

特集 「ゲーム産業における人工知能」

# シーマンは来たるべき会話型エージェントの福音となるか? : 斎藤由多加インタビュー

Will Seaman Lead AI Technology for Interactive Agent?:  
Yutaka Saito Interview

斎藤 由多加  
Yutaka Saito

オーブンブック株式会社  
OPeNBook  
yutaring@gmail.com, <http://openbook.jp/>

聞き手：大澤 博隆（筑波大学，人工知能学会誌編集委員）

三宅 陽一郎（株式会社スクウェア・エニックス，人工知能学会誌編集委員）

構成：高橋 ミレイ（編集者）

**Keywords:** Seaman, conversation, natural language processing.

## 1. はじめに

人間と心を通わせて親しい会話ができる人工知能はいかにして開発できるのだろうか。そのヒントとなるゲームが今から 18 年前にリリースされていた。ドリームキャスト向けのゲーム『シーマン～禁断のペット～』（セガ，1999）は、架空の生物シーマンを育成しながら、シーマイクコントローラに付属するマイクを通してシーマンと会話できるのが特徴だ。時にふてぶ

てしく、時に心を砕く友人のようにユーザの気持ちに寄り添うシーマンの言葉は、「シーマン語録 現代人への 185 の賢言」（斎藤由多加 著，ダイヤモンド社，2001）というタイトルで単行本にもまとめられている。いわゆるスモールトークは現代の会話型エージェントの不得意とする部分であるにもかかわらず、当時の技術でなぜ可能だったのか？ 開発元のオーブンブック株式会社の代表であり、シーマンシリーズのプロデューサーの斎藤由多加氏にシーマン開発の裏舞台についてうかがった。

話エージェントが出てきて自然言語処理の技術は向上しています。それにもかかわらず、マトモに人と会話をしようとする、どうしても不自然になってしまいます。斎藤さんのつくられた『シーマン』がユーザとの自然な会話をなぜ実現できたのか、今、人工知能分野の研究者の間でも注目されておりますので、『シーマン』の会話エージェントをどのように構築されたのかを、ぜひお聞きしたいと思いました。

斎藤：シリーズの中で一番最初に出した『シーマン～禁断のペット～』（以下『シーマン』）は、分岐型のシナリオの会話で構成されています。あた

## 2. 引数管理で自然な会話を実現する

三宅：現在は Siri など、いろいろな会



図1 向かって左から大澤氏，斎藤氏，三宅氏。  
2017 年 1 月，斎藤氏の六本木の事務所にて



図2 大ヒットとなった会話ゲーム『シーマン』  
© OPeNBook Inc.

かも本当に理解しているかのようなセリフ回しはすべて人間の筆の力によるものです。またいくつかのトリックによって、あたかもシーマンが考えているかのようにも見せています。

例えば「お前、何歳？」と聞いたとき、1歳から99歳までの答えを想定しているとしましょう。ただし、なかには「天才」と言う人もいるだろうということも想定します。あるいは「星座は何座？」と聞いたときに、牡羊座などの12星座以外に「銀座」とか「餃子」と答える人もいるかもしれません。そういうことを創作者が先読みして入れ込んであるのです。「お前、学校でもそういうこと言ってるだろ？」とか、あたかも理解して即興で返しているようにコメントを返すのです。それ以降も、とにかく相手の言っていたことをちゃんと記憶しているぞ、という言い回しを「引数」を駆使して端々に出す。そういう引数を充実させるのが一つのテクニクでした。

AppleのSiriだと一問一答に近いような形の会話ですが、シーマンではユーザが前に言った言葉を覚えて次の会話にも反映させる機能を、初期バージョンはシナリオの分岐と「引数」の組合せによって、そして現在試作しているAIではプログラムで実装しています。

最初の『シーマン』では、ですから膨大な分岐という形でアナログ的にセリフを紡いでいったというわけです。次のバージョン『クリスマスシーマン～想いを伝えるもう一つの方法～』（セガ、1999）では、膨大な引数を入れることによってバリエーションを増やすことで原稿の分岐をできるだけ減らしました。というのも、分岐から分岐という形を続けると、あまりにも分岐が広がり過ぎて人間の頭ではリターンする先がそもそもどこだったのか管理がおぼつかなくなってしまうからです。

ですから、中期になってくると、大きなセリフの流れだけつくっておいて引数が入る部分だけ用意しておくのです。

そしてユーザの胸に響くような名セリフを言わせたいときは、あらかじめつくっておいた人間が感動しそうな語録100パターンぐらいのライブラリーの中から、そのときそのときに当てはまりそうなものを検索してきて発話させるのです。そうすることで、引数と名言が組み合わさって、あたかもユーザのことを理解しているようなセリフにするわけです。例えば職業が弁護士だというユーザに対して「×××（職業）のように人の上に立って先生と呼ばれる職業は、人に教わる立場でもあるんだゾ」という言葉を引っ張ってくると、ぴったりとはまってくるわけです。

いちいち「あなたは弁護士ですから……」と職業名を復唱することなく、しかしユーザに理解していることを匂わす言い回しとして言葉の端々に宿らせることができるわけです。

大澤：その辺はアイディアも含めて斎藤さんがすべて管理されていたのでしょうか？

斎藤：はい、そうです。

『シーマン』の制作において、非常に優れたエンジニア達が僕をサポートしてくれていたことは事実です。ですが、シーマンというものが一体何なのか内輪のスタッフですら開発前半の時期は理解できなかったもので、しゃべる内容や、演出、声優まで全部自分でやらざるを得なかった。自分で原稿を書いて、自分で読んで、分岐先のセリフも考えて、と、要するに、AIは企画がすべてです。一貫性を考えると、誰かと分担したり外部のライターに依頼することができませんでした。

言葉には、主語述語といった文法以外にも、その奥にある大きなコンテキストの中にその人の視点みたいなものが反映されているものです。それがずれていると受け手には伝わってしまうんですね。例えば『クリスマスシーマン～想いを伝えるもう一つの方法～』を制作したときに、有名な脚本家の方に手伝ってもらったことになったのですが、原稿を見て

大変失礼ながら全部ボツにしちゃったんです。その方は人気のあるテレビ番組も担当する非常に著名な脚本家でしたが、シーマンのキャラクターと「視点」が違ったのです。

一人で書くからできることに、「首尾一貫とした視点、ものの見方」などがあると思うんですよね。「辛口なことも言うけれど絶対に相手を貶めない視点」を僕はこのゲームでは一貫して守るようにしていたのですが、視点とは、つまりそういうことなのです。そういうものを、一人の人格が書くことで貫いてきたことを、AIがどう引き受けてコントロールしていくのかが、今後の大きな課題になっていくと思います。例えば同じことを言うにしても、丁寧語だったり、標準語だったり、あるいはがらっぱち語だったり、異なるモードがありますが、それはキャラクターの設定ではなくてAIの機嫌によって切り替わるべきだと思います。例えば「おっしゃってください」から「言っていますよ」→「言えつってんだよ」と、この三つのモードを、違う言語系ではなくて同じ言語系の別モードとして捉えていくということです。そういうスイッチングで切り替えていく言語のレイヤのようなものを人間はもっているのですが、人工知能の分野では、あまりそういうところまで議論が行われていないように思います。特に日本語は端々に上下の立場を確認する言葉が入っていますから、そういう“出口縛り”みたいな視点をAIに従事している人間は、もっともつべきだと思います。

大澤：今のお話を聞いて、シーマンが異星人的で、我々の日常とは離れた視点で自分達を見てくれる存在として設計されているところがすごいなと思いました。それは当初からあったコンセプトなんですか？

斎藤：いくつかは意図的にやってきましたが、残りのほとんどは私の直感、つまり偶然の産物だったと思っています。しかし今、当時のドキュメントを振り返ってみると、自分なりに言葉をデザインする一つの枠組み



図3 インタビューに応じる斎藤由多加氏

みたいなものをもたせていた形跡が  
あちこちに残っているのです。不思議  
なことです。たぶん当時の自分より  
も、今の自分のほうがこの AI に関  
しても物事が見えているはずにもか  
かわらず、です。当時どうしてそこ  
までのことに気付いていたのか？  
それがいまだにわからないんですけ  
れど。

### 3. メロディ言語と間、そこから新たな 文法の枠組みへ

斎藤：『シーマン』で声優をしていると  
き（注：シーマンの声はすべて斎藤  
氏自身が演じている）、言葉のメロ  
ディがもつ規則性みたいなものがある  
ことに気付きました。それをベー  
スに今、日本語の「メロディ言語」  
というのを定義しようとしています。  
そういう発想が生まれたのも自分で  
何もかもやってきたことの産物だと思  
います。合計で 200 万本ほど売った  
シリーズのユーザからの反応をすべ  
て受けて、原稿を全部自分で書き、  
声を全部自分で録音したゆえに得  
られた経験はとても膨大でかつ貴重  
です。今、実は「シーマン人工知  
能研究所」というのを開設しよう  
と思っているのですが、そこでやろ  
うとしている研究はこの「メロディ  
言語」の文法が大きな根幹になる予  
定です。ロボットの頭脳と会話につ  
いては、もうとっくに優れた人工発  
話のエンジンがつけられていて、自  
分にできることなどもうないと思っ

いたのですが、この間見学に行っ  
て、僕らが 14 ~ 15 年前にやっている  
ことと同じことをいまだにやってい  
るのを見て、ちょっとびっくりしまし  
た（笑）。そんなことなら、自分のも  
っているノウハウを人工知能やロボ  
ットの開発に貢献できる余地がまだ  
あるかもしれない。まだまだ出番は  
ありそうだと、とそう思って研究所  
を設立しようと走り回っています。

三宅：我々が今回来たのも同じ理由で  
す。言語研究や会話研究が進んでい  
て、ロボットにもいろいろなソフト  
ウェアが入っているけれど、実際に  
会話するとなると、単発の会話しか  
できず、自然な会話の流れというの  
もありませんから。

斎藤：まず皆さんにハッキリ言えるこ  
とが一つあるとすれば、文語をベー  
スにしておられることを抜本的に見  
直すべきだと思います。我々が学ん  
できた国語の文法は、この業界の我々  
が今やろうとしている言葉のやり取  
りに対してはまるで“無力”だとい  
うことに気がつくべきです。では、  
何を拠り所にすればよいかといえば、  
それがいいんです。ないから口語の  
ための文法をつくり直すしかありま  
せん。さっきお話ししたメロディに  
規則性があるメロディ言語も、膨大  
な電話帳 20 冊分（注：過去のシー  
マンのセリフ原稿のこと）ぐらいの  
原稿を読み上げた人間だから気付く  
ことがベースとなっています。

2001 ~ 03 年頃にダウンゴの川上

量生氏と、携帯電話が鳴ると、シー  
マンが実名で所有者を呼んでくれる  
「呼び捨てサービス」という着メロ  
サービスをやっていたのですが、こ  
れは、「おい、三宅、電話だぞ」と着  
信音が言うわけです。

このため、例えば「三宅」と言わ  
せるために何をしたかという、非  
常にシンプルなことを彼は提案して  
きたのです。テキストスピーチを使  
うわけでも何でもなく、「あ」は青  
山から「わ」は渡辺まで日本人の姓  
2000 を全部読み上げろ、と。その中  
から自分の名をユーザはダウンロード  
すればいい、と。つくる側としては  
地獄の特訓みたいな話ですが（笑）。  
そのときに「三宅」という言葉を 7  
パターン用意すると、ほぼ「みやけ」  
に必要なすべてをカバーできること  
に気付いたのです。

「おい、三宅、電話だぞ」、「三宅  
から電話だぞ」と、三宅という一つ  
の単語でもその種類のメロディ音声  
があればあとは組合せでつくってし  
まえることがわかってきたんですね。  
例えばシーマンが「お前、昨日誰と  
デートしたの？」と女の子に聞いた  
ときに、「三宅くん」と答えたときに、  
「三宅!?（すごく嫌そうに）」と答え  
たら、「やめとけ、あんな奴」って意  
味なんですよ。「三宅…?（戸惑った  
口調で）」だと「誰それ?」という意  
味だし、「三宅（棒読み）」だと「へ  
〜」という意味です。これについて、  
「みやけ」という 3 文字、バイト情報  
的には同じ 3 文字の言葉の中に宿る  
意味が、メロディによってだいぶ補  
われているというのが今までの見解  
でした。ところが、実はそうではな  
いんです。メロディが主軸にあって、  
そのメタ情報として「三宅」という  
言葉があるのです。言いたいニュア  
ンスがすべてメロディですでに表現  
されているのが口語なんですね。だ  
けど、このイントネーションの違い  
は文字で書かれた教科書では説明で  
きないですよ。

それから、無音の「間」というの  
は文字原稿にならないから録音して  
も捨てられてしまう。でも、一番大

事なのはこの「間」だったりするわけじゃないですか。例えば「昨日、彼氏に振られちゃった」と言ったときにAIがすかさず「マジ？」と答えるのと、しばらく間を置いてから「…マジ？」と答えるのとでは、全然意味が違う。こういった音楽的な要素が言葉の時間軸には多分にあるのです。しかしそういったものは文字にはならないから文語表現では全く削り取られてしまいます。

だから「間」を上手に駆使する講演者の話は、聞いていて面白いけれど、その内容を文字起こしすると、つまらなく、間抜けにすら見えることがある。なぜかという、上手な話者は人の心を惹くために「間」や「省略」を多用しているからです。これらは文字に変換すると劣勢に回ってしまうということですね。

つまり、口語は文語の代用品だという考えをいったん頭から捨てなければならぬということです。そうすると、私達が知っている従来の文法も捨てることになりますし、文法を捨てるということは、主語述語やSVOCといった構造も捨てなければならぬということです。まあ、こんな文法なんて試験問題以外何の役にも立たないですから。

文脈を宿らせるために僕は言葉を使っているのに、文語にはそれが当てはまらない。だったら宿らせるための文語体系を自分でつくってしまったほうが良い。そうすると、全く新しいやり方で言葉を定義する必要があります。例えば「食べる」という単語があります。「食べてんじゃねーよ」とか「食べたらず殺す」とか、実際に僕はそういう言い回しをします。僕らのつくっているAIの定義では、それを「食べる」の活用形として使っていますから、「食べてんじゃねーよ」は「食べる」の活用形で1ワードとして捉えられます。そのように定義し直しているのです。

大澤：これはすごく面白いですね。言葉に含まれる、抑揚やテンポから得られる情報を、パラ言語と我々は呼んでいます。しかし今のAIの技術だと、どうしても統計的な情報処理手法が主ですから、オンラインでデータのとりやすい文語形式の分析がまだ主流です。

斎藤：統計情報というのはリソースが文字でしょう？ 文字で情報を集めて、例えばそれをディープラーニングとか何かで学ばせたところで、生成されてくるのは文字じゃないですか。だけど今我々が相手にしなければな

らないのは「メロディ付きセリフ」です。特に日本人はフレーズを一塊のパッケージで覚えている、単語一つ一つを組み立てながらメロディとして組み立てているわけではない。「食べてんじゃねーって言っただろ？」というフレーズを例に取ると、多分頭の中に「○○じゃねーって言っただろ？」というテンプレートがあって、○○の中に「食べる」という動詞引数をぶち込んであるだけなんです。慣用句みたいな言い回しがあって、そこにいくつかのパラメータをぶち込んで活用している、それがフレーズです。「いつまで食べてるんだ、この野郎」というのも「いつまで○○…」というフレーズを活用した例ですね。それを一つの単語の活用形と捉えると、活用形が400個ぐらいあるのだけれど、受験問題のための文法としては多すぎる。けれど、人間ではなくコンピュータのためのものだと考えればどうってことないわけですよね。「食べる」の活用形をとにかく全部出してみても、後は「食べる」を「走る」や「笑う」といった他の動詞に置き換えて、変なところがあれば手で直す。そうすると、全く新しい辞書が出来上がってくるわけです。



図4 対談風景

三宅：口語の文法をつくるということですね。

斎藤：そうです。そして、それは可動構造でなければならない。例えば、「食べてんじゃねーよ」を過去形にしたり未来完了にしたり、あるいは否定形にしたり疑問形にできないければ、意味がないのです。数だけ羅列したものではなくて、意味的に変化できるものという意味での、可動構造をしていなければなりませんから。

大澤：例えば映画に出てくるヤクザの脅し文句なんかは、意味が全然わからなくても、怖いと感じたり自分が脅されていることが伝わるといことですね。

斎藤：そうですね。そうそう、海外のホテルでテレビを観ていたら、日本のヤクザ映画をやっていて、なぜか英語で字幕が出ていたんです。そうしたら、「お控えなすって」の英訳が「How do you do?」と出てました(笑)。

大澤：今だとオンラインで文語データを集めることができますが、オフラインの会話はまだ集めづらそうです。斎藤さんは、どうやってサンプルを集められたのでしょうか？

斎藤：これは冗談と思うかもしれませんが、かつては自分でしこしこ集めていまして、具体的にはキャバクラですね。『シーマン』をつくっているときにネタに困ったら、ほとんど毎晩キャバクラに行って飲みながら、ワットとメモを取って帰ってきました。例えば「それちょっとビミョー」というフレーズがあるじゃないですか。多分 2000 年ぐらいに出てきたフ

レーズだと思いますが、それを俳優の方にレコーディングで言ってもらおうとしたのですが、そのフレーズを知らない世代の人なので、発音できないんですよ。結局最後までうまく録れませんでした。言葉ってすごく変化していくものですが、そういうのをサンプリングするために、ゲーム制作という意図は隠してただのお客としてキャバクラによく行っていました。深夜でもやっていて、お金を出せば誰でも入れますからね(笑)(このときの経験値を元に「指名される技術」(ゴマブックス, 2015)という本を堀江貴文氏と共著で出版しました)。

#### 4. テキストにできない言葉をサンプリングする

大澤：『シーマン』はプレイしている様子を多人数で観ていてもすごく面白いですよね。誰かが言ったことがシーマンに突っ込まれるのを観たり、そういう状況で楽しむことも含めてゲームを設計されたんでしょうか？

斎藤：いえ、全く意図していませんでした。これは僕も発売のときにコンパニオンが店頭で『シーマン』のデモをやっているとき、ドッと笑いが起きるのを見て初めて発見したことでした。コンパニオンがシーマンにいじられているのを街頭で見たおじさんがゲラゲラ笑っているのを見て、「ああ、こういう楽しみ方もアリなんだな」と思いました。でも、それは全く想定していなかったことでした。

三宅：基本的にはスクリーンの前でユー

ザが自分で楽しむものとしてつくったんですね。

斎藤：そうです。ところで、映画の『Her』(スパイク・ジョーンズ監督, 2013)はご覧になりましたか？あの映画を観てどう思われましたか？

大澤：音声認識よりも別の点、ジェンダー的な視点なんか面白いと思いましたが……。

斎藤：みんなそうおっしゃるんですよね。僕にとっては、あの映画を観て「はあ……」というため息をつく AI をつくらないといけないと思ったのが大きかったですね。「間」もそうですし。実は『シーマン』では引数化を進めれば進ませるほどこの「間」が失われてしまったんです。ため息って、文字としての情報量が「無」なので音声ファイルとしては存在しないはずの部分なんです。「はあ……」と表記系にする音がないですから。だけど、実生活ではこれがもたらす効果は絶大です。ですから、皆さんには、もう 1 回、音と間、そして口語というものに目を向けてほしい。例えば「はあ……ダメだよ…」と言うのと、ただ「ダメだよ」と言うのとでは伝わる意味がなんとなく違います。どう違うのかはハッキリと言葉にはできませんが、やっぱり違いますよ。言葉にならないものを表現するのが、人間の感情のやり取りだと思います。これまで人工知能を開発してきた人達は、会話を知識のやり取りだと思い込んでつくっている方が多いように見えます。ですが、家の中で人と自然に共存できる人工会話があるとしたら、それは「今日、ジャイアンツ勝ちましたか？」とか「今一番売れている本はどれですか？」とか「今度行く AKB のチケットを取ってください」という事務的なリクエストに答える AI ではなくて、「振られちゃった」と言ったときに「マジ？」と言ってくれる AI が、求められているものの半数を超えると思います。情報を取得する手段や、自分の代わりにインターネットで検索するものではなく、ただ聞いて感情を共有してくれるだけでいい、い



図 5 『シーマン』ゲーム画面  
© OPeNBook Inc.

や、もっと言うと、「頷くだけ」でいい。「頷く」というのは、情報的には無ですよ。でも、そこを共有する仕組みに我々が目を向けていかない限り、冒頭でおっしゃった、人間と親しく会話をする人工知能をつくる手立ては見つからないと思います。エージェントというと検索エンジンの代理人みたいに思っておられるんじゃないかと思いますが、何も検索しない、ただ「マジか…」と言ってくれる(笑)。

そういうことに目を向けるためにシーマン人工知能研究所をつくろうと思いました。

三宅：シーマン人工知能研究所では、そういう文法を汎用的に構築するという計画になるのでしょうか？

斎藤：そうです。いくつかありますが、文法をつくるのは「いろは」の「い」の部分です。そこから後はピッチ認識、いわゆるメロディ認識を併用しないといけませんから。

三宅：Siri もそうですね、今はどちらかというといろいろな人の知識の寄せ集めで質問に答えるという方向に向かっていますね。そうであるにもかかわらず自然な会話ができない。そこはさっきご指摘されたように、視点が会話ごとにバラバラだからだと思います。

斎藤：後は省略の補完の構造。これが相手との共感共有の一つの言語の果たす役割になっています。例えば、「もう別れるわ」、「マジ？ なんで？」、「ウザいし」という女子高生の会話があるとすると、これを文語的にフルな文章にすると、「私、彼と別れます」、「あなたは彼の彼と本当に別れるのですか？ なぜ別れるのですか？」、「その彼がウザいからです」となります。省略のない文章はリズムやテンポが出ない、省略された言い回しだと、省略部を補完することで生じる共有感が、感情共有の大きなエレメントのように思います。つまり省略しても伝わる関係性が同族意識の確認になる。女子高生の言い回しは日本語の崩壊と呼ばれることもあります。僕はそうは思わない。高度に省略化された言語と言外にそ



図6 聴き入る大澤氏(右)と三宅氏(左)

れを補完するやり取りは、それが伝わるという同族意識を確認する儀式でもあるからです。だから、省略してそれを補完することそのものに実は会話の意味があるのではないかと思います。

「別れる」、「マジ？ なんで？」、「ウザい」、「どうすんの？」と、一つの文章を省略して強調したい部分を冒頭にもってくることを強調倒置といいます。それをフルの文章に戻す仕組みをつくることで差異の部分を検知できるなら、その部分を強調形でいうことができます。それを繰り返すことによって、会話を省略したり補完したり、省略したり補完したり…というやり取りが可能なのです。それは文脈を宿らせることとはほぼ同義語だと思っています。だから文脈の確認というのは、言い方を変えると「省略しても補ってくれること」ということだと思います。

大澤：今だと顔文字や LINE のスタンプのようにエモーショナルなものを記号で伝える手段が全世界に出てきていますよね。将来的にはシーマンが、あのようなコミュニケーション手段に応用されていく可能性はあるのでしょうか？

斎藤：むしろ、そちらに応用するほうが簡単なのかもしれません。顔文字は言ってみれば表情付けを目に見える形でもってくるわけですからね。むしろ、それが本当のベースになるかもしれません。ただ、我々は残念ながら口語を使っていますから、そ

のメロディの言語を確立するほうが会話認識においては先行すべき課題だと思います。

##### 5. 「あなたに向けたメッセージ」をいかにしてつくるか

大澤：どんな人にでも刺さる言い回しにもっていくためのノウハウや、制作の際に意識されていることはありますか？

斎藤：少しずるいのですが、人の心理を突いた設定をいくつも仕込みました。ユーザはある程度シーマンの言うことに感服し始めてしまうと、その後何を言っても感服するようになるのです。例えば、「Mac が好きで、タイガースのファンということは、お前 B 型だろ？」とシーマンが言うようにしたところ、当てられた人がびっくりしてネットに書き込みまくるわけです。日本人で B 型の人は十人に二人なので、2 割の確率で当たりますから。A 型だと母数が多すぎてこちらの意図がバレてしまうので避けました。AB 型と言わなかったのは、当たる確率が下がりすぎてしまうからです。ただ、この設定は海外版では却下されました。「血液型？ 意味わからないです」とか言われて、これは文化の違いです。

この設定のヒントになったのは、昔僕がリクルートにいたときの出来事です。当時、大きな汎用コンピュータをユーザに貸し出すセクションにいたときに同僚と銀座を歩いていたんです。そうしたら屋台が出ていて、

「コンピュータによる手相占い」という看板が出ている。手を乗せると手をスキャンして、横にあるドットマトリクスプリンタで占い結果が出てくるわけです。それで僕がやったとき、「あなたは、見た目は大胆そうに見えますが、実はすごく繊細。でも、それが周りから理解されておらず、大変苦しんで悩んでいますね」と出てきたんです。当たっているのびっくりして、どういうコンピュータを使っているんだろう？ と裏側を見たら富士通のオアシスと書いてあったんですね。よく見ると、おじさんが相手の顔を見て 2 とか 3 とか出力されるだけの仕組みなんです。つまりプリセットされた文章が横に出てくるだけで、「一見すると大胆そうですね」と僕を見て判断したのは、そのおじさんだったんです。

つまりほとんどの場合、占いが当たるか当たらないかを判断するのは、占われている本人ということですね。自分が考えていることを言い当てられれば、当たりだと思うものなんです。人間が思っていることは大体同じようなもので、「あなたはすごく優しい側面をもっているんだけど、周りから理解されていないですよ」と言われると「そうなんだよ…!」となるんですよ (笑)。よく考えると誰にでも当てはまることだとしても、その前後に「40 代になった研究職のあなたとしては…」と付いたりすると、急にそれが自分のことのように思えてくるものなんです。

大澤：ある種のバーナム効果ですね。誰にでも当てはまることを言う自分のことだと思う。

斎藤：そういうことなんですよ。先ほどのご質問に戻ると、万人に刺さる言葉とおっしゃったけれど、これは占い師がクライアントを信じ込ませるノウハウにヒントがあると考えています。……って、今度の本で書いていることを全部バラしていますね (笑)。

三宅：今書かれているご著書のタイトルや発売日は、もう決まっているのですか？

斎藤：まだ決まっていませんが…… (笑)。あと、なんといってもシーマン人工知能研究所の設立ですね。人工知能とは言っても、今申し上げたとおり、人間の言葉を理解したフリをする技術だけなので、人工知能というのはおこがましいかもしれませんが、ただ、人工知能のフリをする技術にかけては誰よりも経験値が多いと思っています (笑)。

大澤：知能があるように見せることも含めて、僕達は知能だと捉えています。人に合わせて話すというのもインテリジェンスですから。今まで我々は、これがいろいろな理由でできていませんでした。例えば AI 研究者の多くはコミュニケーションが上手ではなく、そういうことに目が向かなかったことも理由の一つでした。それが賢いことだということは、なかなか理解されないというのも昔からあります。でも、本当にこれは、今すごく求められる技術だと思います。

## 6. 今すべきことは日本語という

### 厄介な言語と向き合うこと

大澤：『シーマン』を海外で売る際に、文化的な違いから斎藤さんのアイディアやノウハウが伝わらなかったり、変えなければいけないところはありましたか？

斎藤：あったと思うんですが、僕は『シーマン』の英語版については何もタッチしていないので把握していません。ただ、多国語展開の経験値はあります。例えば、『ザ・タワー』(OPeNBook, 1994) というゲームをつくったときに「夕方になると、なぜカラスが鳴くのか？」とアメリカ人から言われたことがありました。確かにカラスが夕方に鳴くものだと思っているのは日本人だけです。カラスは昼間から鳴いていますよ (笑)。

大澤：カラスといえば夕暮れというのは日本人のマインドセットに染みついているんですね (笑)。

斎藤：たぶん歌のせいでしょうね。あるいは「なんで夏になると蝉が鳴くのか？」と聞かれたこともあります。向こうでは蝉が鳴くのは 11 年に一度

の大発生のときだと思われているんです。グローバルに出て行ったとき、そういう文化的な表現の差は本当に出るなと思いました。単語に関しても、イヌイットの言語だと雪の名前が十何種類もあるとか、英語でも牛を表す言葉は多いですよ。そういうエピソードはたくさんありますし、話題としては面白いんですけど、それを考え始めると、たぶん一生終わらないので。

僕が言う発話的な研究をされている AI の方々が今取り組むべきことは、グローバルイゼーションではないと思います。日本語という一番厄介な言語をクリアしたら、その中にヒントは自ずとあるはずですから、まずは日本語に向き合って、その中から演繹的に法則を見いだして他の言語に応用すればよいのではないかと思います。

大澤：今は世界中で絵文字が使われているように、間とか文脈のようなものを相手に伝えたいという需要はあるはずなので、おっしゃるように日本語を掘っていくことで、他の言語にも応用できる法則を見いだせると思います。

斎藤：そうですね。何か新しい言葉がボンと出てきたときに、文字を見ただけではその言葉の使い方がわからないことって多くないですか？

メロディがわからないと使い方がわからない。だから、日本流行語大賞というのが出てきたのはテレビとともになんです。

大澤：僕は「神ってる」という言葉を最初に聞いたときはテレビを観ていなかったの、まったく何が起きているのかわからないので面白さがわかりませんでした。それは発話された文脈がわからないからですよ。

斎藤：あとは使い方のタイミングですね。レッドツェッペリンの「Rock and Roll」という曲があるんですけど、そのイントロがすごく印象的なんです。ドラムから入るのですが、それが何拍子なんだかまだにわからないんですよ。ものすごく複雑なビートでコピーするのも大変なんで

す。だから印象に残るんです。プロのミュージシャンに聞いたところ、「あれは、あのまま覚えるしかない」と言われました。つまり、あれ一つがパッケージなんだという考え方ですね。さっき言った「お控えなすって」は、あれ一つでパッケージだから分解して考えたらいかんということですよ。そういうものが日本語の口語にはたくさんあります。だから、そのフレーズが何ワードで出来上がっているかというよりは、全部で1個と捉えましょうということですね。

大澤：そこにパラメータを入れて調整していくということですね。

斎藤：そう、そしてそれは言葉の変形ではなく意味の変形かもしれませんね。多くある慣用的な言い回しのほとんどが、メロディとパッケージとして記憶されている。「おはようございます」や「ぼちぼちでんな」なんかもそうですね。それらを構成している要素をいくつか分解してルール化したのが今の文語の文法だとすると、文語の文法は口語に関しては無力なので、新しいパッケージの組合せの法則を定義していく必要がある気がしています。

今日は偉そうにいろいろと言いましたが、まだまだ僕自身も発展途上で迷いの連続です。今後ともいろいろと皆さんにお力をお借りできればと思います。

三宅：今回のインタビューは学会誌に載りますので、それがきっかけになると良いですね。

斎藤：ありがたいことです。ぜひ、シーマン人工知能研究所と一緒にやってくださる方とつながれればと思います。

三宅：本日はありがとうございました。

2017年2月10日 受理

## —— 著者紹介 ——



### 斎藤 由多加

1962年生まれ。東京都出身。ゲームクリエイター。オープンブック株式会社代表取締役社長。『The Tower』、『シーマン ～禁断のベット～』などの作品を開発したことで知られる。また「アップルコンピュータの研究家」でもあり、マッキントッシュに関する著書である、アップル日本上陸の軌跡を綴ったノンフィクション「林檎の樹の下で」(1993～94年に「DIME」誌に連載)が、2017年夏にコミックにて復刊予定(光文社)。早稲田大学理工学部建築学科卒業。株式会社リクルートを経て1994年オープンブック株式会社を創業。1994年高層ビルシミュレーション『タワー』の国内外のヒット(海外名はSimTower)で全米パブリッシャーズ協会賞ほか受賞。1995年日経BP社『ペンチャー・オブ・ザ・イヤー』最優秀若手経営者部門賞。1999年『シーマン ～禁断のベット～』をドリームキャスト向けに発売。文化庁メディア芸術祭優秀賞、米国GDC「年間キャラクター賞」はじめ受賞多数。2006年ゲーム『大王』を発表。2012年ゲーム『エアロポーター』を発表。2014年大手住宅メーカーの『喋る家』開発など先端技術分野に関与。Twitter @yootsaito



### 大澤 博隆 (正会員)

2009年慶應義塾大学大学院理工学研究科開放環境科学専攻博士課程修了。2009年慶應義塾大学訪問研究員および米国マサチューセッツ工科大学AgeLab特別研究員。2010年日本学術振興会特別研究員PDに採択され、国立情報学研究所へ出向。同年から2011年にかけて、JSTさきがけ専任研究員に従事。2011年より2013年まで、慶應義塾大学理工学部情報工学科助教。2013年より現在まで、筑波大学大学院システム情報工学研究科助教。ヒューマンエージェントインタラクション、人工知能の研究に従事。情報処理学会、日本ロボット学会、IEEE、ACMなどの各会員。博士(工学)。共著『人とロボットの“間”をデザインする』(東京電機大学出版局、2007)、『人狼知能 だます・見破る・説得する人工知能』(森北出版、2016)、『AIと人類は共存できるか? 人工知能SFアンソロジー』(早川書房、2016)。



### 三宅 陽一郎 (正会員)

株式会社スクウェア・エニックス テクノロジー推進部リードAIリサーチャー。1999年京都大学総合人間学部基礎科学科卒業。2001年大阪大学大学院理学研究科修士課程物理学専攻修了。2004年東京大学大学院工学系研究科博士課程(単位取得満期退学)。同年、株式会社フロム・ソフトウェア入社。2011年退職。同年、株式会社スクウェア・エニックス入社。現職。本誌編集委員、日本デジタルゲーム学会理事、芸術科学会理事、国際ゲーム開発者協会日本ゲームAI専門部会代表、CEDECアドバイザリーボード。単著『人工知能のための哲学塾』(BNN新社、2016)、『人工知能の作り方』(技術評論社、2016)。共著『デジタルゲームの教科書』(SBCr、2010)、『デジタルゲームの技術』(SBCr、2011)、『絵でわかる人工知能』(SBCr、2011)、翻訳監修『ゲームプログラマのためのC++』(SBCr、2011)、『C++のためのAPIデザイン』(SBCr、2012)。デジタルコンテンツシンポジウム第4回 船井賞受賞(2008)、CEDEC AWARDS 2010 プログラミング・開発環境部門優秀賞、日本デジタルゲーム学会2011年若手奨励賞受賞。デジタルゲームにおける人工知能技術の理論的確立と実際のゲームタイトルへの具体的導入に従事。



### 高橋 ミレイ

編集者。ギズモード・ジャパン編集部を経て、2016年10月からフリーランスに。デジタルカルチャーメディア『FUZE』創設メンバー。テクノロジー、サイエンス、ゲーム、現代アートなどの分野を横断的に取材・執筆する。関心領域は科学史、哲学、民俗学など。