

印刷業界の新技术情報を三美印刷がお届けするメールニュース

sanbi-i-com (No.191)

オーディオブック (4)

可変か固定か

今回は、「人間か機械か」の他に「可変か固定か」という視点を用いてオーディオブックの制作手順を図にまとめます。また、オーディオブックの隣接分野である音楽と書籍についても同じ見方で整理してみます。

1. オーディオブック

オーディオブックの基となるテキストデータは、「可変か固定か」で言えば「可変」です。どんな声でどう読ませることも可能だからです。一方、読み上げて録音した後のオーディオデータは、声色も読み方も確定しているので「固定」です。

「可変」は「未加工で用途が広い」「汎用」、「固定」は「加工後で用途が定まっている」「専用」と言い換えることもできます。いわば原料と製品の関係です。

「人間か機械か」については言うまでもなく、朗読は「人間」、音声合成で読むのは「機械」です。以上を図にまとめます。

＜軸の取り方＞ ・横軸：「可変か固定か」、左が可変、右が固定
・縦軸：「人間か機械か」、上が人間、下が機械



以下の二例のように読んでも録音・固定化しない場合は、オーディオブックにはなりません。

- ・(人間の例) 大人が幼児に絵本を読み聞かせる
- ・(機械の例) 電子書籍内のテキストを読者が手元の読み上げソフトに読ませる

なお、音声合成には、人間の声の波形を使うもの(A)と、声を使わずに波形から合成する(B)があります。(A)は、声を使うとはいっても人間による朗読ではないので、これもやはり音声合成の一種です。近年著しい性能向上を見せているのが、この(A)で AI(人工知能、機械学習)を使うタイプです。

続いては、この縦軸、横軸を使って、音楽と書籍(目で読むタイプ)を見てまいります。

2. 音楽

オーディオブックと音楽はどちらも「音を聞いてもらう」ものであり、オーディオデータ(MP3 や WAV 等)に固定するという共通点があります。

●参考：自動演奏がなかった時代

譜面(演奏者の解釈によってテンポ、強弱、音色等に変わる余地のある可変情報)を見て人間が演奏し、録音して固定化する一本の道筋しかありません。



●現代

MIDI データ+音源による自動演奏(いわゆる打ち込み)で人間の演奏と区別がつかないくらい高品質なものも作れるようになり、次頁の図のように人間の演奏と自動演奏を選択したり、混ぜ合わせたりしてオーディオデータを作るようになりました。オーディオブックの図とよく似た、複数の道筋がある構図です。適用する音源に、(A)サンプリング録音した音を使うタイプと、(B)波形から合成する(物理モデルと呼ばれる)タイプの両方がある点も似ています。



※MIDI データは、音の高さ、強さ、長さ、タイミング、その

他を記述するデータですが、音源(楽器)がなければ音(演奏)になりません。

※譜面は必須ではありません。左図の MIDI から右側だけで完結するケースも多々あります。

※通信カラオケのように、オーディオデータ化せず、MIDI データで配信して現地の音源で鳴らす例もあります。

3. 書籍(紙または電子)

オーディオブックと書籍(紙または電子)は、写真やグラフ等を度外視すれば、どちらも「言葉を伝える」ものであり、テキストデータが基になるという共通点があります。テキストデータはフォントなどの書式情報を含まない「文字のみ」の情報ですが、オーディオブックはこれを聴覚化し、書籍は視覚化します。

を右の固定側に並べてみます。



●参考：文字、書籍がなかった時代

言葉は音で口伝するしかありません。いわばオーディオブックの源流です。書籍がまだないのでオーディオブックの方が歴史は古いという言い方もできそうです。ちなみに、文字がある時代になっても口伝は残ります。琵琶法師が音楽とともに口述で平家物語を伝えたのが有名な例です。



●参考：写本が主流だった時代

文字が発明され、書籍が誕生した後でも、印刷が利用できなかつた頃は、手書きで写すしかありません。



●現代

一足飛びにDTPが普及した現代に進みます。写本は極めて少ないと見て無視します。

音楽の自動演奏で音源がなければ音が鳴らないように、テキストデータの文字はフォントがなければ形になりません。音源とフォントには聴覚と視覚という大きな違いこそありますが、可変のものを固定化することでよく似た役割を担っています。そこで下図では、フォントが未確定のものを左の可変側に、確定したものを

<上記①～④の補足>

・①のテキストデータは、このままでは紙書籍にはなりませんが、電子書籍には成り得ます。例えば、青空文庫はテキスト形式です。どんなフォントで表示するかは読者にお任せ(手持ちのフォント次第)となります。

・②の DTP ソフト等でフォントを指定しますが、データを受け渡した場合、渡された相手側にも同じフォントがないと渡した側の意図通りに表示されないのが、フォントが確定したとはまだ言い難いです。

・③のフォント埋込みの PDF にすれば確定します。テキストとフォントデータがセットになっているので、データを受け渡しても確実に意図通りに表示されます。ここまで(①～③)はテキストが残っているので読み上げも可能です。

・さらに字形の確定度が高いのが④のアウトライン化です。フォントを変更しようにも、もはやテキスト+フォントではなくなっているので変えようがありません。読み上げはできなくなります。

以上、オーディオブックとその隣接分野(音楽と書籍)の制作手順を「可変か固定か」の視点で図示、整理してみました。何かのご参考になれば幸いです。

以上

(第 191 回: 2022 年 5 月 26 日)