

早稲田大学大学院 基幹理工学研究科

博士論文概要

論文題目

快適な情報享受を可能とする
音声対話システム

A Spoken Dialogue System for
Enabling Comfortable Information
Acquisition and Consumption

申請者

高津	弘明
Hiroaki	TAKATSU

情報理工・情報通信専攻 知覚情報システム研究

2018年12月

快適な情報享受は、質問応答のような意図性の高い能動的な情報行動だけでも、文書閲覧・聞き流しのような意図性のない受動的な情報行動だけでも実現できず、これらの情報行動のモードを自由に行き来することにより達成される。本研究では、意図性の変化を捉えるのに音声対話が適しているという立場から、ニュース記事を例題としてまとまりのある情報を利用者が快適に享受するための枠組みについて検討し、意図性の異なる多様な情報行動を高頻度に即応性高く切り替えながら情報にアクセスできる会話システムを開発した。

ここで提案するシステムは、主計画と副計画と呼ぶ、二種類の発話計画に沿って会話を進める。主計画は、ニュース記事など、伝達対象の文書を要約し、口語化することで作られるもので、文書の要点を伝えることを目的としている。ユーザーが受け身で聴いている限り、システムは主計画に従った情報伝達を行う。一方、副計画は、会話進行の各時点におけるユーザーの反応予測に基づいて、それらに対する応答を計画したもので、ユーザーの興味・理解度に応じた補足情報の提示を目的としている。ユーザーが情報要求を示してきた場合、システムは副計画に遷移し、補足説明を行う。これらの発話計画を事前に準備することで、目的とする即応性の高い円滑な会話インタラクションを実現した。システムの評価のために、情報伝達効率（**EoIT; Efficiency of Information Transfer**）という指標を提案した。EoITは、一連の会話によって、興味のある文を過不足なく効率的に伝達できたかどうかを図るための指標である。この指標を用いて評価を行った結果、要約のみの情報伝達や従来の一問一答型の対話システムに比べて、提案システムの方が情報伝達効率が顕著に高く、快適な情報享受が可能であることを確認した。

本論文の構成は以下の通りである。

第1章では、情報行動および音声の揮発性・同時性の観点から望まれる会話システムについて説明し、提案する会話システムの概要について述べる。

第2章では、会話によるニュース記事伝達のための要約手法について説明する。本システムでは、円滑な情報伝達を実現するために事前に発話計画を作成しておく。発話計画は主計画と副計画から構成される。主計画はニュースの要点に関する発話計画であり、ユーザーが受け身の姿勢で聴いている限り、システムはこの主計画に沿って発話する。主計画はニュース記事を要約した内容を基盤に生成される。ここで、要約は重要文抽出、整列、文圧縮の3段階の処理を経て行われる。重要文抽出は、文書から重要な文（核となる情報・見出し的な内容を含む文）を数文抽出する問題である。重要文抽出の手法として最大被覆モデルを利用した。最大被覆モデルは、制限要約長に収まる範囲で重要な単語をできるだけ多く被覆するような文集合を抽出する手法である。そのため、単語の重要度の与え方が肝となる。従来、単語の重要度としては、頻度やTF-IDF、ロジスティック回帰で推定した値などが用いられてきた。本研究では、BERT（Bidirectional Encoder

Representations from Transformers) を基本構造とした双方向モデルにより，文脈を考慮して単語の重要度を推定する手法を提案した．要約コーパスを作成し，評価実験を行った結果，従来の文脈を用いない手法よりも，提案モデルで推定した単語重要度を用いたときの方が高い ROUGE-1 を示した．整列は，重要文抽出で抽出した文の提示順序を決める問題である．要約の対象が単一記事である場合は，記事の出現順で問題ないことを確認した．複数記事からの要約である場合は，ある記事の重要文と別記事の重要文前後の文章との類似度に基づいて前後関係を決める手法を提案し，有効性を確認した．文圧縮は，文自体を短く要約する問題である．文圧縮の問題を文節レベルの系列ラベリングの問題として定式化し，BERT に基づく文圧縮モデルを提案した．実験の結果，単純な双方向 LSTM モデルに比べて，提案モデルの方が高い精度を示した．

第 3 章では，会話によるニュース記事伝達のための口語化手法について説明する．ニュース記事は書き言葉で書かれているため，会話で伝えるのに適した話し言葉に変換する必要がある．本研究では，友達と会話するようなカジュアルな感覚で利用できるシステムを目指している．そこで，敬体（ですます調）ではなく，くだけた表現に書き換える．中性的なキャラクターを想定し，ルールに基づく書き換えを行った．口調は「情報のなわ張り理論」に基づき伝聞口調に書き換える．その他，同格の展開，連体修飾節の展開，長い重文の分割などの書き換えを行う．口語化によって生成された発話文の言葉遣いに違和感がないか確認するため，主観評価実験を行った．実験の結果，システムが生成した発話系列に対する評価結果が，人間が口語化したものに対する評価結果と同程度であることを確認した．

第 4 章では，副計画の生成方法について説明する．要約で省かれた補足情報や予想される質問に対する回答を事前に副計画として用意しておくことで，ユーザーの情報要求に対して迅速な応答が可能となる．Wikipedia などを用いた用語の定義説明や，固有表現抽出結果を利用した日時，場所，人物などに関する回答計画，「ので」などの手がかり表現を利用した原因・理由に関する回答計画，含意・矛盾表現を用いた真偽判定，「何それ」といった抽象的な質問に対する回答計画を行う．質問応答として「何それ」に特化したシステムであることを被験者に伝え，応答の適切さについて評価したところ，68% 程度の応答能力があることを確認した．副計画の事前設計による迅速な応答の有効性については 7 章の評価で確認した．

第 5 章では，システム発話の文脈を考慮した発話意図認識手法について説明する．より高い情報伝達効率を実現するためには，ユーザーからのフィードバックを正しく理解することが重要となる．従来，発話意図の識別は，F0 などの韻律情報に基づいて行われてきた．しかしながら，フィードバックに込められる意図を

ユーザー発話の情報のみから判断するのは難しい。そこで、本研究では、システム発話の文脈を考慮した発話意図認識モデルを提案した。

本システムを用いて収集した対話データに対して発話意図のアノテーションを行ったデータセットを使用し、評価実験を行った結果、システム発話の文脈を考慮したモデルの方が考慮しないモデルに比べて高い識別精度を示した。

第6章では、談話構造上の発話の役割に応じたメリハリのある話し方を可能とする音声合成手法と間（ま）の推定手法について説明する。ニュースのようなまとまった量の情報を伝える会話システムでは、重要な情報がユーザーに伝わる話し方が求められる。この問題は、音声合成の分野では、これまで一文内におけるプロミネンスの付与という形で扱われてきた。しかしながら、発話の強調に影響を与える韻律は、談話構造に応じても変化することが古くから指摘されており、一文より長い単位で韻律や間（ま）を制御する必要がある。そこで、本研究では談話構造上の文の役割を、核（最も伝えたい内容を含む文）、前置き（核を伝えるためのリード文）、補足（情報を補足する文）の3つに分類し、これらをDNN音声合成の補助情報として利用することで、発話系列全体で韻律や間（ま）の制御が可能な音声合成システムを開発した。実験は次の3つの観点で行った。(1)情報を伝える話し方としてのふさわしさに関する主観評価。(2)ニュースの理解度に関する客観評価。(3)質問しやすさに関する評価。実験の結果、従来の1文単位の読み上げ音声で学習したモデルの合成音声よりも、情報伝達を意識させ段落単位で発話させた音声で学習したモデルの合成音声の方が、情報伝達のための発話としてふさわしいと感じるとともに、ニュースの理解度という面においても優れていることが分かった。さらに、発話の段落における位置や談話構造上の役割を補助情報に用いて合成することにより、メリハリのある聴きやすい話し方になることが分かった。また、発話の間（ま）に関しても、提案モデルで推定した間（ま）で発話した方が、短い単調な間（ま）で発話するよりも、聴きやすく頭に入りやすく、さらに質問しやすい話し方を実現できることが確認できた。

第7章では、提案システムと従来システムの比較実験結果について説明する。主計画と副計画からなるシナリオに基づいて、割り込みを許容しながら、質問に即座に回答できる点が本システムの特徴の一つである。このような即時応答機構を持つシナリオドリブンな提案システムと、要約を伝えた後にQAのモードに移行する従来の一問一答型の対話システムを比較評価した。その結果、情報伝達効率（**EoIT; Efficiency of Information Transfer**）という、提示できたほしい情報の被覆率と提示しなかったいらない情報の除外率の調和平均で定義される提案尺度において、提案システムの方が優れていることを確認した。また、本システムを継続して利用したいかどうかについて、アンケート調査を行ったところ、過半数が継続して利用したいと回答した。

早稲田大学 博士（工学） 学位申請 研究業績書

氏名 高津 弘明 印

(2019年2月現在)

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む）
論文	○高津弘明, 福岡維新, 藤江真也, 岩田和彦, 小林哲則: “会話によるニュース記事伝達のための音声合成,” 人工知能学会論文誌, Vol. 34, No. 2, pp. 1-15, 2019年3月. (掲載決定)
論文	○高津弘明, 福岡維新, 藤江真也, 林良彦, 小林哲則: “意図性の異なる多様な情報行動を可能とする音声対話システム,” 人工知能学会論文誌, Vol. 33, No. 1, pp. 1-24, 2018年1月.
講演	Katsuya Yokoyama, <u>Hiroaki Takatsu</u> , Hiroshi Honda, Shinya Fujie, and Tetsunori Kobayashi: “Investigation of users’ short responses in actual conversation system and automatic recognition of their intentions,” in Proceedings of the 2018 IEEE Workshop on Spoken Language Technology, Dec. 2018.
講演	横山勝矢, <u>高津弘明</u> , 本田裕, 藤江真也, 小林哲則: “会話によるニュース記事伝達のための発話意図の分類と認識,” 第123回音声言語情報処理研究会, pp. 1-8, 2018年7月.
講演	<u>高津弘明</u> , 横山勝矢, 本田裕, 藤江真也, 林良彦, 小林哲則: “会話によるニュース記事伝達のための発話意図理解,” 人工知能学会第32回全国大会論文集, 4Pin1-29, pp. 1-4, 2018年6月.
講演	横山勝矢, <u>高津弘明</u> , 本田裕, 藤江真也, 林良彦, 小林哲則: “会話によるニュース記事伝達のための発話意図分類とデータベースの構築,” 人工知能学会第32回全国大会論文集, 3Pin1-23, pp. 1-4, 2018年6月.
講演	<u>高津弘明</u> , 横山勝矢, 本田裕, 藤江真也, 林良彦, 小林哲則: “会話によるニュース記事伝達のための間の調整,” 言語処理学会第24回年次大会発表論文集, pp. 212-215, 2018年3月.
講演	<u>高津弘明</u> , 福岡維新, 藤江真也, 林良彦, 小林哲則: “会話によるニュース記事伝達のための文圧縮,” 言語処理学会第23回年次大会発表論文集, pp. 154-157, 2017年3月.
講演	<u>高津弘明</u> , 福岡維新, 藤江真也, 林良彦, 小林哲則: “意図性の異なる多様な情報行動を可能とする音声対話システム,” 人工知能学会第30回全国大会論文集, 204-0S-23a-3in2, pp. 1-4, 2016年6月.
講演	福岡維新, <u>高津弘明</u> , 藤江真也, 林良彦, 小林哲則: “会話による情報伝達における発話系列の韻律分析,” 人工知能学会第30回全国大会論文集, 2P1-5in1, pp. 1-4, 2016年6月.
講演	<u>高津弘明</u> , 福岡維新, 藤江真也, 林良彦, 小林哲則: “快適な情報享受を可能とする音声対話システム,” 言語処理学会第22回年次大会発表論文集, pp. 302-305, 2016年3月.

早稲田大学 博士（工学） 学位申請 研究業績書

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む）
講演	高津弘明, 福岡維新, 藤江真也, 林良彦, 小林哲則: “会話によるニュース記事伝達のための口語化における述語の書き換え,” 言語処理学会第 22 回年次大会発表論文集, pp. 1033-1036, 2016 年 3 月.
講演	高津弘明, 福岡維新, 藤江真也, 林良彦, 小林哲則: “会話によるニュース記事伝達のための情報選択,” 言語処理学会第 22 回年次大会発表論文集, pp. 199-202, 2016 年 3 月.
講演	高津弘明, 小林哲則: “多人数会話における議論支援のための状態把握,” 言語処理学会第 22 回年次大会発表論文集, pp. 497-500, 2016 年 3 月.
講演	Shinya Fujie, Ishin Fukuoka, Asumi Mugita, <u>Hiroaki Takatsu</u> , Yoshihiko Hayashi, and Tetsunori Kobayashi: “A spoken dialog system for coordinating information consumption and exploration,” in Proceedings of the 1st ACM SIGIR Conference on Human Information Interaction and Retrieval, pp. 253-256, Mar. 2016.
講演	林良彦, 藤江真也, 福岡維新, <u>高津弘明</u> , 小林哲則: “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,” 人工知能学会第 11 回インタラクティブ情報アクセスと可視化マイニング研究会, pp. 25-32, 2015 年 11 月. (研究会優秀賞受賞)
講演	藤江真也, 福岡維新, 麥田愛純, <u>高津弘明</u> , 林良彦, 小林哲則: “聞き手の反応を考慮した発話計画を用いた音声対話システム,” 日本音響学会 2015 年秋季研究発表会, pp. 1-4, 2015 年 9 月.
講演	藤江真也, 福岡維新, 麥田愛純, <u>高津弘明</u> , 林良彦, 小林哲則: “効率的な情報伝達を指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, pp. 7-12, 2015 年 7 月.
講演	福岡維新, 麥田愛純, <u>高津弘明</u> , 藤江真也, 林良彦, 小林哲則: “漸増的な情報補完機能を有する音声対話システム,” 人工知能学会第 29 回全国大会論文集, 2L4-0S-07a, pp. 1-4, 2015 年 6 月.
講演	<u>高津弘明</u> , 小林哲則: “分散表現を用いた動詞・フレーズの含意関係認識,” 言語処理学会第 21 回年次大会発表論文集, pp. 928-931, 2015 年 3 月.
講演	<u>高津弘明</u> , 小林哲則: “対話エージェントのための性格モデル,” 言語処理学会第 21 回年次大会発表論文集, pp. 191-194, 2015 年 3 月.