

# 漸増的な情報補完機能を有する音声対話システム

A Spoken Dialogue System for Enhancing Interactive and Incremental News Content Delivery

福岡維新 \*1  
Ishin FUKUOKA

麥田愛純 \*1  
Asumi MUGITA

高津弘明 \*1  
Hiroaki TAKATSU

藤江真也 \*2\*1  
Shinya FUJIE

林良彦 \*1  
Yoshihiko HAYASHI

小林哲則 \*1  
Tetsunori KOBAYASHI

\*1 早稲田大学  
Waseda University

\*2 千葉工業大学  
Chiba Institute of Technology

This paper proposes a spoken dialogue system that can effectively deliver the content of a news article. The effectiveness is achieved by the system's ability to present the core content and further provide the details upon user's responses. To enable this interactive capability, the system generates two kinds of plans in advance: the primary plan for delivering the core content of an article, and the subsidiary plans that organizes detailed information that could answer user's possible inquiries. Moreover, the system is capable of promptly reacting to user's non-verbal responses which may signal his/her state of understanding and interest, achieving pleasant conversational rhythm in the dialogues.

## 1. はじめに

ユーザが行う暗黙的要求にテンポよく応えることで、ニュース記事のようなまとまった量の情報をユーザに、効率よく、過不足なく伝達する音声対話システムを提案する。

新聞などの視覚メディアでニュースを読む際には、斜め読みをして不要な情報を飛ばし、必要なところのみを拾うことができる。これに対し、ラジオなどの聴覚メディアでニュースを聞く際には、内容を最初から最後まで聞き続ける必要がある。興味や知識によってユーザが必要とする情報は異なるにも関わらず、全てを聞かなければならない聴覚メディアでの情報伝達は、フラストレーションの大きいものであった。本研究では、このような聴覚メディアにおけるまとまった量の情報伝達の問題に対し、対話的要素を導入することで解決を図る。

古く、音声対話システムとしては、飛行機のフライト案内 [Senef 00] や京都の観光バス案内 [駒谷 03]、天気案内 [西村 06] など、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。これらのシステムでは、少量で断片的な情報を、ユーザの明示的な要求に応じて提供するとどまっていた。

近年、ユーザ主導の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 [翠 07], [杉山 12], ニュースの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 [Pan 12], [吉野 14]。

ニュースの伝達というタスクでこの枠組みを利用すると、ニュース記事を要約してユーザに伝えた後、ユーザから質問を受け付け、それに応じた情報をシステムが提供することが可能となる。しかし、質問のような明示的な情報要求は、与えられた情報をユーザが理解した上で初めて発することが可能となる行為であり、ユーザにとって比較的負荷が高い。インタラクションの敷居が高ければ、ユーザが質問を行わず単に要約が読み上げられるのみとなって、ユーザにとって有用な情報が伝えられることなくシステムによる情報伝達が終わってしまう可能性が生じる。

一方、日常会話において、常に明示的な情報要求が行われるわけではない。藤江らのレストラン推薦システム [藤江 05] でユーザの意思の推定にパラ言語情報が用いられているように、会話では表層的な言語情報以外にも多くの情報がやり取りされる。間投詞による聞き返しや相手の発話の復唱といった発話は、与えられた情報を受理できなかったことを表す暗黙的な情報要求と捉えられる。暗黙的な情報要求を受け取った相手は、直前の会話の流れや発話の内容、情報要求の内容から適当な情報を選び補足することでそれに応えることができる。

聞き返しや復唱は相手の発話の表層的な理解・判断に基づいて行われる負荷の低い行為である。このためユーザが行う暗黙的な情報要求によって提示する情報をテンポよく補足することができれば、システムとユーザのインタラクションが活性化しユーザが情報を効率よく受け取ることが可能になると考えられる。しかし、ここでの補足を円滑に行うためには、ユーザが行う様々な暗黙的な情報要求を予め想定し、それに応じたシステムの応答を事前に用意しておく必要がある。

そこで、ニュース記事の解析に基づいてユーザの暗黙的な情報要求を想定した発話計画を生成し、それに従って対話を行うことでインタラクションを活性化させニュース記事の伝達を効率よく行う音声対話システムを提案する。ニュースを伝えるというタスクの性質上、システムが発話する内容がある程度想定できることや、ユーザの理解を促進するためには対話のテンポを損なわない必要があることなどから、対話計画を予め用意することは有用である。

本研究ではニュース記事の文章から、発話計画に必要な情報の構造化を行い、そこから発話計画を生成する。発話計画はニュースの要点となる情報を伝える主計画と、それらを補足する副計画からなる。システムはまず主計画に従ってユーザに情報を提供し、ユーザの暗黙的要求に応じて適宜副計画へ展開し補足を行う。これによりユーザはニュースの要点を押さえつつ自分が求める情報を受け取ることが可能となる。

## 2. システム構成

対話システムの概要を図 1 に示す。システムは構造解析部、計画生成部、対話制御部、音声認識器、音声合成器の 5 つのモ

連絡先: 早稲田大学基幹理工学研究所, 東京都新宿区喜久井町 27, fukuoka@pcl.cs.waseda.ac.jp

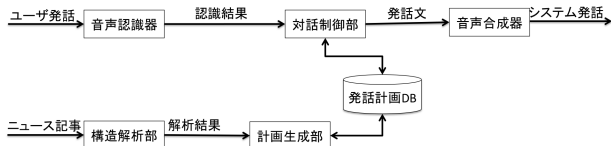


図 1: 対話システム構成図

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 2: 文節単位に分割された文章

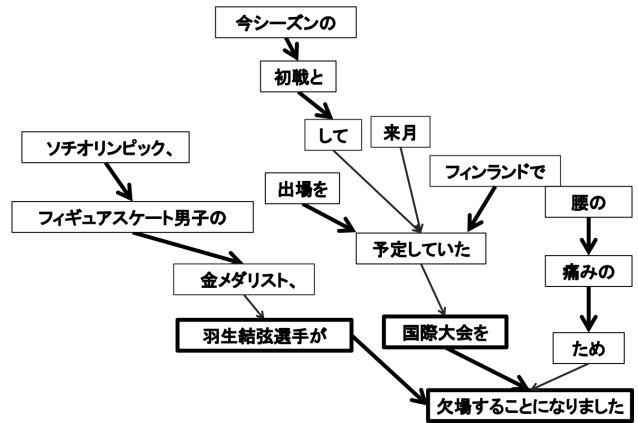


図 3: 構造木

ジュールで構成されている。計画生成部で作られた発話計画は発話計画 DB に保存され、対話制御部ではそれを受け取って対話を行う。音声認識器は、ATR-Trek 製のものを使用した。また、音声合成器は Open JTalk\*1 を基に開発したものをを用いている。

### 3. 構造解析部

構造解析部では、発話計画を生成するために必要な情報を構造化する。本来、記事全体の内容を再構成して発話計画を立てることが望まれるが、ここでは、記事内の文毎に独立して発話計画を立てる。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ため、主計画に含める情報(主情報)の選択、主情報に対する補足関係の把握という二つの課題が存在する。本研究では、文に含まれる語の係り受け関係を用いることで情報の構造化を行う。主情報の選択や補足関係の把握に関しては、ニュース記事の性質を踏まえた重要語抽出手法を検討した。形態素解析、係り受け解析には Juman\*2, KNP\*3 を使用した。

#### 3.1 構造化

構造化を行うために、文を文節単位に分割する。例として、ウェブニュース記事\*4 の 1 文を分割した例を図 2 に示す。

係り受け解析に基づいて文節単位に分割した文の構造木を作成する。文節一つ一つが構造木のノードとなる。図 3 に、図 2 を元に作成した構造木を示す。各文節は係り先の文節と、自身の文法的情報を持つ。例えば、「羽生結弦選手が」という文節は「欠場することになりました」という文節に係っていて、「人物、主題」という情報を持つ。

#### 3.2 重要語と省略不可避性

図 3 中、太枠で囲まれた文節がこの文における主情報である。これは、文における最後の述語と、ニュース記事において重要だとされた文節(後述)が選択される。

太い矢印は、係先の文節に対して係元の文節が省略不可避性を持つことを表す。省略不可避性とは、「係先の文節が発話される際に、省略されてはならないこと」である。例えば「ため」という文節は、それだけを発話したのでは意味を持たず、「腰の」「痛みの」を省略してはならない。従って、「腰の」は「痛みの」に、「痛みの」は「ため」に対して省略不可避性を持つ。

つ。省略不可避性の判定は、野本の研究 [野本 08] で述べられている係り受け構造の「刈り込み」による文の要約問題に近い問題である。

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は全て省略不可避性有りとする。図 3 の例では「羽生結弦選手が」「国際大会を」は主情報であるため、最後の述語である「欠場することになりました」に対して省略不可避性を有する。

重要語の抽出や省略不可避性の判定は、現状では自動化されておらず人手で行っている。このうち、重要語の抽出に関しては次節に述べる検討を行った。

#### 3.3 重要語抽出方法の検討

主情報に含める文節の選択のため、語の統計的な情報を利用してニュース記事から重要語を抽出する方法を検討した。重要語には「事実を伝える語」と「事実の特徴を伝える語」があると考えた。後者は、一般的には重要語として捉えにくいですが、ニュースを伝えるには欠かせない語である。「事実を伝える語」の抽出については、松尾らの手法 [松尾 02] を以下に示すニュース記事の性質を踏まえて改良し、アンケートにより評価した。「事実の特徴を伝える語」についてはアンケートを基に考察を行った。

松尾らの手法では、文書の傾向を示す頻出語との共起の偏りを  $\chi^2$  検定で測り、語の重要度としている。 $G$  を文書に出現する頻出語群として、語  $w$  の頻出語  $g \in G$  との共起の総数を  $n_w$ 、頻出語の生起確率を  $p_g$ 、語  $w$  と頻出語の実際の共起頻度を  $\text{freq}(w, g)$  として、 $\chi^2$  値は次式のように計算される。

$$\chi^2 = \sum_{g \in G} \frac{(\text{freq}(w, g) - n_w p_g)^2}{n_w p_g} \quad (1)$$

多くのニュース記事は次の性質を持つ。

- 1 文目がリード文として記事の要約の役割を持つ
- 記事内で語の言い換えが多い

これらを踏まえ、ニュース記事の 1 文目の語を頻出語として利用する、語をクラスタリングするという 2 つの改良を行った。クラスタリングは、文字列の類似度を表す Jaro-Winkler 距離 [Winkler 90] を用いて行う。その後、生成されたクラスタの中から頻出語群  $G_C$  を選出する。この手法では  $n_w$  の代わりに

\*1 <http://open-jtalk.sp.nitech.ac.jp/>

\*2 <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

\*3 <http://nlp.ist.i.kyoto-u.ac.jp/index.php?KNP>

\*4 NHK NEWSWEB, <http://www3.nhk.or.jp/news/>

表 1: 頻出語の選出に 1 文目を利用する手法

評価軸	提案手法	従来手法	TF-IDF
適合率	0.567	0.515	0.564
再現率	0.547	0.479	0.511
F 値	0.562	0.506	0.551

表 2: 被験者が選んだ語の品詞ごとの割合

	名詞	動詞	形容詞	副詞
指示 1	0.394	0.249	0.247	0.276
指示 2	0.208	0.057	0.154	0.457

して  $n_c$  を, あるクラスタに含まれる語  $w$  が出現する 文の語数の総和とする.  $G_C$  は, 1 文目中出现する語を含むクラスタを,  $n_c$  が大きい順に取り出し, 含まれる語数が全体の  $1/6$  程度になるように構成する.

### 3.4 重要語抽出に関する評価

「事実を伝える語」と「事実の特徴を伝える語」を調べるためにアンケートを行い, その結果に対して提案手法と従来手法を比較した. ウェブニュースからランダムに選んだ 9 記事を用い, 被験者 4 人を対象に行った. 次の指示に従って, 各記事から単語を選択させ, 2 人以上が選択した語を重要語とした.

指示 1 ニュースのタイトルの内容を伝えるために必要と考えられる語を選んでください (10 語程度)

指示 2 自分がニュースを伝える側として, ニュースの特徴をより伝えやすくするために強調したい語を選んでください (5 語程度)

指示 1 で選択された重要語は「事実を伝える語」, 指示 2 で選択された重要語は「事実の特徴を伝える語」とした.

記事に対して, 提案手法, 松尾らの手法, 一般的な重要語抽出方法である TF-IDF を用いた手法の, 3 通りの手法の比較を行った. 重みづけされた単語 (クラスタ) の上位 10 個を抽出し, 指示 1 で選ばれた語との適合率, 再現率, F 値の 9 記事における平均を求めた. 結果を表 1 に示す. いずれの値も提案手法が他の手法を上回った.

「事実の特徴を伝える語」についてアンケートで得られた回答から考察を行う. 表 2 に指示 1 および指示 2 で被験者が選んだ語の品詞毎の割合を示す. この表から, 各品詞の選ばれやすさの傾向が指示 1 と 2 で異なることがわかる. 指示 2 では, 副詞が多く選ばれている. 「事実を伝える語」としては「何が・何を・どうした」という情報が選ばれやすいのに対し, 「事実の特徴を伝える語」としては「どのように」という情報が選ばれやすいことがわかる. これより, 「事実を伝える語」の抽出で用いた手法に品詞情報や動詞との係り受け関係を利用することで「事実の特徴を伝える語」の抽出ができることが期待できる.

## 4. 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発話計画を作成する. 発話計画は, ユーザ発話を入力, システム発話を出力として遷移する状態遷移構造を持つ. 例を図 4 に示す. ユーザの発話として, 以下に示す応答を想定し, 発話計画に組み込む.

肯定応答 「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発話

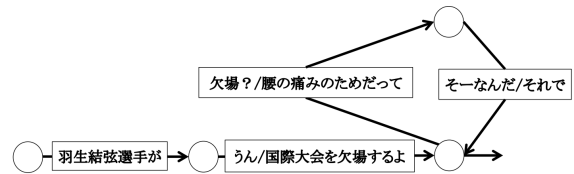


図 4: 発話計画の状態遷移

否定応答 「えっ?」といった, システムの発話進行に否定的な態度を表す発話

反復応答 ユーザがシステムの発話の一部を反復する発話

否定応答と反復応答のはたつきは似ているが, 反復応答はその内容によってユーザの焦点 (何がわからないのか) が具体的に示される点で否定応答と異なる.

システムは肯定応答を受け取る限り主計画に従って発話を進める. 否定応答や反復応答のような暗黙的な情報要求を受けると副計画を展開して発話する. 副計画に従って発話しているときにユーザから肯定応答が返ってくれば, 要求を満たすことができたと判断し主計画に復帰して発話を進める.

### 4.1 主計画の生成

主計画には, ユーザの肯定応答を受け取る限りシステムが述べるための, ニュースの要点を含める. ここに含まれるのは 3.1 で述べた主情報に他ならない. 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 つの文節を選択する.

次に, 選ばれた文節をもとに作成した文を, 適切な長さに分割する. 文節を連結した文をそのまま読み上げると, ユーザが暗黙的な情報要求を挟む間を奪ってしまうことになり, インタラクションの機会を損失することになる. 一方で, 短い単位毎に発話を区切って反応をうかがうのは, ユーザに煩わしさを感じさせかねない. このため, 文法的な規則を適用し, 文節を連結することで計画に適した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 [海木 96] を参考に次の条件を満たす箇所ではポーズを設け, それ以外では文節を連結する.

- 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, 「国際大会を」と「欠場することになりました」は結合され, 最終的に「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.

### 4.2 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が, 副計画となる. 主計画と同様に, 各文節から省略不可避性を有する文節を再帰的に取得した文節を計画に含める. 図 3 の例で「国際大会を」に対する副計画を生成する場合, 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」や「今シーズンの」「初戦と」「して」という文節は, 「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.

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んだよ」(主計画)
U	「え？」(否定応答)
S	「欠場するんだ、腰の痛みのためだって。」 (副計画への展開)
U	「へー」(肯定応答)
S	「それで....」(主計画への復帰)

図 5: 発話計画による対話例

副計画へは、ユーザの否定応答または反復応答という暗黙的な情報要求を受けて主計画から展開する。そこで、システムは補足となる情報を提示し新しい状態へ遷移する。補足後にユーザから肯定応答が得られると、元の発話計画へ復帰する。図 4 の計画に従った対話例を図 5 に示す。

## 5. 対話例と課題

提案システムとユーザとの対話例を図 6 に示す。ユーザは、聞き返しや復唱などを挟まなくてもニュースの要点を得られる。更に、知らない言葉や理解できない表現が出て来たときに暗黙的な情報要求を挟むことで、適宜システムから補足情報を提供してもらうことが可能となった。

一方で、課題も残された。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が、それぞれ別の補足情報が存在しており、情報要求に対してどちらに関する補足を行うかを決定しなければならない。また、一つの文節に複数の文節が係ることがあるため、補足対象の文節が決まったとしても、どの補足情報を提示す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、「発話中で後ろの文節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」といった規則で処理を行っているが、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の重要度は、それらの一般的な知名度や、ユーザ個人の知識や嗜好などによって変化する。これらを踏まえ、ユーザの暗黙的な情報要求に対して、適切な補足情報を提示できるように計画を立てる方法を確立することが今後の大きな課題である。

## 6. まとめ

ニュース記事から、ユーザの暗黙的な情報要求を想定した発話計画を作成し、それに沿って対話を行うことで効率的な情報伝達を行うシステムを提案した。提案システムは、ユーザにとって負荷の軽いインタラクションで、過不足のない情報伝達を、聴覚メディアによって初めて実現したものであると考える。

今後は提案システムの枠組みを踏襲し、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、ユーザ発話からの暗黙的な情報要求の認識、重要語抽出など個々の技術における精度の向上が課題として挙げられる。また、構築したシステムを実際にユーザと対話させることで評価を行うことも課題である。

## 参考文献

[Seneff 00] S. Seneff, J. Polifroni: Dialogue management in the Mercury flight reservation system, *Proc. 2000*

S	「羽生結弦選手が...」
U	「うん。」
S	「国際大会を欠場することになったんだよ。」
U	「え？」
S	「欠場するんだ、腰の痛みのためだって。」
U	「へー。」
S	「それで、日本スケート連盟によると...」
U	「うん。」
S	「羽生選手は全治 4 週間と診断されたんだって。」
U	「そうなんだ。」
S	「スケート連盟は...」
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...』」
U	「中国大会？」
S	「うん、グランプリシリーズのだよ。」
U	「ふーん。」
S	「それで、万全の状態ですべて望むためだ」としているよ。」
U	「がんばってほしいね。」

図 6: 対話例

*ANLP/NAACL Workshop on Conversational systems*, 3, 11–16 (2000).

- [駒谷 03] 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バ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, 59–64 (2003).
- [西村 06] 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*, 46, 21–26 (2006).
- [藤江 05] 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)*, J88-D-II(3), 489–498 (2005).
- [翠 07] 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, *情処学研報*, 48(12), 3602–3611 (2007).
- [杉山 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な情報要求の推定, *信学論 (A)*, J95-A(1), 74–84 (2012).
- [Pan 12] 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*, 20(2), 632–645 (2012).
- [吉野 14] 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*, 70, 53–58 (2014).
- [野本 08] 野本 忠司: 係り受け構造の刈り込みと CRF による文の要約, *言語処理学会第 14 回年次大会*, 488–491 (2008).
- [松尾 02] 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*, 17(3), 217–223 (2002).
- [Winkler 90] W. E. Winkler: String comparator metrics and enhanced decision rules in the Fellegi-Sunter model of record linkage, *Proc. the Section on Survey Research Methods, American Statistical Association*, 354–359 (1990).
- [海木 96] 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)*, J79-D-II(9), 1455–1463 (1996).