

ブリコラージュ風発話生成システム〈PoKeBo Talk〉の提案

Proposal for a Bricolage-Style Utterance Generation System “PoKeBo Talk”

佐々木 冬威¹ 長谷川 孔明¹ 大島 直樹² 岡田 美智男¹

Toi Sasaki¹, Komei Hasegawa¹, Naoki Ohshima¹, and Michio Okada¹

¹ 豊橋技術科学大学 情報・知能工学系

¹ Department of Computer Science and Engineering, Toyohashi University of Technology

² 豊橋技術科学大学 エレクトロニクス先端融合研究所

² Electronics-Inspired Interdisciplinary Research Institute, Toyohashi University of Technology

Abstract: Our daily speech is not planned or rationalized, but rather is a haphazard process of making things up as we go along. In this presentation, we will introduce the concept and prototype of PoKeBo Talk, a bricolage-like utterance generation system that defines a style of speech production in which the speaker gathers fragments of words as they wish and produces speech.

1. はじめに

誰かに思いを伝えたい、けれどもうまく言葉にすることができない。それでも誰かとコミュニケーションを取りたいと覚えることもある。もしかしたら、そんな瞬間を日常生活の中で体験したことがあるのではないだろうか。私たちの日常での発話は必ずしも計画的、合理的になされるものだけでなく、むしろその場を取り繕うような行きあたりばったりなことも多いのではないだろうか。

例えば、「今日は遊園地に出かけようと思う」という直接的な言葉を聞いたらどうだろう。私たちはその一方的な発話を何も考えずに受け取ってしまう可能性がある。一方で、「今日ね」「出かけようと思う」「あそこ、えーと」「あの遊園地に！」のように断片的な発話ではどうだろうか。私たちはこの断片的な発話を再構築しながら、そのメッセージを受け取ることができる。外国人の片言な日本語も同様に受け取って解釈をしている。このように私たちの日常の会話では、十分に整えられた発話文でなくとも理解し、コミュニケーションが成り立つのである。むしろ断片的な発話のほうが伝えようとする気持ちまでも感じ、話し手に寄り添い合えるのではないだろうか。

本研究では、話し手が思うまま、断片的な発話を自ら寄せ集めて発話を生み出す様式のことを、「ブリコラージュ(bricolage)」と「モノログ(monologue)」の合成語として「ブリコログ(bricologue)」と呼んでいる。本稿では、この「ブリコログ」の発話様式を備えた発話生成システム〈PoKeBo Talk〉の基本

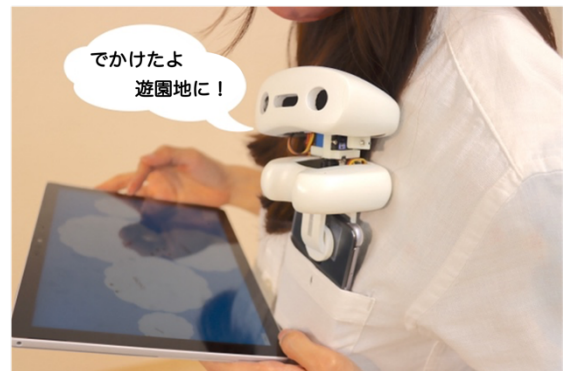


図1 〈PoKeBo Talk〉

コンセプトとプロトタイプのインタラクションについて述べる。

2. 研究背景

2.1. 発話における漸次性、漸次的精緻化

私たちの日常の発話には「漸次性(incrementality)」と呼ばれる性質があり^[1]、これに対して2つの意味が含まれているといわれている。

1つは「相手の知らない情報だけを伝える」という意味である。お互いが共有する知識や情報については繰り返さず、そのとき相手がまだ知らないことだけを伝えるのである。もう1つの意味は、発話を構成していく際に思いつくまま、その順序で新たな情報を追加していくということである。発話の断片を思いつくままに伝え、受け手側はその断片から発

話を自ら再構築しながら、そのメッセージを解釈するのである。

自らが表現したいイメージと発話となって表現されたものとの間には「ずれ」をとまなう。段階的に時間の許される範囲で、これらの「ずれ」を小さくし精緻なものにする。その結果、与えられた時間によって「精緻化の程度」は変化する。このことを **anytime** 性といい、こうした振る舞いを発話における「漸次的精緻化(incremental elaboration)」の振る舞いと呼ぶ。

精緻化は完全ではなく、「ずれ」は最後まで残る。発話の中で私たちは「精緻化の程度」を使い分けている。これらのことから日常の発話は必ずしも最適なものでなくとも、満足のするものであれば十分であると考えられている。

2.2. ブリコラージュ

ブリコラージュとは、クロード・レヴィ=ストロースが『野生の思考』の中で紹介した言葉であり^{[2][3]}、日曜大工・素人大工という意味を持つフランス語であるが、ここでは「ありあわせでその場を凌ぐこと」という意味で用いる。

ブリコラージュによって作られたものは、作り手の意図を実現したものであると同時にそれを超えた何かとしてある。つまり、ブリコラージュによって出来たものは作り手の思いとは不可避免的にずれるのである。なぜなら、用いられる様々な材料の用途というのはひとつだけではなく、いくつもの潜在的な用途と組み合わせの可能性が開かれ、それぞれの状況で意味や価値が変化するからである。

ブリコラージュによって生み出された発話も何かしらの意図を感じられ、相手に対して気持ちを伝えることが可能なのではないかと考えられる。

2.3. ブリコロージュ(bricologie)

コミュニケーションの形態を大きく分類すると 2 つに分けられる^[4]。

1 つは、話し手の独り言や、話し手と聞き手の関係が固定化して一方の話し手が長い時間聞き手に向かって話し続ける「モノローグ(monologue)」であり、もう 1 つは話し手と聞き手が交互にお互いの発話をやり取りする「ダイアローグ(dialogue)」である。

著者らは話し手が自らの伝えたい思いを断片的に表現し、その断片的な言葉を寄せ集め、聞き手に伝える新たなコミュニケーション形態として「ブリコロージュ(bricologie)」を定義した^[5]。本研究では、「ブリコロージュ」を備えたシステムの実現を狙っている。

3. 〈PoKeBo Talk〉の構成

3.1. 〈PoKeBo Talk〉

本システムは、断片的な言葉を寄せ集めて自らの伝えたい情報を持つアプリケーション〈MoCoMo〉と自身の代わりに聞き手に向かって発話を伝える〈PoKeBo〉から構成される。この 2 つを合わせた全体システムを〈PoKeBo Talk〉と呼ぶ。

3.2. 〈MoCoMo〉

ユーザとの相互的なインタラクションを経ることで興味や関心を生じさせ思いがけない出会いをもたらす情報メディアとして〈MoCoMo〉が提案されてきた^[6]。〈MoCoMo〉の外観を図 2 に示す。

〈MoCoMo〉のコンセプトは、「おしあいへしあいして、自らを主張する情報クリーチャ」である。本システムではブリコロージュとして自身の伝えたい言葉を表現するためのインタラクティブなクリーチャとして活用する。



図 2 〈MoCoMo〉の外観

3.3. 〈PoKeBo〉

〈PoKeBo〉はユーザーへ情報を提供するソーシャルインターフェースとして開発されたロボットである^[7]。図 3 に〈PoKeBo〉の外観を示す。

本システムでは〈PoKeBo〉が選択された言葉を自身の代わりにお話ししてくれるエージェントとして活用する。



図3 〈PoKeBo〉の外観

3.4. システム構成

〈PoKeBo Talk〉のシステム構成を図4に示す。

〈MoCoMo〉側は、タッチ操作及びキーボード入力によるテキスト情報の取得とボタンによる操作を実装している。〈PoKeBo〉側は、テキスト情報を受け取りスピーカーによって文章を読み上げ、モータによって動きの振る舞いを実装している。

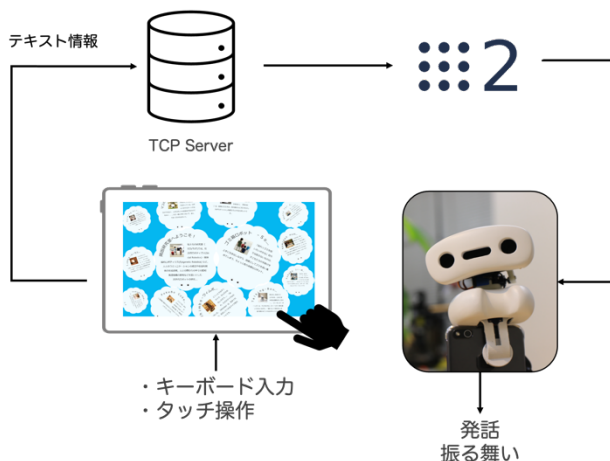


図4 〈PoKeBo Talk〉システム構成

3.5. ソフトウェアデザイン

〈PoKeBo Talk〉のソフトウェア構成について説明する。〈MoCoMo〉側はUnityで、〈PoKeBo〉側はRobot Operating System (ROS2)を用いて構成されている。図5にUnity-ROS2のシステム構成を示す。

Unity-ROS の連携ではUnity内の「ROS TCP Connector」とROSの「ROS TCP Endpoint」を介して行っている。「ROS TCP Connector」はROSの「ROS TCP Endpoint」間の通信を行うパッケージであり、「ROS TCP Endpoint」は他のROSのノードと直接データ通信を行うためのパッケージとなっている。

Unity側は自らの伝えたい言葉や文で入力することができる。入力された言葉や文は〈MoCoMo〉が情報として持っており、この情報をROS側にPublish

して情報を渡している。

ROS側はテキスト受信ノードで〈MoCoMo〉からの情報を受信した後テキスト処理を行い、ロボットの制御をする。発話ノードでは受信した情報をWizardVoiceと呼ばれる合成音声ソフトに送信し発話を行うノードである。動作ノードは発話に合わせて〈PoKeBo〉の振る舞いの制御を行うノードである。

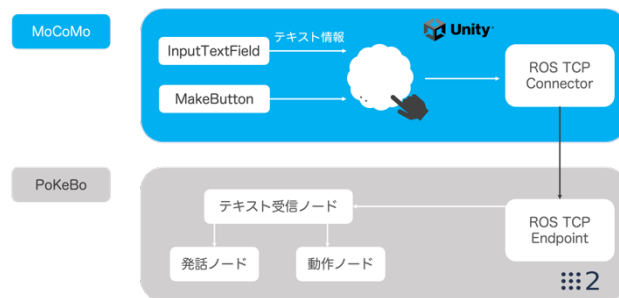


図5 Unity-ROS システム構成

3.6. ハードウェアデザイン

本システムの一部である〈PoKeBo〉のハードウェア構成を図6に示す。

〈PoKeBo〉にはDCモータが2個とそれを制御するための制御回路が組み込まれている。

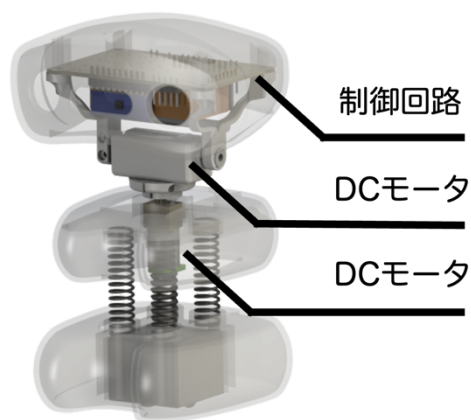


図6 〈PoKeBo〉ハードウェア構成

4. インタラクションデザイン

〈PoKeBo Talk〉のインタラクションは、〈MoCoMo〉のアプリケーションの操作と、これによる〈PoKeBo〉の振る舞い、発話を中心となり行われる。インタラクションの過程は以下のようになっている。

- ① 〈PoKeBo〉は待機している。
- ② 〈MoCoMo〉のアプリケーションを操作し、自らの伝えたい言葉をキーボード入力、または発話片を寄せ集める。
- ③ ボタンを押して〈MoCoMo〉を生成させる。

- ④ ②～③を繰り返す。話す内容が決まったら次のインタラクシオンに移る。
- ⑤ 〈MoCoMo〉を画面内に寄せ集め、伝えたい発話片を担う〈MoCoMo〉を2回タップしロボットに情報を送る。
- ⑥ 〈PoKeBo〉がおしゃべりをする。

インタラクシオンとして2種類の形態を搭載している。選択された発話をそのまま話すブリコログ形態と選択した発話片を英語で伝える翻訳形態が備わっている。

5. まとめと今後の展望

本稿では「ブリコログ」を備えたブリコラージュ風発話生成システムである〈PoKeBo Talk〉の基本コンセプトとプロトタイプについて紹介した。「ブリコログ」による発話はユーザとのコミュニケーションにおいてより親密な関係が築けるのではないかと期待している。

今後の展望として発話の入力を音声認識で取得し、よりユーザとのインタラクシオンが行いやすい設計にしようと考えている。また、音声認識を活用することで発話片の内容を GPT-3 や ChatGPT に生成させることで幅の広いコミュニケーションが可能になるのではないかと考えている。ユーザが選択した発話をフィードバックすることでより主体的なコミュニケーションが可能になると考える。そのためにもシステムのブラッシュアップを行っていく必要があると考えている。

謝辞

本研究の一部は、愛知県が公益財団法人科学技術交流財団に委託し実施している「知の拠点あいち重点研究プロジェクト第IV期（第4次産業革命をもたらすデジタル・トランスメーション（DX）の加速）」により行われた。

参考文献

- [1] 岡田美智男：『口ごもるコンピュータ』，共立出版；(1995).
- [2] クロード・レヴィ＝ストロース(著)：大橋(訳)：『野生の思考』，みすず書店；(1976).
- [3] 小山：現代社会の日常生活における野生の思考：未開人たちのもたらしたものと A.ブルトンに関する考察；東京商船大学研究報告.人文科学, 51 巻, pp.75-88 (2000).
- [4] 田島：教育実践を理解するためのバフチン・ダイアログ論 豊かな異文化交流の実現；言語文化教育研究会, 第 16 巻, pp.260-278, (2018)

- [5] 佐々木, 長谷川, 大島, 岡田：〈PoKeBo Talk〉思いつきをよせ集めるブリコラージュ風発話生成システムの提案；ヒューマンインタフェース学会 2022 論文集 2T-D2(2022).
- [6] 新保, 石川, 香川, 岡田：〈MoCoMo〉：宛名性を伴う情報環境が ユーザに与える効果について；Human-Agent Interaction シンポジウム 2016 論文集, G-17 (2016).
- [7] 真弓, 上野, 大島, 岡田：PoKeBo3 共構築型コミュニケーション空間の提案, ヒューマンインタフェースサイバーコロキウム, 67-70 (2020)