

英語のリズムを教えるコミュニケーションロボットを目指して

Towards a Communication Robot that Teaches the Rhythm of English

船越 孝太郎*¹ 中野 幹生*¹ 永田 亮*²
 Kotaro Funakoshi Mikio Nakano Ryo Nagata

*¹ (株) ホンダ・リサーチ・インスティテュート・ジャパン
 Honda Research Institute Japan Co., Ltd. (HRI-JP)

*² 甲南大学
 Konan University

As our economy gets global, second language education (SLE) attracts increasing attention but only a few spoken dialogue systems feature it, while SLE has several advantages in spoken dialogue systems research. In this paper, we introduce our SLE robot which features a rhythmic English training method called Jazz Chants.

1. はじめに

本稿では、我々が研究を進めている第二言語教育ロボットを紹介する。

第二言語教育 (SLE) に対する関心は従来にも増して高まっている。その中で、中国語やスペイン語なども存在感を増しては来ているものの、実質上の世界標準語である英語が、依然として第一の学習対象言語として選ばれる状況には変わりがない。日本においても 2011 年から小学校高学年での英語教育が必修となる。小学校教員の英語能力が限定的である現状では、家庭での学習補助だけでなく、学校における教員に対する補助・支援技術も強く期待されている。このような状況で、複数人と同時並行的に会話のできる音声対話ロボットは、強力な英語教育支援となる可能性を秘めている。

ロボットを SLE へ応用する研究はいくつかあるが [Kanda 04, Han 09, Movellan 09, Han 10, Lee 10], 発音訓練への音声認識の利用に比べると、SLE に着目した対話システム研究は多くない。

この背景には、対話を SLE に活用しようとする従来の研究が主に成人を対象にし、対話の中身や文法的側面に着目しがちであったために、音声対話で、かつロボットで複数人で行うことを想定すると、現状の技術水準では様々な面でハードルが高くなりすぎてしまうという事情があると考えられる。しかし、だからといって単純な語彙学習や発音訓練だけに限ってしまうと、わざわざ研究プラットフォームに対話システムやロボットを用いる意義や魅力が薄れるため、基礎研究レベルで取り組むことがなくなる。その結果、ソフトウェアとして開発した英語教材を僅かに修正して既製のロボットに搭載するだけとなっているのが現状ではないかと推察する。

一方、音声対話ロボット研究、特に非課題指向の研究で着目されることの多い前言語的コミュニケーションやヒューマンロボットインタラクションの研究素材として見たとき、第二言語教育・学習支援は非常に魅力的である。(1) ある種の身体的スキル獲得を目的とする第二言語学習においては、長期的な繰り返しが非常に重要であるため、対話システム技術の向上に必要なデータの収集・評価に有利なドメインであると考えられる。対話システムを使ったデータ収集・評価を行う際によく問題になるのが、対話システムの前にユーザを座らせて対話させようとしても、ユーザが何を話したらよいか分からず固まってしまう

うという事態である。ユーザ自身にはシステムを利用する明確な目的がないため、予め課題等を与えないと対話を続けることが難しい。SLE では、何かの要求・課題を達成することではなく、インタラクションを繰り返すこと自体が目的となるため、このような問題を回避し易い。もちろん長時間・長期間続けさせるためには、ゲーム的な要素を加えるなどの、飽きさせないための様々な工夫も必要であるが、そのような工夫はさほど難しくないとされる。(2) エンゲージメント形成、多人数対話管理、視線管理、ジェスチャー生成、動機付けや共感形成のための感情的コミュニケーションなど、多様な研究を行うことができる。特に、小学校程度までの幼児・児童を対象にしたレベルの学習内容に限定し、ロボットがレッスンを主導する形にすれば、語彙や表現のバリエーションを限定して、高度な言語理解・生成も回避することができるため、システムの実装や性能の確保も容易になるし、前言語的コミュニケーション研究に集中し易くなる。また、幼児・児童を対象とすれば、システムの理解・対話能力に問題があっても、ロボットやインタラクションのデザイン次第 (見かけやキャラクター設定など) で相当な部分をカバーできるとされる。(3) 実用性・商業的価値が高く、研究予算の獲得、産学連携が行い易いと思われる。

このような背景において、我々は Jazz Chants という教育手法に着目し、研究を進めている。Jazz Chants は、四行詩のような比較的簡単な教材を用いて、リズムに併せて体を動かしながら学習を行うので、主に児童を対象に利用されており、上記のように語彙や表現を限定し易い。その中で対話的なやりとりも好んで行うので、音声対話システムを用いる意義が十分にある。体を動かしてリズムを取りながら複数人で同時的に行うので、身体を持つロボットで行う意義も十分にある。

2. Jazz Chants

Jazz Chants*¹[Graham 06] は Carolyn Graham によって提唱され、世界的に普及してきている英語教育手法で、英語のリズムを獲得させることに重点を置いている。言語をそのリズム (心理的な等時性) で分類すると、英語は強勢拍 (stress-timed)、スペイン語やフランス語は音節拍 (syllable-timed) に分類される [窪園 91]。日本語は音節拍に近いが、より細かく分けてモーラ拍 (mora-timed) とされる。いわゆる日本人英語が英語らしく聞こえないのは、個々の単語の発音だけでなく、このリズムの違い (より正確には、英語らしいリズムを生

連絡先: 船越 孝太郎, HRI-JP, 埼玉県和光市本町 8-1, funakoshi@jp.honda-ri.com

*¹ Jazz Chants は Oxford University Press の登録商標である。



図 1: Jazz Chants を行うロボット: Chanty Bear

み出す強弱のパターンの欠落)が矯正されていないことによる。母語のリズムが第二言語の発声にも影響することは最近の研究でも示されており [Yuan 10], 第二言語学習者にとって重要な学習要素である。

Jazz Chants では、4 拍子の Traditional Jazz を BGM に、以下のような短いパッセージ (以降、チャンツ詞とよぶ) を繰り返し発声することで学習を行う。

```
* * * *
Two, four, six, eight.

* * * *
I don't wanna be late. I don't wanna be late.

* * * *
Two, four, six, eight, ten.

* * * *
Say it again. Say it again.
```

*マークを付けた場所が強拍の置かれる音節で、これが Jazz の 4 拍リズムと時間的にそろえるようにする。その他の音節は弱音化し、拍間に収まるようにする。CD 教材を用いて個人学習もできるが、一般的には指導者とともに、手拍子をしたりステップを踏みながら多人数で行う。

3. 実装例: Chanty Bear

上記のような英語教育手法を実践する音声対話ロボットの実現可能性を確かめるため、既存技術を用いたプロトタイプシステム (図 1) を構築した。音声認識には Sphinx-4^{*2}を用いた。通常の Text-To-Speech では Jazz Chants の等拍性を表現できないので、音声合成には歌唱合成ソフトウェア VOCALOID2 (Sweet Ann)^{*3}を用いた。チャンツ音声を合成するために必要な時間情報及びピッチ・パワー曲線は、Jazz Chants 指導者の実演を BGM 無しで独自に収録したデータから取得した。ロボットにはイワヤ製の IP RobotPhone を用いた。ロボットは事前に決められたタイミングでチャンツ詞に併せて頭部を動かすことでリズムの存在を教える。また、ユーザがチャンツを行うときには、ロボットは一定間隔で腕を振って拍のタイミングを提示し、ユーザを補助する。また、ユーザの実演を音声認識した結果から、等拍性を基準としてスコアを提示する。より高度で多様な対話を実現するためには多くの研究課題があるものの、基本的な形のインタラクションであれば、既存の技術でも実用化できる見込みがあることが分かった。

このロボットを身近にいた 3 歳半の日本人女児 1 名 (英語学習経験はないが、ディズニーアニメを通じて英会話を耳にしたことはある) に見せた所、非常に強い関心を示し、ロボット

が動作を開始するとその動きに併せてすぐに一緒になって全身で上下運動を始め、つたないながらもチャンツ詞をまねし始める様子が観察された。時折目にするテレビの英語教育番組ではただ見入るだけでまねをするようなことはなかったことと対照的であった。このことは開発したロボットが幼児・児童の英語教育に多大な効果を発揮できる可能性を示唆している。しかしながら、前述のように固定のチャンツ詞をただ繰り返すだけのロボットなので、女児は直ぐに飽きてしまった。児童が飽きないで学習を続けられるようにするためには、ロボットとの対話を (前言語的コミュニケーションのレベルで) もっと豊かにする必要があると考えられる。

4. 幼児・児童教育における倫理的問題

幼児・児童の教育にロボットを用いようとする際、そこには倫理的問題が潜む可能性に十分に注意を払う必要がある [Sharkey 08][田中 09]。そこには、完璧ではない合成音声にさらされることで児童の発音に悪影響を与えるのではないかと比較的程度の懸念から、親がロボットに子供の教育を任せっぱなしにして育児放棄に近い状態を多発させてしまうのではないかと深刻な懸念まで様々なものがある。

ゲーム機、パソコン、携帯電話、テレビなど人間や社会に対して影響力を持つものは必ず何かしらの批判を受けるものなので、どこからも何も言われなくすることはできないだろうから、一定の批判が来ることは覚悟して取り組む必要がある。それでも、子供達に深刻な害が及ばないように最大限努めること言うまでもなく、研究の妨げとなる無用の障害を作らないためにも、社会のコンセンサスを得易いロボットの利用方法・環境などを考えておくことは重要である。

第一には、ロボットと子供だけがずっとインタラクションをすることがないように、積極的に大人が関与するようにすることである。学校教育であれば、あくまで指導の主役は教員であり、ロボットはその補佐をするという形態を目指すのが理想的であろう。しかし、特に家庭への導入を考えると、常に大人が付き添うことを前提とするのは無理がある。これを補う方法として、田中 [田中 09] が指摘するように、ロボットを教師役・養護役として子供の上に位置づけるのではなく、寧ろ子供が教える・面倒を見る対象として位置づける枠組みとすることが有望であると考えられる。例えば、ロボットの使用が英会話スクールでの人間の講師による Jazz Chants レッスンとセットになっていて、帰宅後に子供がロボットを相手にレッスンを再現する (子供が講師役) という形態が考えられる。これには副次的な効果として、学習効果が高いとされる Learning-by-Teaching [Gartner 71] の形態をとれるという利点もある。

5. 今後の展望

現在のシステムでは、事前に与えたチャンツ詞を繰り返すことしか出来ないが、我々は詞中の強勢拍位置の自動推定にも取り組んでいる [Nagata 10]。これと Wang らがやっているような小節への自動割り当て技術 [Wang 10] を組み合わせ、抑揚と強弱の付与を自動化することで、学習者のリクエストに応じてロボットがその場で自動生成したチャンツ詞 (例えば花の名前を集めたチャンツ詞など) で Jazz Chants を行えるようになり、語彙学習やフレーズ学習への幅も広がる。対話的に即興詞を掛け合うチャンツも可能になる。これらの要素技術を開発しつつ、エンゲージメントや多人数対話管理などの研究を今後進めていきたい。

*2 <http://cmusphinx.sourceforge.net/sphinx4/>

*3 <http://www.vocaloid.com/product.html>

また、現在の教示方法はチャンツ詞を丸ごと繰り返すだけなので、学習者は事前にある程度チャンツ詞を覚えておかなければついていけない。上記のようにロボットが作成したチャンツを使うなら、ロボットが行やフレーズの単位に細かく分けてその場で教えて行く必要がある。試みに、人が人にチャンツ詞を1から教える対話を収録し、その過程を対話システムに実装してみたが、Chanty Bearよりも単調で退屈な印象の対話になってしまった。どのようにすれば学習者が楽しんで対話を続けられるようにできるのかも重要な研究課題である。

頑健なインタラクションを行うためには、一般的な音声対話システムのように音声認識駆動を基本とするのではなく、音や画像などの信号レベルでの処理を基本とした対話管理の枠組みも必要である。

参考文献

- [Gartner 71] Gartner, A., Kohler, M. C., and Riessman, F.: *Children teach children: Learning by teaching*, Harper & Row, New York (1971)
- [Graham 06] Graham, C.: *Creating Chants And Songs*, Oxford University Press (2006)
- [Han 09] Han, J. and Kim, D.: r-Learning services for elementary school students with a teaching assistant robot, in *HRI '09* (2009)
- [Han 10] Han, J., Lee, S., Kang, B., Park, S., Kim, J., Kim, M., and Kim, M.: A trial English class with a teaching assistant robot in elementary school, in *HRI'10* (2010)
- [Kanda 04] Kanda, T., Hirano, T., Eaton, D., and Ishiguro, H.: Interactive robots as social partners and peer tutors for children: A field trial, *Human-Computer Interaction*, Vol. 19, No. 1, pp. 61–84 (2004)
- [窪園 91] 窪園 晴夫, 溝越 彰: 英語の発音と英詩の韻律, 英潮社 (1991)
- [Lee 10] Lee, S., Noh, H., Lee, J., Lee, K., and Lee, G. G.: Cognitive Effects of Robot-Assisted Language Learning on Oral Skills, in *INTERSPEECH 2010 Satellite Workshop on "Second Language Studies: Acquisition, Learning, Education and Technology"* (2010)
- [Movellan 09] Movellan, J. R., Eckhardt, M., Virnes, M., and Rodriguez, A.: Sociable robot improves toddler vocabulary skills, in *HRI'09* (2009)
- [Nagata 10] Nagata, R., Mizumoto, T., Funakoshi, K., and Nakano, M.: Toward a chanting robot for interactively teaching English to children, in *INTERSPEECH 2010 Satellite Workshop on "Second Language Studies: Acquisition, Learning, Education and Technology"* (2010)
- [Sharkey 08] Sharkey, N. E.: The Ethical Frontiers of Robotics, *Science*, Vol. 322, (2008)
- [田中 09] 田中 文英: ロボット技術の早期教育応用にまつわる諸問題, 第23回人工知能学会全国大会 (2009)
- [Wang 10] Wang, H., Mok, P., and Meng, H.: MusicSpeak: Capitalizing on Musical Rhythm for Prosodic Training in Computer-Aided Language Learning, in *INTERSPEECH 2010 Satellite Workshop on "Second Language Studies: Acquisition, Learning, Education and Technology"* (2010)
- [Yuan 10] Yuan, J.: Linguistic Rhythm in Foreign Accent, in *INTERSPEECH'10* (2010)