

はじめに

「この発明は科学史における画期的なできごとです。エジソン氏の装置の音を聞くのは大いなる喜びであり……それに感謝を捧げたいと思います」¹⁾

「貴殿のすばらしい発明のことで頭がいっぱいで、仕事を手につかないほどです。その成果は(科学において)広範におよび、その可能性は計り知れません」²⁾

一八七七年一二月、トマス・エジソンは自作の蓄音機で〈メリーさんのひつじ〉を録音し再生したことで歴史に名を残した。単に「科学史における画期的なできごと」であるばかりでなく、人間の声にとって革命的なできごとだった。それまで、人が話すのを聞くのは常に「生」^{なま}の経験であり、話手の口から声が出る瞬間に聞くしかなかった。エイブラハム・リンカーンのゲティスバーグ演説など、

蓄音機の発明以前のすぐれた演説を文字で読むことはできるが、このときの大統領の話しぶりを正確に知ることはできない。一方、蓄音機は話し方をとらえることができた。話し方というのは、語られる言葉そのものに劣らず大きな意味をもちうる。誰かが「大丈夫だ」と言った場合、その言い方によってじつは大丈夫ではないということが伝わるかもしれない。

声はそれを発する人が何者であるかを如実に語る。友人や家族が電話をかけてきたとき、ほんの数語だけ聞けば、相手が誰だかわかる。これは驚くべきことだ。赤の他人が電話をかけてきた場合には、すぐさま訛りやイントネーションを聞き取り、相手の身分や出自や教養の程度を推測する。年齢や体格や性格についても想像するが、間違っていたり、先入観や偏見によってゆがんでいたりすることも少なくない。人に与える印象を変えるために、話し方を加減することもある。私たちはいわば言葉のカメレオンであり、生まれ故郷にいるときには無意識のうちに地元の訛りを丸出しにし、別の土地ではその場に溶け込めるように訛りを抑える。話し方の特徴というのは意外に柔軟なのだ。

声がどれほど私たちの自意識を形成しているかについては、自分の声の録音を聞いたときの狼狽を見ればよくわかる。自分の声はほかの人が聞いているよりも低く聞こえる。これは骨の振動によって音が喉頭から耳へ伝わる際に、低音が強調されるからだ。録音した声を聞けば、他者に対して示している声の個性が自分の思うそれとは違うということがたちどころにわかる。エジソンが蓄音機を発明するまで、人間はそのような違いを知らない幸せな時代を生きていた。

人類の「声の歴史」は三つの時代に区分でき、その一つの幕を開けたのが蓄音機だ。人類も昔はほかの動物と同じく、他者に影響を与えるために単純な声を出すだけだった。敵を撃退したり、危険を知らせたり、配偶相手の気を引いたりするために声を上げていた。言語の誕生とともに、第二の時代

が始まった。言語のおかげで、集団で大規模な仕事をしやすくなり、世界を支配することも可能となった。それでも依然として、話し声の多くは他者の考えや行動に影響を与えることを目的としていた。たとえば親が幼い子どもに道路へ飛び出してはだめと言ったり、ヘンリー五世が「もう一度あの突破口へ突撃だ」²「ヘンリー五世」シェイクスピア著、小田島雄志訳、白水Uブックス」と言って兵士を鼓舞したりした。しかしその一方で、私たちは楽しみのため、あるいは気晴らしをしたり、世界とかかわったり、愛を打ち明けたりするために、言葉を発することもある。それから第三の時代の始まりを告げた蓄音機などの技術のおかげで、多数の人に声を届けることが可能になり、ときには破滅的な影響をもたらされるようになった。あるドイツの大臣はニュルンベルク裁判で、ナチスの独裁政権は「国を支配するためにあらゆる技術的手段を完全に使いこなした」最初の例であり、「ラジオや拡声器などの機器を利用して、八〇〇〇万人の国民から個々人の自由な考えを奪った」³と述べた。私たちは今、人工知能（AI）によってコンピューターと対話できる刺激的な時代を迎えようとしている。人類に限られていた言語コミュニケーション能力が人類だけのものではなく、よかれ悪しかれ、機械もこの能力をもち始めている。

本書は、「話すこと」と「聞くこと」がどのように進化したか、人間が幼児期にこれらの大事な能力をどう発達させるのか、そして人間のコミュニケーションがテクノロジーによってどう変わるうとしていくのかについて語る。私たちは会話をするのにすっかり慣れきっているので、会話など単純な作業だと思っている。しかし実際には、話すことと聞くことは私たちの体と心が担う作業のなかでもとりわけ複雑なものなのだ。話すには、さまざまな脳領域からの細かい指示に従って、各器官がきちりと協調して働く必要がある。人の発した言葉を聞き取ったり、声のトーンから言葉の真意や気分

のヒントを読み取ったりするのは、ものすごく込み入った仕事だ。ふつうはこれらのプロセスは隠れていて目に見えないが、心理学者、神経科学者、生物学者たちはそこで起きていることを次々に解明している。かつて会話とは媒介を挟まずに直接交わされるものだったが、今日の世界では会話の多くがテクノロジーによって伝達され変形されるものへと移行しつつある。コンピューターとの会話が身近になるにつれて、テクノロジーの影響は拡大していく。私たちはどんな秘密を機械にうっかり漏らすことになるのだろうか。AIは、どんな仕組みで話したり聞いたりするのか。そしてこの先、人間の話し方をどう変えていくのだろうか。

話すことや聞くことに技術がどう影響するのか、それを教えてくれる好例が蓄音機だ。一八七八年二月、蓄音機が初めてイギリスで聴衆に披露された。実演の舞台となったのは王立研究所。ヴィクトリア朝時代に偉人たちが集い、科学や工学の最先端に触れることのできる場所だった。英国郵政省電信部の主任技師、ウィリアム・H・プリースがエジソンの発明品のレプリカを使ってデモンストレーションを行なった講堂には、客が詰めかけていた。そのレプリカは、アメリカから取り寄せた蓄音機の到着が遅れたため、前の週に急いで作製したものであった。エジソンと同じく、プリースも蓄音機の公開実験に童謡を使うことにして、「ヘイ・デイドル・デイドル」というマザーグースの童謡を再生した。「歌詞ははっきり聞き取れたが、声は非常に弱々しく不気味なカリカチュアであった」と『ロンドン・ウィークリー・グラフィック』誌は伝えている。工学の画期的な成果を披露するのに童謡を使ったのは、賢明な選択だった。聞き手は歌詞を知っているから、針がスズ箔を引っかくときの雑音に言葉がかき消されても、無意識のうちにそれを補ってくれるからだ。この新しい発明品は大いに注目を集めた。「蓄音機とはどんな形をしているのか、声をかけたらどうなるのか、どんな音がするの



トマス・エジソンと彼の作った蓄音機⁽⁵⁾

かと、テーブルのまわりに人が集まってきた。講堂にはいつまでも人が居残っていたが、一時になり、もう帰れと言わんばかりにガス灯が消えると、ようやく帰っていった⁽⁴⁾。それから二週間経たぬうちに、エジソンの作ったスズ箔蓄音機がイングランドに届いた。通常は公開されていないが、私はBBCのラジオ番組の制作にかかわった際に間近で見る機会に恵まれた。装置の中心にスズ箔で覆われた円筒があり、装置の右側のハンドルでそれを回転させる。左側には、動作をなめらかにするための大きなはみ車がついている。単純なじょうご型の管に向かって声を出すと、音が小さな膜に集められ、この膜が振動する。膜の裏側には針がついていて、回転するスズ箔に溝をらせん状に刻んでいく。精巧だがシンプルな設計となっていて、音声を構成する空気の振動を針の振動に変換し、この針の動きを波状の溝としてスズ箔に刻むことにより

記録する。音声を再生するときはこのプロセスを逆にたどり、針が溝の凹凸をなぞることで膜を振動させ、それによって空気分子が振動し、これが音となって聞き手の耳に届く。

エジソンの蓄音機はもはや実際に使われることのない博物館の収蔵品だが、私は王立研究所を訪れた折に、別の蓄音機でアルフレッド・テニソンの詩「庭に出よ、モード」を録音することができた。これ以外の詩は考えられなかった。というのは、蓄音機が王立研究所で初めて披露されたときにテニソン本人の立ち合いのもとで吹き込まれたのが、まさにこの詩だからだ。十分に大きな刻みをつけるためには集音ホーンのすぐそばで大声を張り上げなくてはだめで、そうしないと、再生したときに針と溝から生じる表面雑音で言葉がかき消されてしまう。私の声を再生してみると、音は弱かったが、針が溝を引っかくことでどうしても生じてしまう雑音の中でも、言葉ははっきりと聞き取れた。

初期の蓄音機の実演では、愉快な実験も行なわれた。当時は音声を再生しているときにハンドルの回転速度を変える遊びが人気だった。それを実際に聞いたある人は、回転が速すぎると「老婦人が怒っている」ように聞こえ、遅すぎると「よぼよぼの老人が水を口いっぱいに入れてしゃべっている」ように聞こえたと述べている。ビートルズは、音声を重ねたり、録音した音声を逆再生したり速度を変えたりといった音声実験をした先駆者としてよく知られている。一九七〇年代には、レッド・ツェッペリンの〈天国への階段〉などの楽曲を逆再生すると冒瀆的な歌詞が聞こえるという噂が流れ、宗教団体が不快感を抱いたこともあった。しかし曲の逆再生を最初に試みたのはエジソンで、このときには自分で録音した「マッド・ドッグ」を使った。

声に対するテクノロジーの影響は、録音した言葉と戯れることができるようになったことにとどまらず、それよりはるかに大きい。人の話し方や歌い方まで変えているのだ。私は先日、「われらをお

おつていた不満の冬も」『リチャード三世』シェイクスピア著、小田島雄志訳、白水Uブックス」というセリフを一九世紀の俳優ヘンリー・アーヴィングが蓄音機で録音したものと、現代の俳優デイヴィッド・モリッシーによる録音とを比べてみた。一九世紀の録音では、アーヴィングは大劇場の舞台に合った発声法を用い、よく響く低音の声を出している。一方、マイクのおかげで声を張り上げる必要がなくなったモリッシーは、少人数の観客を相手に演じているかのように語っていて、細かな息遣いまではっきりと聞こえる。語り方よりもさらに顕著な変化を見ているのが歌いだ。オペラ界のスーパースター、アデリーナ・パッティの残した初期の蓄音機録音と、エイミー・ワインハウスのような現代の歌の名手を比べてみるとよい。パッティの歌うオペラは清らかで甘やかに響くのに対し、ワインハウスの歌声は歌い手の性格や情念をはるかにたっぷり表現する。パッティは大きな声を出せるように、発声器官をきっちり鍛え上げる必要があった。ワインハウスのほうは、音量は電子機器に任せ、もっと自由な表現ができた。テクノロジーのおかげで、現代の音楽ではじつに多様な声が聞かれるようになった。

音は何もしなければはかなく消えてしまうが、それも録音によって変わった。今では話し方の歴史を伝える音声記録が大量に存在し、科学者はそれを分析することができる。それによって、文化と結びついた変化が明らかになっている。たとえば、ここ数十年で女性の声が低くなっているとか、ロンдон訛りのコックニーの特徴だった鼻声がロンドンのイーストエンドで聞かれる多文化の混ざり合った訛りに駆逐されたことなどが挙げられる。人間の声は人類史全体にわたって変化してきたと思われるが、そのような変化を直接観察できるようなものはいない最近だ。科学者は昔の録音と最近の録音を比べて、生涯にわたって話したり聞いたりすることで人の声がどう変化するかを調べることも



1916年、北米先住民ピーガン族の男性と民族学者フランシス・デンスモア

できる。ありがたいことに、私たちの発声器官は加齢の影響を受けにくい。声の老化が始まるのは、しわや白髪が現れる時期よりもう少し遅いのだ。

エジソンは自身の発明品についてさまざまな用途を思い描いたが、最も関心をもっていたのは人の声を生前に録音しておくことだった。声は「生き物」なので、写真よりも鮮やかに人の性格を伝える。エジソンの予想では、「死期の迫った家族や偉人の口癖や声や最期の言葉を記録するために、蓄音機は間違いなく写真機より重要になるはず」だった。この予言は外れたが、今ではビデオで大切な人の声を記録することが増えている。携帯電話などの機器のおかげで、静止した写真と張り合えるほど、動画や音声の存在感が増してきた。近ごろではAIのおかげで、亡くなった人との会話をシミュレートする「声の形見」も実現している。AIが労働者から仕事を奪う

可能性についてはさまざまな議論があるが、降霊術の会を仕切る霊能者がAIのせいで失業の危機に直面するとは、誰が予想しただろうか（もともと本人たちはとくに予見し、転職に備えて職業訓練を始めたりしているだろうが）。

二〇一五年、デジタル雑誌編集者のロマーン・マズレンコは、モスクワ市内を走っていたスピード違反の車のせいで痛ましい死を遂げた。親友でIT起業家のエフゲニア・クイダは、ロマーンともう一度話したくてチャットボットを作った。彼からもらった何千通ものテキストメッセージが残っていたので、それをコンピュータプログラムに読み込ませ、AIを利用してロマーンと同じ言葉遣いをするボットを作成したのだ。テニスの言葉は蓄音機の蝸管に刻まれたまま永久に変わらないが、ロマーンのボットはもとのテキストメッセージとは違う新しい言葉で応答してくれる。たとえば典型的なやりとりはこんなふう(9)に展開する。

エフゲニア 元氣？

ロマーンのボット まあね。ちょっと落ち込んでる。僕抜きで何かおもしろいことなんかやってないよね？

エフゲニア いろんなことが起きて、毎日が過ぎていくけど、みんなあなたに会いたがってる。

ロマーンのボット こっちも会いたいよ。これが愛つてもなんだろうね。

このやりとりをどう受け止めたらいだろう。大事な人を亡くした場合に、その人に話しかけたり頭の中で会話を交わしたりするのはめずらしいことではない。しかし、答えているのが機械だと考え

ると不気味な感じもする。この技術をめぐっては、ロマンの友人や家族のあいだでも意見が分かれていて、喜んでいる人もいれば嫌悪する人もいる。ここからさらにもう一步進んで、録音した音声を使ってロマンの声を再現したらどうなるか想像してみよう。実現可能性については、ならん問題はない。運動ニューロン疾患などの病気で声を失った人のために、本人の声に合わせた合成音声を作成するのは一般的になってきている。形見のチャットボットがテキストメッセージを送ってくることに気味悪さを感じるというのなら、大切な人の声を薄気味悪い機械がよみがえらせるのはどうなのだろう。これによって、たとえば亡くなった人にいわば永遠の命を与えるためにデジタル記録を漁ってもよいのかといった、数々の倫理的な問題が生じる。

AIは、私たちの会話を根本から変えようとしている。人にとって、話すことと聞くことは事実にもとづく情報をやりとりするだけの手段ではない。「愛してる」というフレーズには、さまざまな意味合いが込められている。コンピューターにこんな言葉を告げることなどありえないと思われそうだが、じつは毎日たくさんの人が、アマゾンの販売する音声認識ホームアシスタントの「アレクサ」に向かつて愛を打ち明けている⁽¹⁰⁾。感情を理解し表現する（あるいは、ただ巧妙に感情を模倣しているにすぎない）機械が開発されるにつれ、私たちとこれらの機械との関係は根本から変化している。二〇一三年の映画『Her/世界でひとつの彼女』では、孤独な男性がサマンサと呼ばれるAIオペレーターینگシステムに恋をするが、私たちもこの映画の筋書きからさほど遠くないところにいる。

テクノロジーが自然な会話をマスターしたら、危機にさらされるのはどんな仕事だろう。一九世紀の初頭、産業革命に反対するラッダイトは自分たちの生活を脅かす新しい機械を打ち壊した。二〇世紀の序盤に音楽の録音が普及すると、作曲家のジョン・フィリップ・スーザは、やがて「音楽を学ぶ

という高貴な修練に進んで身を捧げようとする者がいなくなる⁽¹¹⁾」のではないかと危惧した。二〇一四年、コネティカット州ハートフォードでリヒャルト・ワグナーの楽劇『ニーベルングの指環』が上演されることになっていたが、オーケストラの代わりにコンピューターを使うことに対して抗議の声が上がリ、それを受けて公演は延期された⁽¹²⁾。機械が感情に働きかける力を獲得したら、ラッダイトの役者はグロブ座に押し寄せて、シェイクスピアを演じるアンドロイドを打ち壊そうとするだろうか。AIがさらに進歩してシェイクスピアになり代わり、ロボット用の脚本を書く日が来たりもするのだろうか。

劇場では古くから、人間の真実を明らかにするために動物や亡霊や操り人形を使う伝統が受け継がれてきた。コンピューターが人間と会話するようになれば、テクノロジーも人間についてさまざまなことを明らかにするだろう。子どもが聞く力や話す力を自然に獲得するプロセスと、コンピューターにそれらの力をもたせようとする科学者の苦勞を比べてみよう。私たちは、複雑な計算をするのは難しく、会話するのは簡単だと思っている。ところが機械にやらせるなら、計算のほうがじつは簡単だ。人間の会話能力などごく単純なものだと思われているが、実際には驚異的なものなのだ。

昨今では話すことと聞くことをテクノロジーと結びつけることも増えてきたが、人間の会話能力を理解したければ、まずは蓄音機の発明よりもはるか昔に起きたことを振り返る必要がある。人間の話す能力はどのように進化したのか。ネアンデルタール人が「現生人類」であるホモ・サピエンスに遭遇したならば、互いに言葉を交わすことはできるだろうか。第1章では、この活発に議論されているトピックをテーマとして扱う。