

大学業務における合成音声活用方法の検証

長倉宏樹

北海道大学大学院工学研究院 工学系技術センター技術部

1. はじめに

工学部の放射線施設である「瞬間強力パルス状放射線発生装置研究室」は、利用前に年1回開催している施設主催の教育訓練(ライナック講習会)の受講もしくはその動画の視聴が必要となっている。これまでは教員と技術職員が講習会を行ない、それを録画して動画を作成していたが、昨今の情勢により放射線の規則変更に伴う施設利用方法変更が多々行われることから、合成音声を利用した講習会動画を予め作成することで業務の省力化を行なえるか検証を行なった。また北大技術研究会 2023 で発表したアナウンス用デバイス作製について少々改良を行なったのでこれらをまとめ、合成音声の大学業務へ活用方法の検証として報告する。

2. 合成音声について

本発表での「合成音声」とは、テキストを読み上げる音声を機械的に作り出す技術(音声合成技術)によって作られた音声とする。音声合成の仕組みの一例は以下に示した通りで、PC に入力した文字をソフトウェアが解析し、音律を生成、そして音声を生成することで合成音声を作成している。身近な例を挙げると、スマートフォンの音声アシスタント、駅・空港の車内放送や案内放送などで使われていることが多いが、近年では個人の趣味で気軽に合成音声を利用できるソフトウェアが販売されているほか、手軽に商用利用できるもの、外国語に対応しているものなど様々な種類があり、北海道大学でも一部の研修や講習会用の動画で利用されている。本検証では、合成音声の作成に株式会社 AHS が発売している文字読み上げソフトウェア VOICEPEAK を使用した。



図1 合成音声の仕組みと身近な例

3. 検証内容とその結果

ライナック講習会の動画作成は、今までの講習会形式からあまり変えないよう **PowerPoint** を使ったスライド形式に合成音声を組み合わせて行なうこととした。作成の流れとしてはスライド作成後、**VOICEPEAK** にて発表原稿の内容を合成音声化したナレーションを作り、アニメーションとしてスライドに組み込むことで、説明している箇所を矢印や線などで示すと同時にナレーションが流れる形式として、**PowerPoint** の録画機能を利用して動画を作成した。ナレーションに合成音声を利用することで、撮影時に係る発表者の負担(作業手順の暗記、雑音などへの注意、声の聴きやすさへの配慮等)を軽減できた。合成音声の声質も技術の進歩により実際に人が話している声と区別出来ないものとなっているため、調整の手間もかからずに筆者が喋るよりも視聴者の集中力が欠ける要素を減らした動画を作成できたと思う。また、利用しているソフトウェアは **PowerPoint** と **VOICEPEAK** のみなので、後に規則の変更があった場合でも手軽に動画編集を行なえる点も大きいように感じた。これにより派遣先業務の一つである講習会に係る作業が省力化できたと言える。

アナウンス用デバイスの作製では、前回作製した「**Arduino** と **DFPlayer mini**(+microSD カード)、人感センサーを用いて、施設に入室した際に利用案内の音声を再生するデバイス」について多少の更新を行なった。主な変更点は音声再生のトリガーとなる人感センサーをマイクロスイッチに変更して、誤作動を少なくすると同時にコードを単純化することなどである。このデバイスは、施設職員が正面玄関前の管理室を不在にしている際に、講義室利用等で施設を訪れた来訪者が誤って放射線管理区域に入らないように設置することを目的に作製しているが、**2024 年度 4 月**より開始した施設改修工事の影響で検証を行なえていないため本発表ではテスト運転までの状況について発表する。



図2 施設主催教育訓練動画の作成及びアナウンス用デバイス作製

4. 今後の展望

動画作成の方に関しては合成音声ソフトウェアで様々な声質を選択できることを活かして、学習・技術習得を分かりやすく行えるようなバリエーションの動画を試す、そして近年増加している海外からの施設利用者のために英語ナレーションの動画の作成などを検討している。アナウンス用のデバイスにおいては、改修工事終了後に施設内で実証テストが行なえることを目指して改善を続ける予定。また合成音声の大学業務におけるその他の活用方法として、最近製作に参加している加速器 VR の案内音声などへの利用や、大規模言語モデル (LLM) と組み合わせた業務内容省略化等を検討していきたい。