音声合成技術がもたらす
朗読の新展開

定武禮久（じょうぶのりひさ）

はじめに

　この原稿は、拙著『朗読と著作権』（アピアランス工房刊）からの抜粋です。
高山恭介さんが作られた優れた電子書籍作成ツールの「青空キンドル」を使って、自筆原稿を原稿画像にしたものです。

　合成音声と朗読との関係について書いていますが、以下の合成音声をお聴きいただいた上でお読みいただけると、イメージしていただきやすいかと思います。

①　ＮＨＫの森田美由紀アナの合成音声によるニュース読み上げ
　　http://www.nhk.or.jp/strl/group/ningen_joho/ningen_joho07.html#
　　（真中辺のサンプル音声２つ）

②　声帯を失った大阪芸大牧泉教授の合成音声による講義
　　http://www.oki.com/jp/ap_plat-u/polluxstar/event20080415.html

③　大沢たかおの合成音声朗読サイト
　　http://www.rodoku.jp/

青空
キンドル

（1）飛躍的進歩を遂げつつある合成音声技術

ここでは、合成音声技術の可能性について、朗読と絡めて書いてみたいと思います。

合成音声技術について、最初に触発されたのは、次の3つのことからでした。

年末に日経新聞に載った、声帯を失った大学教授の原稿を、合成音声で本人そっくりの口調で読みあげて、会場からすすり泣きが起こったという記事。

「電子かたりべ」のプレーヤーから流れる合成音声の自然さ。

萌えボイスで自作曲を歌ってくれる16歳のバーチャルアイドル「初音ミク」の記事。

上記の3つは、それぞれは関連はしていないのですが、つなぎ合わせると、今後の朗読分野での発展可能性が見えてくるような気がします。

合成音声というと、これまで機械音声そのものという感じの平板なソフトしか思い浮かびませんでしたが、最近の技術進歩には著しいものがあるようです。

もちろん、生の肉声とはくらぶべくもありませんが、それでも朗読分野においても、十分に実用に耐えるところまで来ています。

声帯を失った大学教授が、「自分の声」の合成音声読み上げソフトで復帰

これは、昨年末の12月23日の日経新聞社会面に載った記事です。

がんで声帯を失った教授が、事前に声を録音して、それをもとにした音声合成ソフトで、「自分の声」で模擬講義を行ない、成功を収めたというものです。その後、教授は、この音声合成ソフトを使って講義に復帰したそうです。

「『人は死を前にすると、自分の人生は何だったのかを考えます』。大阪芸術大で化学を教える牧泉教授（56）がパソコンに文章を打ち込むと、低く落ち着いた声がスピーカーからなめらかな口調で流れた。（平成19年12月）21日、大阪市内の病院で約10人の看護学生を前に行われた模擬講義。『人は時々、映画のワンシーンのような、とても良い時間が持てます。私は死を意識した時、そんな時間をできるだけ家族と友人と共有しようと思いました』。闘病体験を基にした牧教授の思いに、室内にすすり泣く声があふれた。」

「このソフトは、あらかじめ録音しておいた音声を組み合わせ、パソコンのキーボードで入力した文章の通りに再生する仕組み。機械的な合成音しか使えなかった従来型と異なり、録音した人の声質や口調を再現できるのが特徴」とのことです。さまざまな例文を多数読み込んで、パターンを記憶して合成するという方法とのこと。牧教授は、7時間かけて肉声を収録したそうです。これは、市販されている音声入力ソフトと同じイメージかと思います。あれは、「音声→文字」ですが、これは、「文字→音声」という転換です。

このソフトは、通信機器メーカーのOKI（旧・沖電気）が開発した「ボルックスター」という音声合成ソフトだそうです。販売代理をして

いるe-bookランド社やAQUARIUS最先端研究所のサイトから、サンプルボイスの試聴、お試しサービスが利用できます。

http://www.e-bookland.net/campaign/pollux.html
http://www.memorial-voice.com/

料金は、音声の精緻度によって3段階あるそうで、料金は20万円、40万円、80万円とのことです。大学教授が模擬講義で使った音声は、一番精緻なクラスのものだった由。100万円強とのことです。上記サイトで、サンプルボイスがありますが、本人の口調にかなり「そっくり」な域まで達している印象ですから、たいしたものです。

このソフトによって手に入れた「自分の声」で、自分の原稿、文学作品などを読み上げることができる、ということが、このソフトの狙いです。このソフトを使うと、声帯喪失した人だけでなく、母親による読み聞かせ用、亡き肉親の声の再生など、幅広く使えそうです。問題は、料金で、自分用ということでどうしても高くなってしまいます。

なお、『音声図書館』というサイト（http://www.onsei.jp/）では、ボルックスターの簡易版を使っているのでしょうか（販売代理店のサイトからリンクが張られていますので）、アナウンサーの声を合成して、新聞やニュース、政府広報を読み上げています。反響音的な部分が多少気になりますが、これはまさにアナウンサーの読みそのものに近いかと思います。

なお、ご紹介した牧教授ですが、平成20年4月に、この音声合成ソフトを使って、教壇に復帰したそうです。「チャンスを与えてくださった大学の皆さまに感謝いたします」との冒頭の挨拶から始まり、満席の80人の学生を相手に、1時間15分にわたる量子化学の講義を無事終えたとのことです（平成20年4月15日付け朝日新聞）。

「電子かたりべ」プレーヤーは、既成の合成音声で読み上げが可能「電子かたりべ」というのは、青空文庫のサイトで、広告をしていますので、ご覧になった方、試して見た方もおられると思います（http://www.e-kataribe.com/portal/index.jsp）。見るところ、大きく分けてサービスは三つあります。

一つ目は、これで「朗読」をしたオーディオブックを、数百種類以上、ダウンロード販売しています。文学作品に限らず、資格試験用に、法律の条文の読み上げなどのオーディオブックもあります。作品ごとに、試聴もできますから聴いてみますと、若い女性の合成音声で、それなりに自然な感じで読み上げられます。「しみじみと聴き入る」というところまでには、まだまだ達していませんが、少々なまりのある外国人が日本語を流暢に読むとこういう感じになるんでは？　という印象です。しかし、これは文学作品でも、ストーリーを追うには、十分実用に耐えるすぐれ物だと思います。

これはペンタックスの合成音声ソフトを利用しているようです。ペンタックスのサイトで、自分で作った文章をその合成音声で読み上げさせるデモサービスを行なっていますので、試してみてください。サンプル音声が、4ケ国語で聴けますが、ほとんど肉声と同じ雰囲気なのには驚嘆します（http://voice.pentax.co.jp/）。

二つ目は、プレーヤーとエディターを購入すれば、自作原稿やテキストファイルなどの文章を読み上げさせることが可能にするというものです。

青空文庫のサイトでの広告では、「青空文庫が聴ける」というPRコピーを流しています。無料プレーヤーソフトもありますが、これは5万字まで。有料プレーヤーでは、そのような制限はありませんから、青空文庫に収録されているような長い作品のテキストファイルを読み

上げさせることができるわけです。

三つ目は、有料プレーヤーソフトを使うと、音声ファイルを出力し、私的利用の範囲内で使えるとのこと。料金は、1年間で3150円です。

なお、その語りべプレーヤーを使って自作原稿を読み上げた音声ファイルを、「電子かたり部.com」のサイトで販売受託サービスも行なっています。つまり、自分が書いた作品のオーディオブックの販売をしてもらえるというわけです。

朗読音声版「初音ミク」の出現に期待する！

「初音ミク」というのは、上記の「電子かたりべ」の音楽の萌え～版のようなものでしょうか。ずいぶん人気があるそうです。メロディと歌詞を入力すると、その曲を女性の声の合成音で歌う16歳のバーチャルアイドル「初音ミク」が歌うという設定で、声優の藤田咲さんの声を元に作られた合成音で再生する由。

人間の声を元にリアルな歌声の合成音を作ることができるヤマハの技術「VOCALOID 2」を活用した音声製作ソフトだそうです。このボーカロイドの価格は、15000円ですから、十分一般に手が届きます。

※　試聴は、http://blog.avenew.jp/archives/76

http://www.crypton.co.jp/mp/pages/prod/vocaloid/cv01.jsp

などなど、作品によって、気分によって、朗読してくれる声を替えることができたら、語りべプレーヤーがもたらす世界は、もっと豊かなものになることでしょう。

そして、現在、いろいろな方の朗読ブログがあり、それぞれ特徴のある、ユニークで魅力的な声調で朗読しておられます。それぞれに、根強いファンがいることでしょう。できれば、それらの「魅力的な声」を組み込んだプレーヤーソフトを、販売なり仲介なりしてくれるサービスが提供されることを期待したくなりました。つまり、

以上、3つの事象をご紹介しましたが、まとめると、以下のように分類されます。

「自分なり肉親などの声そっくりの合成音声で」、各種原稿や作品を読み上げるソフト

「既製の合成音声で」、各種原稿や作品を読み上げるソフト

「魅力的な人の声を元に作った「魅力的な合成音声」で、自分の作品を歌わせるソフト

「魅力的な人の声を元に作った各種の合成音声で、文藝作品を読み上げるソフト」

です。今の「語りべプレーヤー」は、それはそれで画期的なものだと思います。ただ、できることなら、その声は、既成の1種類のものだけではなくて、さまざまな魅力的な音声のプレーヤーを用意してもらえないものだろうか、ということなのです。

女性の声、男性の声

高い声、低い声

お爺さん・おばあさんの声、若い人の声、熟年の声

落ち着いた声、元気な声

ビスができないものだろうか、ということでもあります。

今の「語りべ.com」のサービスは、既成プレーヤーを使って自作の作品を読み上げた音声ファイルの販売仲介サービスですが、「自分の声自体を提供したい」という朗読者の方もおられるのではないかと思いますし、逆に、「あの人の素晴らしい声で、いろいろな作品を聴いてみたい！」というファンもいるのではないかと思います。私もその一人です。それをつなぐ仲介サービスがあればいいと思うのですが、どんなものでしょうか・・・。

もちろん肉声にはかないませんが、それでも、さまざまな作品をオーディオブックの形で、聴く読書をしたい！　というニーズに一定の範囲で応えることができるものと考えます。

音声認識ソフトでも、かつては何千万円もしたものが、1万円前後で買えるほどにまで普及しています。そして、高性能の音声入力ソフトも同様の価格で買えるようになっています。自分の声で文章を多数読み上げて記憶させ、そのパターンから文字で出力するというものです。これができるのであれば、音声合成・出力ソフト（＝プレーヤー）まで、そう時間はかからないような気もします。

ZHK が開発した驚異的な音声合成システム

NHKが数年前に開発した合成音声システムについてを、初めて聴いたときは衝撃でした。こんな滑らかで自然で、本人同然の声が合成できるものだろうか?!　と呆然とする思いでした。

百聞は一聴に如かずで、試聴サイトから聴いてみてください。NHK技術研究所（世田谷区）のホームページの研究成果紹介のサイトから聴くことができます。そこでは、以前ニュースでおなじみだった森田美由紀アナウンサーの声から合成した声で、ニュース原稿が読み上げられています。まるでこれは、森田アナウンサーの声そのもので、合成音声だといわれなければ、区別できないほどのすごい出来です。

※ 試聴　http://www.nhk.or.jp/strl/group/ningen/oho/ningen/oho

これを読むと、容量が大きくなりすぎるという難点はあるようですが、今後さらに改善されていくことでしょう。これが、一般に普及するような価格帯で提供されるようになれば、音声の世界はがらりと景色が変わって来るような気がします。ZHK の底力に期待したいと思います。

（2）音声合成技術と音訳

2倍速、3倍速の合成音声プレーヤー、ソフトの出現

「音訳」の究極の目的は、聴き手に「文字情報を正確に伝える」ことにあり、あとは聴き手の解釈に任せる、ということなのだろう、と思います。

そのような「文字情報を正確に伝える」ということから言えば、これまで書いたように、技術進歩が著しい「合成音声ソフト」もかなりのものです。お試しコーナーがいろいろあるので、試しにいろいろな文章をいれて聴いてみると、なかなかどうして、駅や電車のアナウンス並みの自然さは出ていますし、男女の声、若い声と、声の種類もいくつか出てきています。最近は、ボルックスターというソフトのように、特定の人間の声とほぼ同じ声で、さまざまな文章を聴くことができるようになっています。一番レベルの高いものだと、文章によっては、ですが、ほとんど肉声と区別できないものもあります。4～5年前とは隔世の感があります（ピアノでいうと、ヤマハなどの自動演奏ピアノのようなものが、音声読み上げの世界でも出現しつつある一ま

だまだ完成の域には達してはいませんが」ということですね。）。

「電子かたりべ.COM」(http://www.e-kataribe.com/portal/index.jsp)のサイトをみると、

合成音声ソフトを使った音訳作品が、続々と売り出され、ジャンルも、文芸作品はもちろん、実用書、法律書、趣味の本等々、あらゆる分野にわたっています。そして、無料でダウンロードできる「かたりべプレーヤー」も、倍速モードが選択できるようになっていますし、ゆっくりにもできるようになっています。

また、「PeBe」というオーディオブックを多数出しているサイトに、「皆様のリクエストにお答えして、遂にオーディオブックの【倍速版】が登場しました！」というPRが載っていました。「皆様のリクエストにお応えして」「遂に登場」・・・ということは、倍速版の需要が高い様子が伺えます。実用書、啓発本などは、早く読んで理解するにはそれがいいのでしょう。文芸作品などでも、味わうというよりは、筋を追うという使い方なのでしょうか・・・。

合成音声プレーヤーと音訳との競合

こういう状況になってくると、「音訳」との差が、微妙になってくるような気がします。

「音訳」は、「勝手な解釈が入ることなく、文字情報を聴き手に正確に伝える」ということだといいます。「音訳ボランティア全国ネットワーク」のサイトでは、次のようにあります。

朗読が「自己表現の手段」であり「主として文学を」「感情豊かに読む」ことに対して、音訳は「音訳者の主観を入れることなく」「書いてあることをできるだけ忠実に音声化する」ことが原則です。利用者の求めに応じてあらゆる分野の情報（図・表・写真なども含む）を伝えることが求められます。誤読が許されないのはもちろん、音訳者の勝手な解釈で原本を読み替えたり余分なことを付け加えたりすることはできません。

「上手に読む」ことではなく「正しく伝える」「正しく伝わる」ことが大事なのです。

もし、そうだとすると、

「文字情報を聴き手に正確に伝える」という「音訳」の目的は、合成音声技術によって達せられつつあるのではないか？　代替可能なのではないか？

スピードの調整も自在だから、速く聴きたい人にもゆっくり聴きたいひとにも合わせることができるから、そのほうが便利ではないか？（音訳テープを2倍、3倍速にして聴かれる方もおられるそうなので、そうであれば、なおのことでしょう。）

「勝手な解釈が入ることはがあってはならない」という要請は、より忠実に果たされるのではないか？

聴き手が好きなタイプの声で自由に選択できるのだからそのほうがいいのではないか？

・・・といったことも言おうと思えば言えないこともないかもしれません。

しかし、それに対しては、「そんなことはない！」と反発する方も多いことでしょう。「肉声ならではの読み方があるはずだ」、「感情を持つ人間ならではの読み方があるはずだ」「無機的な合成音声など、人間の声

にかなうはずがない！」等々・・・。

私もそう思います。ただ、「音訳」の世界で、「感情抜きで淡々と」「情報を伝えるのが最大の目的」ということに過剰にこだわりすぎることがあるとすると、上記のような指摘に反論しきれなくなるような気が一方ではします。

※　ただ、「音訳」と「朗読」の違いというのは、人によっても捉え方が異なるようです。「音訳」と称している『声の花束』に収録されているさまざまな作品をお聴きすると、それぞれの読み手の個性が存分に出ていると感じます。他方で、いわゆる「棒読み」に近い感じで速く読んでいるものもあります。誰もが一致する「模範的な音訳」というものがあるのか、よくわかりません。

合成音声システムによって負担軽減、役割分担も

音訳サークルと朗読サークルと双方に関わっておられる方から、「『ボランティアならそれでよくても・・』とか『押し付け』だとかそれぞれの側から言われ放題」という嘆きを漏らされて、「今ちょっと自分の力を信じられない状況で落ち込んでしまう」という状態に陥っているとのお話をと伺うと、音訳のほうは、合成音声を活用して負担軽減を図ることも選択肢としてはあるような気がします。いくら文字情報を正確に伝えることが役割だとしても、それを2倍速、3倍速で聴かれるとすると、何か割り切れ気がするのですが・・・。

もちろん、長年、この世界で取り組んできた皆さんにとっては、「音訳」か「朗読」かは、アイデンティティ（自己規定）に関わる話で、部外者が軽々に言及する話ではないのかもしれません。ただ、朗読鑑賞ファンとしての立場から言えば、もちろん感情過多は抵抗がありますが、それぞれの読み手の個性と感情とで自然に朗読していただき、そういう方々が多数朗読を提供していただける、ということが、朗読文化の厚みをもたらすのではないかと思います。

もし、「個性や感情の押しつけを排する」というところで、悶々としながら埋もれている皆さんが多数いるのだとすれば、それは、大げさに言えば、「国民的損失である！」ともいえると思います。