

# 情報アクセスにおける受動性と能動性: 音声対話によるニュース記事アクセス

## Intentionality in Information Access Behavior: A Spoken Dialogue System for Interactive Access to News Articles

林 良彦<sup>1\*</sup> 藤江 真也<sup>2,1</sup> 福岡 維新<sup>1</sup> 高津 弘明<sup>1</sup> 小林 哲則<sup>1</sup>  
Yoshihiko Hayashi<sup>1</sup> Shinya Fujie<sup>2,1</sup> Ishin Fukuoka<sup>1</sup> Hiroaki Takatsu<sup>1</sup> Tetsunori Kobayashi<sup>1</sup>

<sup>1</sup> 早稲田大学 <sup>2</sup> 千葉工業大学  
<sup>1</sup> Waseda University <sup>2</sup> Chiba Institute of Technology

**Abstract:** Passive information consumption would an adequate type of information behavior for receiving the content of, for example, a news article. It may however be boring in many cases and even painful in some cases, especially when the information content is delivered by employing speech media. The user of a speech-based information delivery system, for example a text-to-speech system, usually cannot interrupt the ongoing information flow, inhibiting her/him to confirm some part of the content, or to pose an inquiry for further information seeking. We thus argue that spoken dialogue is a suitable media for enabling interactive information access that coordinate passive information consumption and active information seeking. This paper shows that a carefully designed spoken dialog system could remedy these undesirable situations, and further enables an enjoyable conversation with the users. The key technologies to realize such an attractive speech-based interactive information access system are: (1) pre-compilation of a dialog plan based on the analysis of a source content, and (2) the dynamic recognition of user's state of understanding and interests during the course of conversation. This paper illustrates technical views to implement these functionalities, and discusses a dialog example to exemplify our approach.

### 1 はじめに

人間の情報に関する行動 (information behavior) のうち、情報獲得・収集 (information acquisition) に関する行動は大きく、意図的な情報探索 (intentional information seeking) と、意図性のない受動的な情報行動 (unintentional passive information behaviors) に分けられるとされ [1]、情報学の分野では、主に前者を導く動機や状況に関するモデルの研究が行われてきた [3].

コンピュータサイエンスの領域においても、その焦点はもちろん前者にあり、情報検索 (information retrieval)、あるいは少し広い概念としての情報アクセス (information access) のシステムについて、様々な観点からの研究開発が活発に行われてきた。

以上の研究状況の背景を推察するに、情報遭遇 (information encountering) などの受動的な情報行動は、主として偶発的な状況によることから、研究的な要素に乏しいと考えられてきたのではないかと考えられる。

しかしながら、我々の日常の情報行動の実際をみれば、両者の区別は必ずしも明白ではなく、むしろ、これらの情報行動の状態を自由に遷移する過程であると考えるのが妥当であろう [2].

さて本研究では、音声対話によるニュース記事アクセスシステムをとりあげる。ユーザ側からみれば、ニュースに関する情報を音声メディアを用いて獲得し、あるいは、消費する情報アクセスシステムであるが、システム側の観点から言えば、音声メディアを用いて、ユーザに伝えたい・伝えるべきニュースを伝達するという情報伝達システムである。

システムが一方向的に記事の内容を読み上げるとすれば、ユーザは読み上げ音声を黙って聞き続ける必要がある。ユーザにとって内容的に冗長である可能性もあるし、そもそも記事の内容に興味がないことに気づく場合もあるだろう。

このような情報提示システムの対極に、記事に関する簡単な内容 (例えば記事の見出し) を与えた後に、ユーザからの質問を一問一答形式で受け付けるモードに移行する質問応答型のシステムが考えられる。このよう

\*連絡先: 早稲田大学理工学術院 実体情報学博士プログラム  
〒169-0072 新宿区大久保 2-4-12 ラムダックスビル 3F  
E-mail: yshk.hayashi@aoni.waseda.jp

S\_1: 羽生結弦選手が  
 U\_1: うん  
 S\_2: 国際大会を欠場することになったよ  
 U\_2: え?  
 S\_3: 欠場するんだ、腰の痛みのためだって  
 U\_3: 腰の痛み... って?  
 S\_4: 練習中に腰を痛めたということなんだ  
 ...

図 1: 想定する対話の断片例. “S\_n:” はシステム発話, “U\_n:” はユーザ発話を表す.

なシステムのユーザは、適度な量の情報を得るまで、質問を発し続けることが必要になる.

先の情報獲得における意図性の議論からすれば、ユーザは両者のモードを、その「状況」に応じて、しかも簡単な手段によって、行き来できることが望まれる. そこで、本研究が想定するような対話を単純化した断片の例を図 1 に示す. システムから伝達された情報に対してユーザは、必ずしも言語的ではない即応的な反応 (U\_1=肯定的, U\_2=疑問) によって理解状況を示したり、さらに対話の過程で生じた情報要求 (国際大会を欠場する理由) をある程度明確な言語表現を用いて示し (U\_3=問い返し) たりする. システムは、必要に応じてこれらのユーザの状況を推定し、適切と思われる情報を付加しながら応答を返す (S\_4).

本研究の前提、あるいは、主張は、このような受動的な情報獲得を主体としつつもインタラクティブ性を要する・有する情報行動の支援形態として、音声対話が適しているという点にある.

本稿の以下では、まず情報学における関連研究を参照しながら、上記の議論を補強し、本研究のスタンスを明確化する. 次に、現在開発中のニュース記事を対象とする音声対話システムについて述べ、最後に今後の課題や研究の方向性について論じる. なお、音声対話システムに関する内容は、当研究グループにおける既発表 [17] の内容によっている.

## 2 情報行動における受動性と能動性

これまでに提案されている情報行動の分類には様々なものがあるが、Erdelez による分類 [2] を図 2 に示す. ここでは、非意図的な情報行動は機会主義的情報獲得 (Opportunistic Acquisition of Information: OAI) と呼ばれており、その主な下位分類として、情報遭遇 (information encountering) が位置づけられている.

Erdelez はさらに、(1) 気づき (noticing); (2) 停止 (stopping); (3) 検討 (examining); (4) 獲得 (captur-

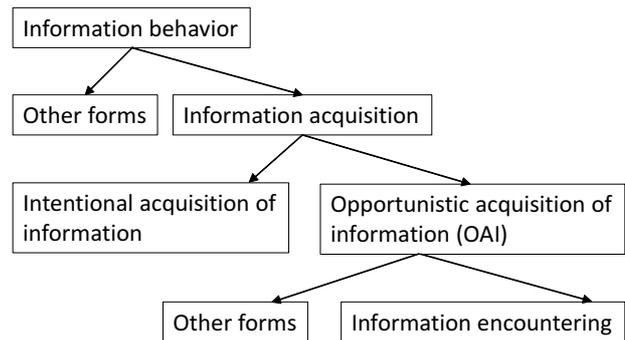


図 2: Erdelez による情報行動の分類 ([2] より作図).

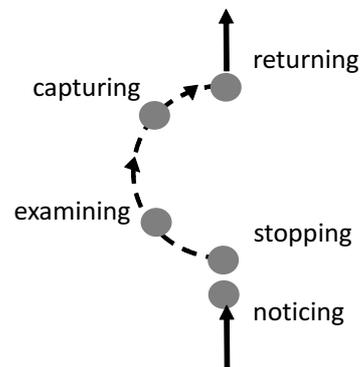


図 3: Erdelez による情報遭遇の機能モデル. 実線部が foreground interest, 点線部が background interest を表す. ([2] より作図).

ing); (5) 復帰 (returning). の各段階からなる情報遭遇における機能モデル (図 3) を提示した.

このモデルでは、ユーザは彼/彼女の主要な関心 (foreground interest) に関わる能動的な情報探索タスクを実行していることが仮定されているが、重要なことは、ユーザはこの過程の中で関連する関心 (background interest) に気づき、foreground の情報探索タスクを一旦停止したうえで、background に関する情報行動 (検討・獲得) を行い、その後に foreground の情報探索タスクに復帰するという点である.

以下、図 1 に示した対話例をこのモデルをと照らし合わせて考えてみる. この対話における「主要な関心」は、対象のニュース記事により定まる「羽生結弦選手の国際大会欠場」にある. これが主要な関心となる契機については問わない<sup>1</sup>が、システムが記事内容の伝達を行っている間は、ユーザは基本的には受動的な情報消費のモードにある.

本研究では、この対話でのユーザによる U\_3 の発

<sup>1</sup>すなわち、システム側が勝手に見つけた記事 (受動的あるいは機会主義的な情報獲得) かもしれないし、ユーザによるある種の情報検索の結果として選択されたもの (意図的な情報探索) であってもよい.

話(「腰の痛み... って?」)の背景には、図3のモデルにおける noticing に相当する過程があると考えられる。すなわち、foreground に対する情報行動の過程の中で background に対する情報行動へのシフトが起こっている。ただし、図3の Erdelez のモデルでは、能動的で意図的な情報探索の過程において偶発的な情報遭遇が生じているのに対し、図1に示す対話においては、ある程度の明示性のある言語表現(「腰の痛み... って?」)によって、モードのシフトが起こっている点が異なる。しかしながら、background に関する情報行動のモードが一段落した後は、foreground の情報行動のモードに復帰することは共通している。

以上にみたような、ユーザにとって自然で効率の良い情報行動は、受動的な情報消費から、意図的な情報要求による情報探索までの情報行動のモードをその「状況」に応じて、しかも簡単な手段によって、行き来することにより達成される。本研究では、このような情報アクセス・情報伝達を実現するために音声対話が自然で効率よいメディアであるという前提にたち、次節で述べるような音声対話によるニュース記事アクセスシステムを提案する。

### 3 音声対話によるニュース記事アクセスシステム

#### 3.1 システムの要件

理想的な対話システムの実現へ向けては様々な課題が存在するが、少なくとも以下のような要件を考慮する必要がある。

**伝えるべきニュース記事の選択:** どのようなニューストピックを対象とするかの決定は、本研究の範囲外とする。すなわち、「本日の重大ニュース」でも、「本日のおすすめニュース」でも、ユーザによる情報検索の結果として選択されたものでも良い。

**伝えるべき内容の選択と構成:** システムは、対象とするニュース記事(群)が与えられたとき、最低限どのような内容を伝えるべきかを決定する必要がある。ここでいう最低限伝えるべき内容とは、対話においてユーザが完全に受動的である場合においても、システム側がとにかく伝達しようとする骨格的な情報内容である。要約と言っても良いだろう。このような要約に相当する情報内容を補足する補助的な内容は、ユーザからの具体的な情報要求に基づいて提示することになる。

**ユーザの状況の把握と対応:** ユーザは、システムの発話に対する自身の理解状況や伝達内容への興味の状況

を反映して、肯定的・否定的な短い即応的な発話(以下、即応的情報反応と呼ぶ)や、もう少し明示的な情報要求を発することが想定・期待されている。したがって、システムはこれらのユーザの反応・発話からユーザの状況を適切に把握し、さらにはそれに応じた応答を返す必要がある。

**リズムのある対話の実現:** リズムのある対話を実現するためには、即応的情報反応を含むユーザの短い発話に対して、システムは素早く応答できることが望まれる。よって、対話の過程で提示されるユーザからの情報要求をある程度見越して、「こう聞かれたら、こう答える」ということを定めておくことが必要となる。

#### 3.2 システムの構成

以上のような要件を(ある程度)満たすものとして、我々が提案する音声対話システムの構成を図4に示す。システムは大きく分けて、ニュース記事をもとに対話に利用する発話計画を生成する事前処理部と、ユーザを相手に対話を行う対話システム部の二つから構成される。

事前処理部は、インターネットからの取得などによって与えられるニュース記事を解析して構造化する構造解析部と、その結果をもとにユーザの反応を織り込んだシステム発話計画を生成する計画生成部からなる。発話計画は、記事における主要な内容を伝達するための主計画と、それを補足する補助的な内容を伝達するための副計画からなる<sup>2</sup>。構造解析部については3.4で、計画生成部については3.5でそれぞれ詳細を述べる。生成された発話計画は、発話計画データベースに保存される。

対話システム部は、発話計画を読み込み、それに従って対話を進める。音声認識器は、ユーザの短い反応を認識する。対話制御部は、発話計画に従ってシステムの発話内容を含む発話文を音声合成器に出力する。また、システムの発話に対するユーザの反応に応じて、事前に生成した発話計画に従って発話内容の調整を行う。音声合成器は、対話制御部から生成された発話文を音声に変換してユーザに提示する。

本システムにおいて、音声認識器はATR-Trek製のものを使用している。また、音声合成器はOpen JTalk<sup>3</sup>を基に開発したものをを用いている。

#### 3.3 発話計画とユーザの反応・応答

ここでは、システム側の視点から発話計画とユーザの応答の関係について述べる。本システムでは、想定

<sup>2</sup>これらはさらにネストしていてもよい。

<sup>3</sup><http://open-jtalk.sp.nitech.ac.jp/>

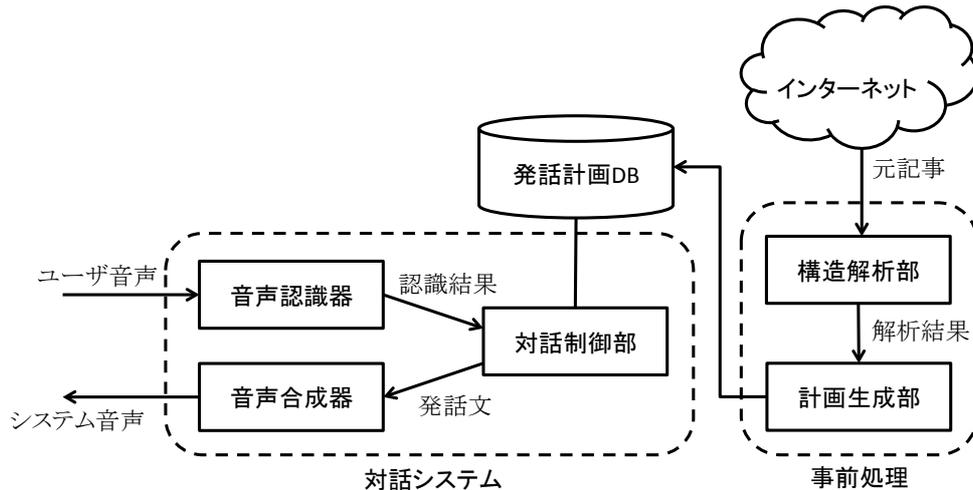


図 4: システム構成

されるユーザの反応を織り込んだ形で発話計画を事前に生成しておき、実際の対話時にはそれに従うことで効率的な情報伝達を実現する。ここではその発話計画がどのようなデータ構造を持つか、また、それによって対話制御部がどのように対話を進めていくかについて述べる。

まず、発話計画のデータ構造を図5を用いて説明する。図に示した通り、発話計画は状態遷移構造として表される。各アークに示された  $U_{i,j}$  はユーザ発話を表し、 $S_{i,j}$  は、システム発話を表す。ユーザ発話は音声認識部へ送られ、システム発話はそのまま音声合成部に渡される。

ユーザ発話  $U_{i,j}$  は即応の情報反応を含む短い応答を想定しており、下記の2つのカテゴリのうちいずれかに分類される。

**肯定的応答 (ACK)** 「うん」「へー」といった相槌など、システムの発話進行に肯定的な態度を表す反応。システムの発話の一部を下がり口調のイントネーションにより反復する場合を含む

**否定的応答 (NACK)** 「え?」といった、システムの発話進行に否定的な態度を表す反応。上がり口調のイントネーションによるシステム発話の一部の反復を含む

すでに述べたように、システム発話の一部を反復することは、ある程度明確な情報要求がユーザにおいて発現している状況、すなわち、情報行動のモードにシフトが起こっている状況を示唆する。また、否定的な応答は、システム発話のいずれかの部分が聞き取れなかった状況、あるいは、発話された内容に対して明示的でない情報要求が生じている状況を表していると考えられる。

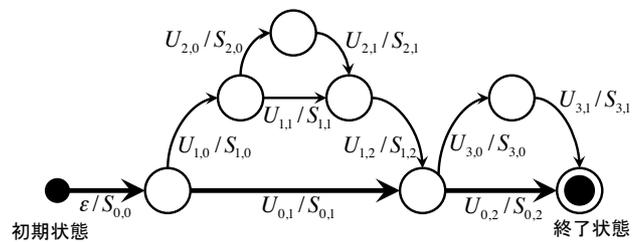


図 5: 発話計画の構造.  $U_{i,j}$  はユーザの反応,  $S_{i,j}$  はシステム発話を表す。

$\epsilon$  はユーザの反応を見ずに状態が遷移することを表す。図中で太いアークにより表されている部分を主計画と呼ぶ。主計画は、システムが最低限伝えるべき、記事の骨格をなす内容情報を表す。ユーザが受動的な態度 (すなわち、反応がないか、ACKのみを示す) を取り続ける限り、システムは主計画に従って淡々と、この内容情報の伝達を進めることになる。

本システムにおける対話制御は非常にシンプルで、アークに与えられたシステム発話  $S_{i,j}$  の内容を音声合成器に出力した上で状態を遷移させる。各状態ではユーザの反応を待ち、その内容に従って次の遷移を行う。例えば主計画上にいる際は、特にユーザの反応が得られなくてもそのまま発話を続けることが好ましいと思われるので、図中、 $U_{0,1}$  や  $U_{0,2}$  などは、特に反応が無くても時間経過 (例えば 0.6 秒無反応で経過) によって遷移させる。

否定的応答が認識された場合 (例えば  $U_{3,0}$  が得られた場合) は、その直前のシステム発話 ( $S_{0,1}$ , あるいは  $S_{1,2}$ ) を補足する副計画に従い、情報を提示する発話 ( $S_{3,0}$ ) を生成する。

このように、想定されるユーザの反応を織り込んだ計画によって対話制御を行うことで、素早く効率的に

ソチオリンピック、／フィギュアスケート男子の／金メダリスト、／羽生結弦選手が／腰の／痛みの／ため、／今シーズンの／初戦と／して、／来月フィンランドで／出場を／予定していた／国際大会を／欠場することになりました。

図 6: 文節単位に分割された文

ユーザの理解や知識に見合った情報を伝達することが可能となる。

なお、ユーザの反応として、上記のカテゴリに含まれない、より明示的な質問により情報要求が提示された場合は、例外処理として質問応答型の対話制御に一時的に切り替えることを考えているが、その詳細は現在検討中である。

### 3.4 構造解析部

構造解析部では、対象とするニュース記事をもとに、発話計画を生成するために必要な情報を構造化する。すでに複数記事を対象とする場合の検討を進めているが、本稿では単一の記事を対象とする。

構造解析の目的は、ニュース記事が持つ情報をもとに発話計画を立てるための情報を抽出することにある。すでに示したように、発話計画はニュースの要点を伝える主計画とそれを補う副計画からなるので、構造解析部では主計画に含めるべき情報(以降、主情報)と、周辺情報と主情報の関係性を抽出するという課題がある。

**構造化:** 構造解析は、文節単位の係り受け関係をもとに行う。そのため、まず文を文節単位に分割し、係り受け解析を行う。本研究では、形態素解析と係り受け解析に、Juman<sup>4</sup>、KNP<sup>5</sup>をそれぞれ利用した。

例として、ウェブニュース記事<sup>6</sup>の一文を文節に分割した例を図6に示す。さらに、係り受け解析に基づいて、文節をノードとする依存構造木を作成する。図7に、図6をもとに生成した依存構造木を示す。図中では省略しているが、各アークは係り受けの関係属性を保持し、各ノードは当該の文節の文法的情報を持つ。例えば、「羽生結弦選手が」に対応するノードは、「人物、主題」という情報を持つ。

**主情報の抽出:** 図7中、太枠で囲まれた文節がこの文における主情報であり、主発話計画を構成する。主情報として、まず、対象文の主辞となる文末の述語文節と、それに対する必須格要素となる文節を選択する。次に

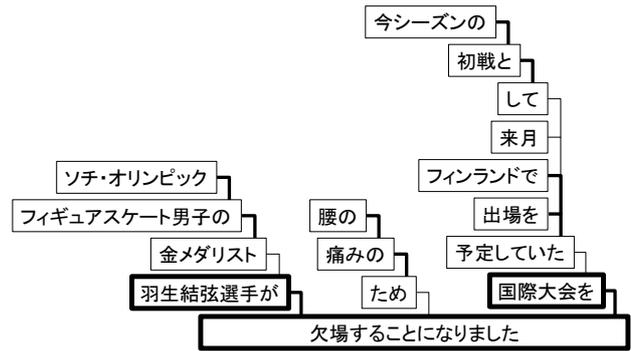


図 7: 図 6 の文に対する依存構造木

これらの文節に対して、次に説明する省略不可避性を有する文節を抽出する。図中では、太線によるエッジが省略不可避性を表す。

省略不可避性とは、「当該文節の係り先の文節が発話される場合は、当該文節が省略されてはならないこと」を表す。例えば「ため」という形式名詞からなる文節は、それだけでは特定の意味を持たないため、「痛みの」を省略できない。また、「痛みの」に対しての「腰の」が省略不可避であるかは微妙であるが、本研究では、野本による統計的な文圧縮の研究 [15] における依存構造木の「刈り込み」に準じた方法により、省略不可避性の判定を行う。主情報は主計画に組み込まれるため、主情報として選ばれた文節から文の最後の述語となる文節にたどり着くために通る経路上の文節の係り受け関係には、全て省略不可避性があるとする。

以上は、言わば非文法的な文の発話を回避するために必要な処理であるが、さらに、内容的に含めたほうが良いと思われる重要語も主情報に含める。このために、松尾らの手法 [18] をニュース記事の性質を踏まえて変更した重要語抽出処理を用いている [17]。

### 3.5 計画生成部

計画生成部では、構造解析部で得られた結果を用いて、3.3で述べた発話計画を作成する。ここでは、3.4での例をもとに生成した、図8に示した発話計画の例をもとに、主計画、副計画の生成について説明する。

**主計画の生成:** 主計画は、前節で述べた主情報により構成する。例に挙げた文では、「羽生結弦選手が」、「国際大会を」、「欠場することになりました」という3つの文節が主情報であった。

選ばれた文節をもとに語順にしたがって配列して作成した文を適切な長さに分割する。これは、文節を連結した文をそのまま読み上げてしまうと、ユーザが短い反応を挟む間を奪ってしまう可能性があるためであ

<sup>4</sup><http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

<sup>5</sup><http://nlp.ist.i.kyoto-u.ac.jp/index.php?KNP>

<sup>6</sup>NHK NEWSWEB, <http://www3.nhk.or.jp/news/>

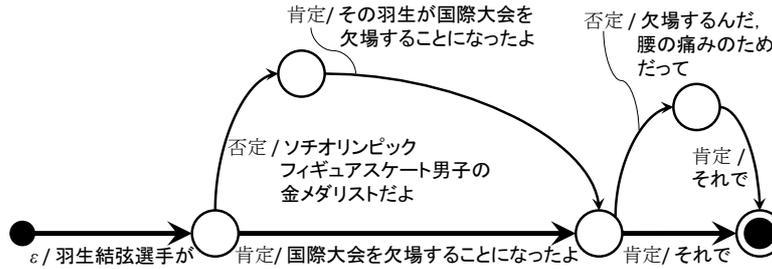


図 8: 発話計画の例

る。一方で、あまり短い単位毎に発話を区切ってユーザの反応をうかがうのは、ユーザに煩わしさを感じさせる可能性があり適切ではない。このため、簡単な文法的な規則を適用し、文節を連結することで適切な長さの発話を構成する。ここでは、海木らのポーズ挿入規則 [9] を参考に、次の条件を満たす箇所では分割し、それ以外では連結する。

- 当該文節の句が右枝分かれ、かつ、先行文節が左枝分かれ
- 当該文節が読点 (、) を含む

この規則を先の例に適用すると、「羽生結弦が」と「国際大会を欠場することになりました」という2つの発話となる。つまり、それぞれの区間をシステムが発話した後には一定のポーズがおかれるので、ユーザが反応を入れやすくなるを考える。

**副計画の生成** 構造化された情報の中で、主計画に含まれる文節に係っている文節のうち、省略不可避性がない文節が副計画を構成する。さらに主計画の場合と同様に、各文節から省略不可避性を有する文節を再帰的に辿って取得した文節を計画に含める。図7の例で、「国際大会を」に対する副計画を生成するとき、まず「国際大会を」に係る文節である「予定していた」が選択され、それに対する省略不可避性をもとに「出場を」「フィンランドで」が選択される。全体として「フィンランドで」「出場を」「予定していた」という文節が「国際大会を」の副計画を構成することになる。ここで選択されなかった「来月」などの文節は、「予定していた」に対する副計画となる。このように発話計画は階層性を持つ。副計画も、主計画と同様の規則で文の分割を行う。

## 4 対話例と課題

提案システムとユーザの対話例を図9に示す。システムが主計画に沿って発話をするので、ユーザは相槌をはさむ(肯定的応答 ACK)、あるいは黙っている

S: 羽生結弦選手が	[主計画]
U: うん	[肯定応答]
S: 国際大会を欠場することになったよ	[主計画]
U: え?	[否定応答]
S: 欠場するんだ、腰の痛みのためだ	[副計画]
U: へー	[肯定応答]
S: それで、日本スケート連盟によると	[主計画]
U: うん	[肯定応答]
S: 羽生選手は全治4週間と診断されたんだ	[主計画]
U: そうなんだ	[肯定応答]
S: スケート連盟は	[主計画]
S: 『症状は重くない。練習が全くできない状況ではない。中国大会に向けて	[主計画]
U: 中国大会?	[反復応答]
S: うん、グランプリシリーズのだよ	[副計画]
U: ふうん	[肯定応答]
S: それで、万全の状態を望むためだ』と	[主計画]
U: そうなんだ	[肯定応答]

図 9: 対話例。「S:」はシステム発話、「U:」はユーザ発話を表す。システム発話には主計画と副計画のどちらから生成されたか、ユーザ発話には反応がどの応答に分類されたか、をそれぞれ付与している。

だけでニュースの要点が得られる(受動的な情報消費)。また、ユーザの知らない言葉、興味を引いた単語、あるいは、理解できない表現が出て来たときには相槌(否定的応答 NACK)や聞き返しを行うことが想定される。このときシステムは、対応した副計画に沿った発話を生成することで情報を補足する(backgroundの情報行動)。このように、ユーザの状況(理解や興味の状態)に合わせてながらニュース記事の内容を音声によって効率よく伝達することが実現できる。

むしろ、課題も多く残されている。例えば、「国際大会を」と「欠場することになった」は、計画上では一つの発話としてまとめられている。二つの情報を一つ

にまとめた発話に対してユーザの聞き返しがあった場合、どちらに関する補足を行うかは自明ではない。また、一つの情報に複数の補足情報が存在することもある。従って、補足対象が定まったとしても、どの補足情報を伝えるべきかは別途決めなければならない。現状ではこれらの問題に対して、「発話中で最も後ろの情報に対して優先的に補足を行う」「時間や場所以外の補足情報を優先する」といった規則を適用しているが、より適切な情報提示のためには、どの補足情報が重要かといったことを考慮すべきである。また、用語や人物の説明といった補足情報は、それらの一般的な知名度や、個人の知識や嗜好などによって変化することを踏まえ、ユーザの反応に対する適切な補足情報の提示となるような発話計画を立てる手法を確立することが求められる。

## 5 関連研究

従来より、特定のタスク(交通案内や天気情報提供など)を対象として、ユーザからの発話に応じて情報を提示する質問応答型の対話システムが研究されてきた[8, 10, 14]。このようなシステムでは、ユーザの明示的な情報要求に応じて限定的な情報を確実に提供することに主眼が置かれていた。

近年、質問応答と組み合わせ、システム側からユーザに主体的に情報を提示する対話システムの研究も進み[19, 11, 12]、文書で表わされるような、まとまった量の情報提供を行う音声対話システムも提案されてきた[7, 20]。しかしながら、これらのシステムも基本的にはユーザの質問に対してシステムが回答を提示するという点では質問応答型の対話になっていると言ってよい。

ところで、ユーザにとって質問という行為は、システムの発話内容を理解した上で、問いかける内容を明示的に言語化する必要があるため、比較的負荷が高い。そのため、この種のシステムにおいてはユーザからの質問がなされにくく、システムは要約が提示することが主な機能となり、対話をとおして必要十分な情報を効率良く伝達することは困難であった。

これを解決するには、ユーザが質問を発しやすい状況を作り出すことが必要になる。音声対話の特性を考えれば、ここで言う質問には、言語表現を用いた明示的な情報要求だけでなく、相槌や聞き返し、相手の発話の一部の反復などの短い反応(本稿では即応的情報反応と呼んだ)を含めて考えるべきであることはすでに論じたとおりである。これらの反応や情報要求を認識するには、発話される語句の音声認識が必要であることは当然であるが、状況や態度を表出する手段としてのパラ言語情報の識別が重要となる。さらには、身

振り手振りや顔の動きなどのマルチモーダル情報も手がかりを与える。当研究グループでは、これまでもユーザの短い反応を韻律情報などを利用して認識するシステムを提案してきている[16, 4, 6]。

## 6 議論

**適用領域:** まず、提案したような音声対話による情報アクセス・情報伝達システムの適用領域についてであるが、やはり、音声対話メディアの特性から、いわゆるユーザが「手が離せない」状況が考えられる。このような条件に合致した適用場面としては、機器の運転中や、料理などの手作業中などが考えられる。一方、システムに対して音声で話しかけたり反応を返したりが行いやすいかという点も問題になる。この問題には、もちろんユーザの嗜好や特性も影響するが、ある種の擬人化エージェント的なインターフェースが有用である可能性も考えられる。また、ユーザの反応を引き出しやすいような、システム発話の生成[5]も有効な要素となる。

**情報アクセスシステムとしての位置づけ:** 通常のテキストを中心とする視覚的メディアを用いた情報アクセスの研究は盛んに行われている。しかしながら、いわゆるサーチエンジンを超えるようなポピュラリティを得ているシステム・インターフェースはほとんどないと言える。一方で、先に指摘したように、音声メディアが適している、あるいは、音声メディアしか使えないような利用状況が考えられる。ただし、記事や文書のようなまとまった情報を音声で伝達したり、ブラウジングすることには困難がある。そのような意味で、音声対話によってもたらされるインタラクションを導入することにより、「基本は受動的だけど、能動的なつっこみもできる」情報アクセスシステムを実現しようとする本研究の方向性には、これまでにはなかった可能性があると考えられる。

**対話システムとしての位置づけ:** 対話システムの分類の軸として、対話の主導権をシステムが持つか、ユーザが持つか、または両者の混合かというものがある[13]。本研究の音声対話システムは、現在の範囲においては、「基本はシステム主導で、必要に応じてユーザ主導」になる。ただしこれは、微妙なコントロールを短い音声反応で行える範囲に限定しての話である。この成約は、対話制御部を非常にシンプルなものにするのに貢献するが、今後もっと明示的で複雑なユーザの情報要求を扱おうとする場合、情報要求をシステム内部の情報検索過程に対応付けるための対話や、答えられない要求に対する対話などの複雑な対話制御が必要となる。もっ

とも、サービス・機能的な側面から、どの程度のことまで行うべきか、行えるかを定めるための検討も必要となる。

## 7 おわりに

ニュース記事から、ユーザの反応を想定した発話計画を作成し、それに従って対話を行うことで、ユーザから見れば必要十分な情報アクセス、システムから見れば効率的な情報伝達が行える音声対話システムを提案した。提案システムのユーザは、音声対話において自然と考えられる反応を返したり、情報要求を提示することにより、過不足のない情報アクセスを実現することができる。

今後は提案システムの枠組みを発展させ、さらに効率的で豊かなインタラクションを実現することを目指す。そのための課題として、パラ言語を利用したユーザの反応の認識、複数のニュース記事群を対象とした、より適切な情報内容の選択と構成、ユーザの反応を引き出すような対話の展開、親しみやすい発話音声の生成など個々の技術における精度や使い勝手の向上が挙げられる。

本システムは明確なタスク達成指向のシステムではなく、また、話を継続することを目的とする雑談システムでもない。また、今後は楽しく役立つ情報対話といった要素も加味していきたいと考えている。その意味で、評価の観点や方法論が現状では未確立である。よって、構築したシステムを実際の対話で評価しながら、これらを確認することも課題である。

## 参考文献

- [1] D.O. Case: *Looking for Information, A Survey of Research on Information Seeking, Needs, and Behavior, Second Edition*, Academic Press (2007)
- [2] S. Erdelez: Information encountering, In [3], pp.179–184, (2005)
- [3] K.E. Fisher, S. Erdelez, and L.E.F. EmKechinie (Eds), *Theories of Information Behavior*, Information Today, Inc. (2005)
- [4] S. Fujie, R. Miyake, and T. Kobayashi: Spoken dialogue system using recognition of user's feedback for rhythmic dialogue, *Proc. Int. Conf. Speech Prosody, OS2-4* (2006)
- [5] K. Iwata and T. Kobayashi: Speaker's intentions conveyed to listeners by sentence-final particles and their intonations in Japanese conversational speech, *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, pp. 6895–6899 (2013)
- [6] T. Kobayashi and S. Fujie: Conversational Robots: An Approach to conversation protocol issues that utilizes the paralinguistic information available in a robot-human setting, *Acoust.Sci. & Tech.*, Vol. 34, No. 2, pp. 64–72 (2013)
- [7] Y. C. Pan, H. Y. Lee, L. S. Lee: Interactive spoken document retrieval with suggested key terms ranked by a Markov decision process, *IEEE Trans. Audio, Speech, and Language Processing*, Vol. 20, No. 2, pp. 632–645 (2012)
- [8] S. Seneff, J. Polifroni: Dialogue management in the Mercury flight reservation system, *Proc. 2000 ANLP/NAACL Workshop on Conversational systems*, Vol. 3, pp. 11–16 (2000)
- [9] 海木延佳, 匂坂芳典: 局所的な句構造によるポーズ挿入規則化の検討, *信学論 (D-II)*, Vol. J79-D-II, No. 9, pp. 1455–1463 (1996)
- [10] 駒谷和範, 上野晋一, 河原達也, 奥乃 博: ユーザモデルを導入したパス運行情報案内システムの実験的評価, *情処学研報, SLP*, 2003.75, pp. 59–64 (2003)
- [11] 杉山 聡, 堂坂浩二, 川端 豪: 音声対話によるテキスト内容の伝達方法, *情処学論*, Vol. 41, No. 6, pp. 1883–1894 (2000)
- [12] 杉山弘晃, 南 泰浩: 情報提示対話を主導するシステムのためのユーザの潜在的情報要求の推定, *信学論 (A)*, Vol. J95-A, No. 1, pp. 74–84 (2012)
- [13] 中野幹生, 駒谷和範, 船越孝太郎, 中野有紀子: 対話システム, コロナ社 (2015)
- [14] 西村良太, 北岡教英, 中川聖一: 応答タイミングを考慮した雑談音声対話システム, *人工知能学研資, 言語・音声理解と対話処理研究会*, Vol. 46, pp. 21–26 (2006)
- [15] 野本忠司: 係り受け構造の刈り込みと CRF による文の要約, *言語処理学会年次大会*, pp. 488–491 (2008)
- [16] 藤江真也, 江尻 康, 菊池英明, 小林哲則: 肯定的/否定的発話態度の認識とその音声対話システムへの応用, *信学論 (D-II)*, Vol. J88-D-II, No. 3, pp. 489–498 (2005)
- [17] 藤江真也, 福岡維新, 麥田愛純, 高津弘明, 林 良彦, 小林哲則: 効率的な情報伝達を志向した音声対話システムの提案, *人工知能学会 第 74 回 言語・音声理解と対話処理研究会, SIG-SLUD-B501-02*. (2015)
- [18] 松尾 豊, 石塚 満: 語の共起の統計情報に基づく文書からのキーワード抽出アルゴリズム, *人工知能学論*, Vol. 17, No. 3, pp. 217–223 (2002)
- [19] 翠 輝久, 河原達也, 正司哲朗, 美濃導彦: 質問応答・情報推薦機能を備えた音声による情報案内システム, *情処学論*, Vol. 48, No. 12, pp. 3602–3611 (2007)
- [20] 吉野幸一郎, 河原達也: ユーザの焦点に適応的な雑談型音声情報案内システム, *人工知能学研資, 言語・音声理解と対話処理研究会*, Vol. 70, pp. 53–58 (2014)