

厚生労働科学研究費補助金（がん対策総合研究事業）
分担研究報告書

合成音声を用いた音声資料作成手法の確立に関する研究

研究協力者 志賀 久美子 国立がん研究センターがん対策情報センターがん情報提供部 看護師
研究協力者 高橋 三智世 堺市立健康福祉プラザ視覚・聴覚障害者センター
研究代表者 八巻 知香子 国立がん研究センターがん対策情報センターがん情報提供部 室長
研究協力者 三村 麻子 国立がん研究センターがん対策情報センターがん情報提供部 研究員

研究要旨

国立がん研究センターがん対策情報センターでは、聴覚障害者向けにがんの冊子の音声図書を作成している。しかし、肉声による作成のプロセスは手間がかかり、公開までに非常に時間がかかっていた。そこで、テキスト DAISY (Digital Accessible Information System) から合成音声を取り出すことで、迅速に音声図書の作成を行うことが可能ではないか、と考え、その方法を確立した。今後、病院や相談支援センター等で音声資料を作成する際に活用されれば幸甚である。

A. 研究目的

国立がん研究センターではがん対策情報センターが発行するがんの冊子シリーズについて、国立がん研究センターと堺市立健康福祉プラザとの協定に基づき、音訳・点訳資料を作成し、インターネット上の点字図書館である「サピエ」および「がん情報サービス」の両方に公開してきた。がん情報サービスでは、視覚障害者向けサービスを知らない人でも利用可能とするため、mp3版もあわせて公開している。しかし、肉声による録音、図表の解説を含む作成プロセスには、何度も人の手を渡る工程となるため、完成・公開までに時間がかかり、がん情報サービス上で公開している音声版が軒並み旧版となる事態が生じていた。世界的には合成音声の利用は広く進んでおり、日本語の合成音声の質も向上し、実用に耐えるレベルになってきている。

そこでテキスト DAISY(Digital Accessible Information System) から、合成音声を取り出すことで、文字情報のみに限られるが迅速に情報を提供できると考え、その工程を確立することで、今後最小限の労力で迅速な情報提供につなげることを目的とした。

B. 研究方法

堺市立健康福祉プラザ視覚聴覚障害者センターの職員にどのような手順や方法でテキスト DAISY を作成しているかのヒアリングを行い、その後も月一回のミーティングを持ち改善を重ねた。

C. 研究結果

1. 資料作成プロセスの確立

作成のプロセスについては、堺市立健康福祉プラザ視覚聴覚障害者センターのボランティアが従来行っている方法で原則行ったが、医療の専門用語や特有の表現に間違いがないかを確認する必要があった。そのため、国立がん研究センターがん対策情報センターが医療用語の読み方の校正を行う工程を追加し、以下のような手順で作成することになった。

まず、元のがんの冊子からテキスト DAISY (Digital Accessible Information System) を作成し合成音声自動的にテキストを読み上げる機能を利用し音声抽出した。その際には、合成音声

が読みあげられるよう、医療専門用語にルビをふった（以上堺市立健康福祉プラザ視覚聴覚障害者センターボランティアが行った）。その後、国立がん研究センターの医療資格を持つものが合成音声と元の冊子を照らし合わせながら、合成音声の正しい読みができていないかを確認した（以上を国立がん研究センターがん対策情報センターが行った）。その後、公開用のテキスト DAISY と mp3 の作成を行い、がん情報サービスやサピエ図書館に公開した。冊子に含まれている図表については自動化できないため、肉声による「完全版」作成時に反映することとした。

これらの情報はすべて公開情報の形態変換であり、個人情報等は一切含まないことから、多くのボランティアと資料共有が簡易にできる方法として、ファイルのやり取りについては、dropbox を利用した。堺市立健康福祉プラザ視覚聴覚障害者センターのボランティアと国立がん研究センターの間の連絡のやり取りは ML を作成し、複数の人が関わっていても、全体で進捗を共有できるようにした。

2. 作成した音声版資料

1の工程により作成したものについては以下のものがある。

まず、「がん患者や家族、周りの人へ：新型コロナウイルス感染症の感染拡大に際して（がん情報サービス一般向け資料、2020年4月10日公開）」と「相談員用、患者支援者用新型コロナウイルス Q&A（2020年4月24日公開）」の2点を2020年5月に作成した。新型コロナウイルスの感染拡大に際し、迅速に資料の作成が必要であったため、即刻作成を開始し、両資料について2020年5月1日には公開することができた。その後のがん情報サービスの情報更新についても、数日以内に更新を行うことができた。

また、がんの冊子については、2020年10月から12月の2か月間で16種類、がん情報サービスホームページ内の「さまざまな症状への対応」については2021年1月から2月の2か月間で7種の合成音声による音声資料を作成することができた。

D. 考察

今回、手順を確立し可視化したことで、作成に携わるスタッフが音声資料作成プロセスの過程で迷う

部分が減り、最小限の労力で流れに乗って作成をすることが可能となり多くの最新資料が更新された。作成に携わるスタッフが、様々な事情で入れ替わることも多く、担当していた者以外でも作成手順を参照すれば作成が可能となり、特定の人がいないと作成ができないという不安定な状況を回避することができる。また、このシステムを利用し作成したことで、今後情報が更新されたり一部が修正されても、その部分だけを修正すればよく、改変も迅速に行うことが可能であると考えられる。障害のあるがん患者にとって、医療情報を得ることはその後の治療や生活を意思決定する命綱であり、その情報が最新のものであることは非常に重要なことであると考ええる。

また、このノウハウを利用すれば病院内の資料(入院案内、説明文書等)の音声資料作成が簡便にできる可能性があり、このプロセスを共有することで応用が可能であると考ええる。

E. 結論

医療情報は日進月歩で更新されており、それに追随して音声図書を肉声で作成するには非常に手間と時間がかかっていたが、音声資料作成手法を確立し

たことでスムーズに最新のがん情報を提供することが可能となった。今後は、この作成プロセスを広く公開することで、様々な音声資料の作成に寄与できると考える。

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表

該当なし

2. 学会発表

該当なし

H. 知的財産権の出願・取得状況

1. 特許取得

該当なし

2. 実用新案登録

該当なし

3. その他

該当なし

(参考)

このプロセスにより作成した情報をごん情報サービスで公開することについて、国立がん研究センター令和2年度第8回がん情報編集委員会内部連絡会(2020年10月7日)にて承認された。

参考1) がんの冊子簡易版更新(16種)

冊子名	簡易版HP掲載日	完全版HP掲載日
子宮頸がん	2020/10/16	2021/03/19
子宮体がん	2020/10/16	2020/12/03
食道がん	2020/11/02	未
大腸がん	2020/11/02	2020/12/15
肺がん	2020/11/08	2021/03/16
胃がん	2020/11/08	2021/03/19
乳がん	2020/11/21	未
小児の横紋筋肉腫	2020/11/30	未
脳腫瘍	2020/12/09	2020/12/18
外陰がん	2020/12/02	未
緩和ケア	2020/12/11	2021/02/22
膵臓がん	2020/12/11	未
胆道がん	2020/12/11	未
肝細胞がん	2020/12/11	未
小児の神経芽腫	2020/12/22	2021/03/16
小児の脳腫瘍	2020/12/22	2021/03/02

参考2) 「さまざまな症状への対応」2021/02/22掲載(7種)

しびれ
脱毛
口内炎
発熱
ほてり・のぼせ・発汗しやすい
発熱性好中球減少
出血しやすい・血小板減少

参考3) 音声図書2019年以前の更新履歴

2014年	掲載開始18種
2015年	8種
2016年	2種
2017年	更新なし
2018年	更新なし
2019年	8種