁾は英語について、instant repair, anticipatory re-tracing, fresh restartの3種類に分類し、それぞれ、42%, 35%, 23%という分析を報告している。日本語については文献(10), (13), (16)などに譲る。

話し言葉はいうまでもなく、会話という状況で使われる。このことを反映する言葉に終助詞がある。図-14では、相手が知らないと思われることを提示するのに「よ」を使い、確認に「ね」を使用している。このような「よ」「ね」は書き言葉にはない表現である。

人間同士の会話では間（マ）が重要である。1人の発話が終り、間ができたとき、同じ話者が発話権をとることもあれば、相手が発話権をとることもある¹²⁾。これに対し会話システムは硬い。ユーザに主導権があれば、ユーザが発話（タイプ）しなくても、システムはいつまでも待つ。

4. 課題

会話システムの現状を見て、人間同士の会話を観察したとき、今後の会話システム、広くは人間機械インタフェースを考える上で、多くの問題点があるが、ここでは報告者なりの課題をあげる。

(1) コンピュータとの会話はどうかあるべきか?

従来の会話システムの会話は、音声言語によるものも含め、人間同士の会話とは異なる次元にある。工学的には、人間とコンピュータとの会話を人間同士の会話の延長線上で追求すべきものか、まったく別のものとして追求すべきかは自明ではない。どちらのアプローチをとったとしても、少くとも、人間の会話の特徴、人間の認知特性は無視できない。

(2) 会話の調整をどのように扱うか?

自然な人間機械インタフェースには調整の階層が重要であり、会話を成立させるためには、基本的なレベルである発話調整が必要である。自然言語インタフェースではないが、現金自動預入支払機ATMの音声ガイダンスで、ユーザがボタンを押すのに手間取っていると、やぶからぼうに「確認のボタンを押してください」などと催促してくるものがある。これなどは、発話レベルの調整さえ守られておらず不快感を抱かせる。あいづちに関する研究が多くなされているが、このような点からも、うなづけることである。今のところ、人間が発話調整をどのようにしているかは自明ではない。話者は一般には相手の反応を予測しながら発話をしていると思われるが、相手の反応に合わせながら発話調整、たとえば、あいづちをしているかどうかは分からない。今後、マルチモーダルインタフェース¹⁰⁾が増えてくるが、音声、文字、絵といくつものモダリティを扱うようになると、この点の設計が問題になる。

(3) 話し言葉のどのような特徴を反映させるか?

システムとの会話に話し言葉の特徴を反映させるのがよいかどうかは自明ではないが、細かい発話単位およびインタラクションにより、情報伝達が正確になり、会話全体で見たときに伝達量が少なくなることが期待される。書き言葉の1文がうまく伝わらないで発話し直したとき、最悪の場合は元の文を再発話することになる。一方、話し言葉は細かい単位に分割されているから、うまく伝わらない部分が特定され、再発話する部分も小さくなると予想される。

(4) 理解の問題をどう扱うか?

実際の会話を見ると理解のされ方にも幅がある。聞き違いや誤解が生じて修復を試みる会話もある。すなわち、共有信念に基づいて理解や発話をしていたとしても、それらが正しいとはかぎらない。ちょうど、吟味したプログラムを走らせて、エラーが陽に現れるのにも似ている。また、発話には言語・社会規範的な

ものもあり、必ずしも、複雑な計算が必要とはかぎらない場合もある。

以上のような面と音声認識の性能を考えると、理解にも調整の階層のように幅を考え、会話の中でユーザの意図を捉えていくことが実際的と思われる。

(5) 実時間の漸進的な処理をいかに扱うか?

会話は実時間でなされる。音声認識、自然言語処理、各レベルで高速化の研究がされているが、発話は漸進的になされ、細かい単位で発話のインタラクションがなされることを考慮した方式を追求する必要がある。

(6) 現象から何を学ぶか?

対話コーパスが少しずつ蓄積され、具体的に種々のことが調べられている¹⁶⁾。多くの場合、場面を設定して被験者の会話を収集している。日常のものをそのままとることはプライバシーなどの問題があるからである。演劇にも会話があるが、これは日常のものとは違うから、会話データは工夫して集める必要がある。

会話コーパスから言語データや知見を得て、会話システムに生かしていくわけであるが、人間同士の会話だといっても幅があり、どれもが模範的とはかぎらない。会話における人間の特性を明かにし、何を会話システムに生かすかを見極めることが必要である。演劇、落語、漫才など、芸人は昔からどのような言い回しとか間(マ)をとれば、観客に受け入れられるかを芸として磨き、それが受け継がれてきている。それに対し、現代の会話研究は会話のパラメータを科学的に見出し、インタフェースに生かそうとしているともいえる。

(7) どのように評価するか?

技術の発展には評価が重要である。会話では自然性や情報量が評価尺度になるが、自然性だけでは必ずしもよい評価とはならないようである。米国に会話システムのコンテストがあるが、人間のタイピングの癖を真似するELIZA方式のシステムが幅をきかせてしまっている。このため、日本ではDiaLeague⁹⁾と呼ぶコンテストが開催され、会話の情報量の面からシステムを評価する活動がある。自然性も単に実際に近いだけでなく、人間機械インタフェースとしての親和感も問題である。評価にあたり、人間の会話自体が規準になるが、どのような特徴を対象にするか、特徴の値はどのようなものであればよいかを明らかにしていく必要がある。

5. おわりに

会話システムの現状、人間の会話の特徴、今後の課題について概説した。会話システムとユーザとの会話では文単位で発話交替を繰り返すなどの特徴をあげた。人間同士の会話では細かな発話単位による表現、会話の調整などの特徴をあげた。このような特徴を踏まえ、会話システムとして考えるべき課題をあげた。

会話システムの研究の歴史は長く進歩は必ずしも速くはないが、種々の研究が進んでいる。会話システムの研究開発は、よい人間機械インタフェースにもつながり、今後期待するところは大きい。

謝辞 資料を提供していただいたNTT基礎研究所に感謝します。

参考文献

- 1) Allen, J. F., Miller, B. W., Ringger, E. K. and Sikorski, T.: A Robust System for Natural Spoken Dialogue, In the Proc. of Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics, pp.62-70 (1996).
- 2) Clark, H. H. and Schaefer, E. F.: Contributing to Discourse, Cognitive Science,13, pp.259-294 (1989).
- 3) Danieli, M.: On the Use of Expectations for Detecting and Repairing Human-Machine Miscommunication, In the Proc. of AAAI-96 Workshop on Detecting, Preventing, and Repairing Human-Machine Miscommunication, pp.87-93 (1996).
- 4) 堂坂浩二, 島津 明: 話し言葉対話コーパスにおける協調的対話原則の分析, 電子情報通信学会技術研究報告 NLC-97-1 (1997).
- 5) 橋田浩一, 伝 康晴, 長尾 確, 柏岡秀紀, 酒井桂一, 島津 明, 中野幹生: DiaLeague - 自然言語処理システムの総合評価 -, 人工知能学会誌, Vol.12, No.3, pp.390-399 (1997).
- 6) Bobrow, D. G., Kaplan, R. M., Kay, M., Norman, D. A., Thompson, H. and Winograd, T.: GUS, A Frame Driven Dialog System, Artificial Intelligence, 8, pp.155-173 (1977).
- 7) Cohen, P. R., Morgan, J. and Pollack, M. (eds.): Intentions in Communication, The MIT Press (1990).
- 8) 川森雅仁, 島津 明: 対話における統御の概念, 電子情報通信学会技術研究報告 NLC96-25 (1996).
- 9) 長尾 確他: 特集「自然言語処理技術の最近の動向」, 情報処理, Vol.33, No.7 (July 1993).
- 10) 中川聖一他: 特集「音声言語情報処理の現状と研究課題」, 情報処理, Vol.36, No.11 (Nov. 1995).
- 11) Levelt, W. J. M.: Speaking, The MIT Press (1989).
- 12) Levinson, S. C.: Pragmatics, Cambridge University Press

(1983).

- 安井 稔, 奥田夏子 (訳): 英語語用論, 研究社出版 (1990).
- 13) 中野幹生, 島津 明: 言い直しを含む発話の解析, 情報処理学会研究報告 SLP97-16 (1997).
 - 14) NAIレポート: 米国でデータベース用自然言語インタフェースの利用が進展, 日経AI, 1.28 (1991).
 - 15) Paris, C. L., Swartout, W. R. and Mann, W. C. (eds.): Natural Language Generation in Artificial Intelligence and Computational Linguistics, Kluwer Academic Publishers (1991).
 - 16) 白井克彦他: 特集「音声対話」, 人工知能学会誌, 12, 1 (1997).
 - 17) Smith, R. W. and Hipp, D. R.: Spoken Natural Language Dialog Systems, Oxford University Press (1994).
 - 18) Winograd, T.: Understanding Natural Language, Academic Press (1972).
 - 19) Young, S. and Bloothoof, G. (eds.): Corpus-Based Methods in Language and Speech Processing, Kluwer Academic Publishers (1997).
 - 20) Zue, V., Seneff, S., Polifroni, J., Phillips, M., Pao, C., Goddeau, D., Glass, J. and Brill, E.: The MIT ATIS System: December 1993 Progress Report, Proc. of the Spoken Language Technology Workshop, pp.66-71, March 6-8 (1994).

(平成9年12月9日受付)



島津 明 (正会員)

1948年生。1971年九州大学理学部数学科卒業。1973年同大学院修士課程修了。同年日本電信電話公社武蔵野電気通信研究所入所。1985年日本電信電話(株)基礎研究所。1997年北陸先端科学技術大学院大学情報科学研究科教授。推論手法, 形態素解析, 構文解析, 意味解析, 文生成, 質問応答システム, 機械翻訳システム, 対話システム等の研究に従事。工学博士。言語処理学会, 計量国語学会, 電子情報通信学会, 人工知能学会, ACL, ACM各会員。