

# AIと言語学

文学研究科英文学専攻博士後期課程満期退学

名倉 秀人

## 1. はじめに

世の中がAIブームで騒がしい。「AIが人の仕事を奪い、失業者が大量に出る」とか、「AIが人の仕事をしてくれるので、人は仕事をする必要がなくなり、人生が豊かになる」など、まるでSFの世界の話のようなことが言われ始めている。では、AIとは何か。AIとはArtificial Intelligenceのinitialsで、人工知能のことを言う。人工知能学会によれば、AIの定義には二つの立場があり、以下のようにまとめられる。

定義A：「人間の知能そのものをもつ機械を作ろうとする立場」

定義B：「人間が知能を使ってすることを機械にさせようとする立場」

その研究の目的はほとんど定義Bである。人型ロボットを作る過程の上で人間の脳を作り出そうとするものではない。しかし、定義Bの研究の上で定義Aの研究は必須である。なぜなら、人の思考の真似事をさせようという点では同じで、究極の定義Aを目指さないわけにはいかないからだ。その上で人の役に立つ方向性を与えていく形になる。ただし、定義AのAIを完成させないまま、定義BのAIを稼働させることはできる。AI搭載の自動掃除機は喋ることはない。人に勝利したことで有名になった囲碁のAIである"AlphaGo"に将棋はできない。それでもAIは私たちの生活に食い込んできていることは確かだ。

私たちの生活に入ってきたAIはよく話す。人の話も聞いてくれる。去年あたりから日本でも流行り出した「スマート・スピーカー」は、キーボードやスマートフォンのスワイプによるテキストではなく、音声のみで命令を受け付ける。天気を聞けば今日の天気を、ニュースが聴きたければ最新のニュースを、今の気分の音楽を聞きたければベストな音楽をチョイスしてくれる。電灯やテレビやエアコンのオン・オフでさえ、音声で受け付けてくれる。ちょっとした未来感覚だ。しかし、「彼ら」はその言葉を「理解」しているのだろうか。「機械が言葉を理解する」とはどのようなことを意味するのだろうか。個人的には人間による「理

解」とは全く違ったものを感じられたが、言葉を扱う研究者としては気になるところだ。本格的な科学的証明や実験は数学者や脳科学者に任せることにして、AIは言葉をどのように扱っており、その発展に言語学者は何を提示できるかを総論的に考察していくことにする。現在の筆者の環境においてできる実験としては3点で、taggingとAlexaと母音の音声認識とについて主に述べていく。

## 2. 言語学で使われているAI

言語学でもAIを使った研究は始められている。最も代表的なものは、コンピュータ・コーパスである。大量のテキストは現代的な言い方をすればビッグ・データだ。コーパスは100万語であったBrown Corpus (1964) からThe Lancaster-Oslo/Bergen Corpus of British English [LOB] (1978) を経てBritish National Corpus (1994) には1億語を突破した。そして現在ではBrigham Young UniversityのMark Daviesらによるコーパス群の中には、進化し続けるコーパスも出来てきた。The NOW corpus (News on the Web) は60億語のネット上のニュースや雑誌のデータを2010年から今までの間集めている。そして今でもひと月に約1億~1億2千万語のニュース記事を収集している。つまり1年に約12億語のデータが可算されていることになる。もちろん、これらのテキストデータは人間がひとつひとつコピー&ペーストしているわけではない。ボットと呼ばれるプログラムが広大なネット上を徘徊し、ニュース英語を選別して自動的に集めてくるのである。新しい英語が常にアップデートされることとなる。ネット上のデータを取り込み、自己増殖を繰り返していくThe NOW corpusは、人工知能というより人工生命体のような所作を示し始めている。

また、textファイルに自動的に英文品詞タグを付け、tagged corpusを生成するauto-taggerというプログラムがある。これもAIと言ってよかろう。定義Bがそのまま当てはまる。auto-taggerによるタグ付けは、単語の品詞の確率から判断される方法と、文法によって適切な品詞を付与する方法がある。Lancaster大学によるCLAWS (Constituent Likelihood Automatic Word-tagging System) はその両方を取り入れたものである。CLAWSは購入可能で、実際の使用感は2016年の東洋大学大学院紀要で示したが、その精度は96-97%であった。一見高い精度に見えるが、100個の単語のうち3、4個のミスがあるのは大きい。これは状況や文脈によって品詞が変わるためである。

- (1) a walking stick
- (2) a walking man
- (3) walking men
- (4) She was walking.
- (5) I like walking.

上の4つの文をCLAWS7でtag付けをし、考察する。

- (1t) a\_AT1 walking\_NN1 dictionary\_NN1
- (2t) a\_AT1 walking\_NN1 man\_NN1
- (3t) walking\_VVG men\_NN1
- (4t) She\_PPHS1 was\_VBDZ walking\_VVG .\_.
- (5t) I\_PPIS1 like\_VV0 walking\_NN1 .\_.

walkingに付加されたTag Setは次のようになっている。

NN1: singular common noun

VVG: -ing participle of lexical verb

文法的に言えば、(2t) は間違いである。複数形である (3t) では正確に現在分詞と判断しているが、(2t) では、名詞として判断している。何故CLAWS7はこのようなミスを犯すのだろうか。理由は単純で、前の不定冠詞を判断基準としているため、NN1のtagを付けたのだ。a walkingで名詞と解釈したのだ。その後ろのmanをNN1とtag付けしたのは、文法的理由ではなく、語彙的なものである。manが名詞以外になる確率は非常に低いため、文法的圧力より語彙の持つ品詞性を重要視する。I would like to water the flowers.という文のwaterにもNN1のタグを付ける。名詞が羅列していても何の不思議とも思わない。それがCLAWS7というAIの限界なのである。CLAWS7に動名詞のtagはない。auto-taggerとしてそこは「逃げた」部分であると思われる。

つまり、CLAWS7は「思考して」「理解して」tagを付けているわけではない。機械的に付けているだけである。AIとしてのレベルは低い。CLAWS7が公開された2010年には、まだ人工知能の概念がauto-taggerに取り入れられていなかった。1対1対応のtaggerにほんの少しだけ文法的な思考法を加えただけのものであった。では、今後auto-taggerの正確性を高めるために必要なことは何か。筆者はdeep learningによる確率判定を導入すべきだと思っている。例えば、waterにNN1とひとつのtagだけを付けるのではなく、VV0の原形動詞の可能性を加える。さらに主語が複数のかの現在の動詞である可能性を加える。CLAWS7には原形動詞と複数主語現在の動詞との区別がなく、全てVV0としているので、仮にVVPというtagを作り、ここを区別する。NN1, VV0, VVPの区別は文法で行うより先に、collocationで行う。would like toと来たら次は名詞の確率はなく、100%VV0だとコンピュータに覚え込ませる。そのパターンはdeep learningによって世界中のあらゆるコーパスを参

照し、学ばせる。そこに「理解」は必要ない。コンピュータの碁が強くなったのは、全て考えなくて済むようになったからだと言う。碁石の白黒のパターンを論理ではなく経験で判断し、この画像パターンなら勝利への一手はこうだと決めている。それと同じようにauto-taggerもdeep learningによるパターンができれば、より正確なtag付与が可能になる。問題は、言語の広がり無限であることにある。碁や将棋は、決まったマスの中で決まったルールで「勝利」という目的を持って進むことができる。言語のルールは曖昧で、ときにはルールから外れた表現でさえ言葉としては正解であることもある。その膨大さに現在のコンピュータはまだ対応できていない。だが、将来今よりもはるかに高い精度のtag付けが可能になれば、自動翻訳機や顧客対応のコミュニケーション・ロボットに応用できると思う。

### 3. AIの歴史

ここでAIがどのように発展していったか、簡単におさらいをしておこう。1940年代、エニグマの暗号解読で有名なAlan Mathieson Turingが初期コンピュータの設計・デザインを提示した。コンピュータは日本語訳でも「電子頭脳」と言われ、感覚的には今のAIに近いものであったことだろう。最初のコンピュータの目的は計算である。「人間が知能を使ってすることを機械にさせようとする立場」という今のAIの立場から言っても、計算機はAIのカテゴリーに入れるべきだ。一方、彼は*Computing Machinery and Intelligence*という論文で「チューリング・テスト」というものを考案した。機械が「知的」であるかどうかを判断するものである。「人間の質問者が機械と会話をして人間か機械か判別できない場合に、その機械が『思考』していると言えるというものである」という考え方だ。これには賛否両論があるが、今でもAIを作っていく上での判断基準のひとつとなっている。2014年に、テストに13歳の少年の設定で参加したロシアのスーパーコンピューターが、30%以上の確率で審査員らに人間と間違われ、史上初の「合格者」となった。これは先に挙げた定義Aを目指すものである。

しばらく大きな発展はなかったものの、1980年代、PrologやLISPといったプログラム言語に新たな機能が加わる。Aのプログラムで上手くいかない場合、Bのプログラムを試す。次にCのプログラムを試す。それでも上手くいかない場合、コンピュータが自分で新たな処理法を作る。それでもだめなら諦める。そのようなルーチンが可能になった。自己解決型の処理は、後のdeep learningに繋がっていく。チェスで人に勝ち、将棋で勝ち、ついには最も困難と言われていた囲碁の世界で人類のチャンピオンに勝利するようになる。ルールを最初に教えただけで、あとはコンピュータ同士で無限の模擬戦を行う。疲れることなく、休むことなく。それは画像として処理され、確率的にどうすれば勝てるかを判断する。そして人類が考え付かなかった手を打つようになる。定義Bの目的であれば、人を超えたAIはすでにいくらかでもあるということだ。

一方、定義AのAIも発達し、先ほど述べたスマート・スピーカーによるコンピュータとの音声コミュニケーションが可能になった。まだ不完全ではあるが、映画『2001年宇宙の旅』中のコンピュータHALのように、人と見間違えるくらいのコミュニケーションが取れる日はそう遠くないように思える。

#### 4. discourseのAIモデル

定義B型のAIは目的がはっきりしているため、その目的に合わせて思考モデルを開発している。掃除なら掃除のための、自動運転ならそのためのモデルだ。これは多種多様過ぎるし、本稿の主旨からも離れるため、定義A型のモデルを考えてみることにする。定義Aは「人間の知能そのものをもつ機械を作ろうとする立場」であった。これはアンドロイドを作る上で必要なAIを人間の脳に似せていこうという目的である。目的と言ってもその範囲は無限に広い。人は何を目的として生きているのかということにも繋がってしまう。もちろん、これは大きな問題だが、これを扱うのもやはり哲学者や脳科学者たちだ。では、言語学者はどこにたずさわることができるか。それはdiscourseであると考え。人は主に音声を使ってコミュニケーションをする。そこにはcohesion（結束性）やcoherence（一貫性）があり、でたために言葉が並んでいるわけではない。そして単に言語使用だけではなく、social practice（社会的実践）もAIに教え込まなければならない。社会の中で人はどのように言葉を選択し使っているか。応用言語学者が常日頃研究していたことがここで役に立ってくる。言葉の部分に限定しているのだから定義Aではなく定義Bなのではないかという考えもある。だが、人間の中で言葉ほど抽象的ではあるが大きな役割を担っているものはないと思う。自動車の運転も掃除も料理も人間にとっては重要なことで、目的もはっきりしているが、人を人たらしめるものとしては弱い。ロボットに人の歩行を完全に真似させることが難しいことと同じくらい、コミュニケーションをさせることは難しい。チューリング・テストに合格することができるAIの完成にはもう少し時間がかかるだろう。最近のコンピュータの発展を見ているとそうも言い切れないが。

さて、コンピュータに音声を聴かせ、反応するモデルは、単純には以下のようになる。



人間の理解には基本的に「テキスト化」という工程はない。人間は抽象的に音声のみで処理できるからだ。一方、コンピュータの処理にはテキスト化は必須だ。コンピュータの基本構造がテキストを必要としているためだ。音声認識からテキスト化の分野は音声学のカテゴリーに任せられる。入力された音声はバイナリデータと言って、その音声を2進法の数値に変える。そのデータ上、発せられた音声の値がどこに位置するかで発音された音を具体化す



る。発音記号化していると言っても良い。その発音記号がどの単語に当てはまるかをデータベースから検索し、テキスト化する。次に処理だが、まず、形態素解析とdeep learningで蓄積したパターンを照合し、目的に合った（人間らしいと言っても良い）テキストを作る。それをtext-to-speechのソフトウェア技術で音声化する。かなり単純化したが、これが人間の発話に対するAIの反応というモデルである。大きな問題は「処理」の定義だ。これは実際にAmazonのスマート・スピーカーのAIであるAlexaを使って検証して行こうと思う。

## 5. Amazon EchoのAlexa

今年（2018年）に入ってスマート・スピーカーの市場が賑わっている。AmazonのAmazon Echo、GoogleのGoogle Home、LineのClovaなどがある。値段も3千円～2万円程度で、手に入らないものではない。コンピュータやスマート・フォンが世の中に出たときと同じように、「こんなもの何の役に立つんだ？」という目で見ている人がまだ多いだろう。筆者も同じような感覚ではあったが、本稿を書くためにAmazon Echoを購入した。Amazon EchoにはEcho Dot, Echo, Echo Plusという3種類があり、スピーカーの性能の差だけで、AIはみな同じである。基本円筒状で、Echoは500mlのペットボトルより少し低く、少し太い。最近Echo Spotという映像付きのものも出てきた。AIにはAlexaという名が付いている。AlexaはAmazon Echoの中にあるのではなく、クラウド上にある。クラウドを説明すると長くなるが、インターネット上にマスターとなるコンピュータがあり、ネットを介して処理をしてるということだ。Echoは、入力された音声をバイナリにしてクラウドに上げる作業と、クラウド上で処理され、送られてきた音声データを再生するだけの能力しかない。他の処理は全てネット上にあるAlexaが行う。使用者にはクラウドの意識などないので、Echo本体が話してくれているような感覚になる。

Alexaは自分から発話することはない。ユーザーが話しかけて初めて反応する。ユーザーがAlexaに話しかけているのか、それとも家にいる誰かと話しているかの判断は、コンピュータにはまだ不可能だ。それゆえ、「君に話しているんだよ。Alexa。」ということを知らせるために、話す前に必ず"Alexa!!"と呼びかけなければならない。それを「プロンプト状態にする」という言い方をしている。コンピュータを昔から使っている方々にとっては懐かしい表現だが、Microsoft Windowsの前のOSであるMS-DOSで、Cドライブでコマンドを待っている状態を"C prompt"と表現していた。まさにこの状態で、ユーザーからの音声コマンドを待っている。よって、一つの質問に対してAlexaが反応したら、そこで会話は終わる。Alexaが反応した内容に対してさらに質問をすることはできない。つまり、本当の意味での会話はできないのである。

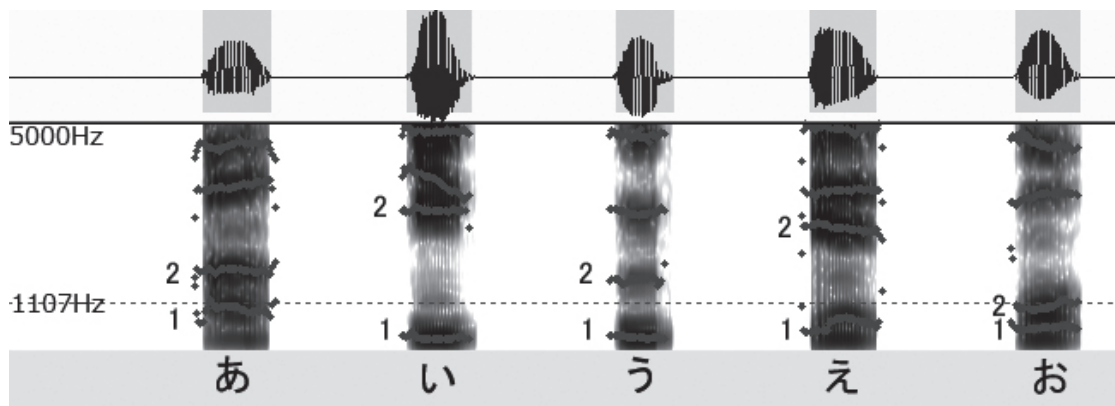
英語モードもあるが、筆者の発音の問題もあるので、日本語で「処理」を検証してみる。例えば、「アレクサ、天気教えて」と言うと、地元の天気を教えてくれる。「天気を教えて」

でも「今日の天気は？」でも「天気？」でも同じ答えをする。だが「今日は暑いかな？」と尋ねると「そうですね。本当に暑いですね」と答えた。尋ねた日はそれほど暑くなかったのだが。これだけでも、Alexaが「天気」という単語が天気予報の情報へつながるトリガーとなっていることがわかる。だから「天気教えなくていいよ」と言っても天気を教えてくれる。つまり、ユーザーの言葉を「理解」しているのではなく、「反応」しているだけなのだ。その意味ではCLAWSと同じ立場にある。しかし、Alexaはコマンドを受け付けることで、その命令パターンの情報をユーザーから回収し、パターンにどのようなものがあるか、その確率を参照し、どう反応すべきかをAlexa自身が考え、反応している。そういう意味ではAIであると言うべきだ。一方、「天気教えなくていいよ」というコマンドを発するユーザーはおそらく極端に少ない。よって間違った反応をする。だからと言ってAIと呼ぶにはあまりにも貧弱過ぎるという評価はできない。

ここで考えたことは、Alexaは何故理解しようとししないのだろうかということだ。人間との大きな差は何か。まず、願望、痛み、喜びなどの感情がない。例えば「明日遠足だから天気が悪かったら嫌だな。だから正確に天気を知りたいな。」という思考は生まれえない。遠足が楽しいものだと判断するには、五感というインターフェイスとそれに伴って生まれる快感がなければならない。そしてもう一つ、その情報を他の人に伝えたいという気持ちがない。言葉が生まれたのは、自分の脳内にある情報を他人と共有したいと人が思ったからだ。または他人の脳の情報を自分に取り入れたいと思うからだ。お腹も空かない、痛くもない、パートナーも欲しいと思わない、それを伝えたいとも思わない現在のコミュニケーション系AIにdiscourseを望むことの方が間違いなのかもしれない。社会性や常識を教えるには、AIに自分が人間と同じであるということを認識させなければならない。そうするためには、本当にアンドロイドを作って社会生活をさせ、そこからフィードバックさせるしか方法はない。コミュニケーションに特化したAIがチューリング・テストに合格するのはいつの日のことになるだろうか。

## 6. 数値による音声認識

音声認識サービスには、MicrosoftのCortana、AppleのSiri、GoogleのGoogle Assistant、そしてAmazonのAlexaなどがある。いずれもスマートフォンの検索機能を中心に開発され、それがスマート・スピーカーに応用されている。音声認識の目的はテキスト化である。アナログな音波として発せられた音をどのように解析し、デジタルであるテキストと変化させるか。これは音声学で研究されてきた分野である。たとえば、Praatという音声分析ソフトを用いれば、人がどのように発音しているかを数値で証明することができる。ならば、逆にその数値を見て、どんな音が発音されているか判断することもできるはずだ。それは音声解析の第一段階となる。人間は数値で直感的に判断することは難しいので、グラフ化し、そこか



ら大まかな音声を推察する実験を試みることにする。

音声は高さ、大きさ、有聲か無聲か、という基本に加えて人の喉から口への調音器官 (articulatory organs) が生み出す、様々な要素の束で出来ている。それを弁別素性 (distinctive feature) と言う。たとえば子音 /t/ の弁別素性は {子音、調声位置は歯または歯茎、無聲、破裂音} という束になる。そこにさらに人間の個々のノイズが乗る。また、その前後にどのような音素が繋がっているかによって変化することもある。年齢、性別によっても変わってくる。これらの分析を全て本稿のみで行うことは難しい。よって、最初の試みとして、個々の特徴をできるだけ削った音声を素材として音声の画像を作成していく。特徴を削った最も良いサンプルはtext-to-speechで発音させた音声だ。これは弁別素性を機械的に作成し、ノイズをできるだけ少なくして聴き易いことを目的としている。これで音声分析の逆の作業ができるかを探る。その音を分析することができれば、そこから徐々に様々な要素を加えていけばよいことになる。

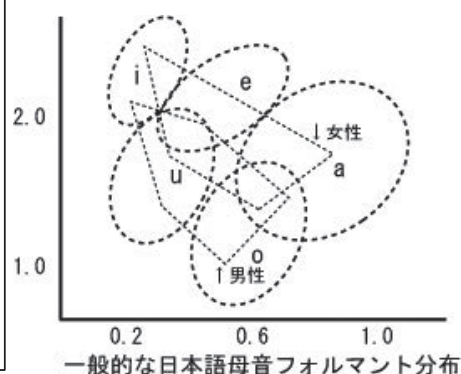
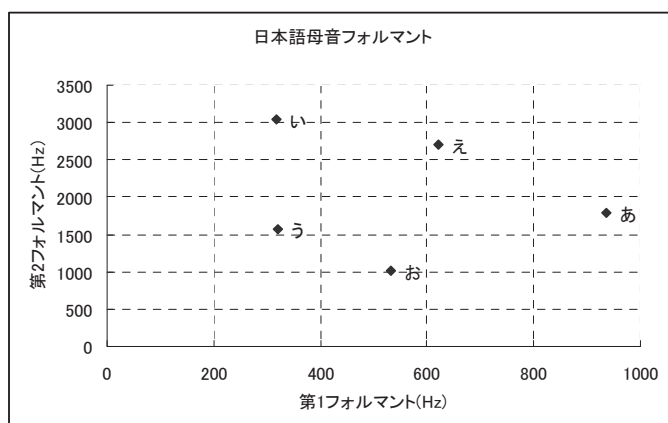
まず、日本語の「あ、い、う、え、お」の5つの母音をPraatで解析してみる。英語は母音の数が多いのでエラーの出る可能性があるためだ。子音は音の要素がさらに多いため、今回は扱わないことにする。母音の弁別は第1フォルマント (約500~1000Hz) と第2フォルマント (約1500~3000Hz) によって行うことができる。text-to-speechのソフトウェアで発音させ、mp3ファイルで保存。性別はデフォルトの女性を使った。それをPraatで解析する。

今回はPitchとIntensityを外して表記した。上部の画像は波形とPulsesである。上部の薄い灰色の部分がPulsesで、全ての音に見られる。有声音と母音にPulsesが現れる。下の濃淡はSpectrumを表し、その一番濃い部分に点の集まりで表されているのがFormantである。下から、第1フォルマント、第2フォルマント…となっている。その高さと幅が各々違うことが確認できる。フォルマントの数値を抽出する際、発音されている部分だけを選択するには、Annotationを使う。selectするとき、silenceを指定すると無音の部分と音声のある部分に分けてくれるので、音声の部分の平均の数値を使う。結果は以下のようになった。

これをグラフにする。



	第1フォルマント (Hz)	第2フォルマント (Hz)
あ	935.3	1780.7
い	317.3	3048.2
う	321.6	1573.2
え	619.8	2702.5
お	531.9	1009.0

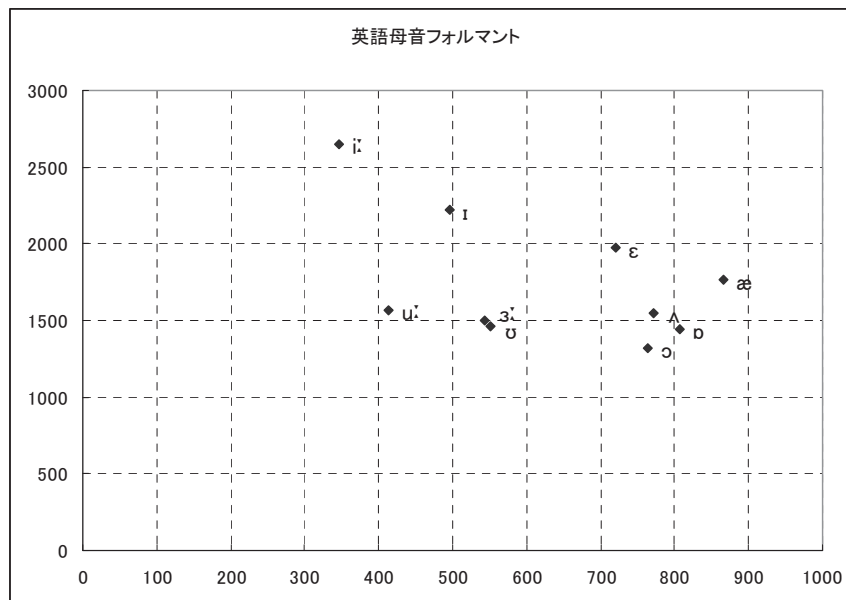


この結果は、一般的な女性の母音フォルマントと比較しても差はなかった。「お」の音の第2フォルマントが少し低かったが、許容範囲である。これはひとつのサンプルである。このデータを大量に集め、確率論的に一般化すれば、普通の女性の話す母音の基本的な数値を定めることができる。つまり、母音を判断できるということになる。また、機械的にこの数値の音波を発生させれば、人の母音に近い音を作ることができる。音声合成である。シンセサイザーでロボットの声のような音を出すのにこの方法を用いている。

では、英語ではどうだろう。英語はテキストに対して母音が1対1対応ではない。TTSから直接母音のみをサンプルとして取ることは難しい。単語でサンプリングすれば、その前後の音の影響を必ず受けるからだ。よって一定の条件である/h/ と /d/ の子音で挟んだ母音を抽出。無声音から始まり、/d/の前で音量が下がるので母音の抽出はしやすいためだ。heed (hi:d), who'd (hu:d), hid (hid), heard (hɜ:d), hood (hʊd), head (hɛd), had (hæd), hud (hʌd), hawed (hɔd), hod (hɒd) の10の母音について調査した。二重母音は除いた。

日本語母音のフォルマントの表と比較すれば、近い位置に英語の母音がある。もう少し精度が高ければ、二つのフォルマントの数値から正確に英語母音を限定することができよう。二つのフォルマントは口の開け方と舌の位置を表している。表を見れば、/i:/ は /ɪ/ を単に音を伸ばしているだけではなく、二つの要素がそもそも異なっていることに気付く。/ɪ/ は /i:/ よりもかなり /ɛ/ に近い。これを利用すれば、音声矯正もできる。実際そのようなソフトウエアも開発されている。正確なテキスト化ができれば、翻訳機も作れる。AIにデータ

	第1フォルマント (Hz)	第2フォルマント (Hz)
heed (hi:d)	345.8	2644.4
who'd (hu:d)	412.8	1565.1
hid (hɪd)	495.3	2218.2
heard (hɜ:d)	543.5	1497.1
hood (hʊd)	551.8	1460.9
head (hɛd)	721.3	1976.9
had (hæd)	865.2	1764.7
hud (hʌd)	771.1	1544.2
hawed (hɔd)	764.5	1323.1
hod (hɒd)	806.2	1442.7



を流し込むこともできる。AIのインターフェイスとして音という要素は非常に重要だ。PCやスマートフォンをいちいち起動する必要はなく、人とコミュニケーションを取っているかのようにコマンドを入れることができるのである。

## 7. AIと言語学との関係

テキストデータの処理の部分の導入部であるTaggingとAIの入力部である音声認識を中心に言語学がどのようにかかわってきたかを述べてきた。言葉を使ってデータを処理する限り、言語学者の力は必要になってくる。1960年代に今のスマートフォンが想像できなかったように、将来AIがどのように発達するか予想できるものではない。あと数年もすれば、この話や考え方も古くなる可能性が高い。だからこそ今の状況を記しておきたかった。しかし、言語学者がAIに寄る必要はない。言語学者は自分の学問を究め、AI開発で必要だと言われた

とき、その知識と思考法を提供すれば良いのだと感じた。

## 参考文献

- 乾健太郎 ほか 2018. 『人工知能のきほん』 ニュートンプレス
- 埋橋勇三 1988. 「英語学研究とコンピューター」 白山英米文学 第13集 pp. 31-69.
- 黒川伊保子 2017. 『アンドロイドレディのキスは甘いのか』 河出書房新社
- 篠田浩一 2017. 『機械学習プロフェッショナルシリーズ：音声認識』 講談社
- 名倉秀人 2016. 「Tagger についての一考察」 東洋大学大学院紀要 第53集 pp.225-239.

## 参考サイト

\*筆者による最終アクセスは2018年9月20日になります。

宇都木昭 2012 「Akira Utsugi's web site Praat 入門 母音のフォルマント」, <<https://sites.google.com/site/utsakr/Home/praat/formants>>

人工知能学会 「人工知能のやさしい説明「What's AI」」, <<https://www.ai-gakkai.or.jp/whatsai/>>

鈴木恭子 ほか 2017 「Amazon Alexaとは何か？アマゾンの音声認識技術の使い方を基礎から解説する」, <<https://www.sbbbit.jp/article/cont1/34026>>

# AI and linguistics

NAKURA, Hideto

In late years, it is sure that linguistics is getting into our life. Vacuum cleaners and microwave ovens talk. Smart phones tell people how to reach the destination by sound. "Smart speakers" can understand speech, process it and react by speech. Speech understanding means the technology of speech-to-text. Processing the text data needs morphological analysis. Speaking of the text data after processing needs voice synthesis. These consist of phonetics, grammar, corpus linguistics and so on.

As long as AI (artificial intelligence) uses texts as a medium between machines and us, linguistic knowledge is indispensable to AI development. On the other hand, linguists are not interested in AI so much, I think. So, I wanted to write an outline about the AI for linguists. I'm not a specialist of AI. I can't argue about machine science or mathematics. But I can tell it from a linguists' point of view.

In this article, first, I investigated auto-tagger as AI. Next, I wrote up the history of AI and thought of discourse. Last, I wrote "Alexa" and voice-to-text.