

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.190)

オーディオブック (3)

音声合成に関する補足

現在市販されているオーディオブックのほとんどは人間が朗読したものです。音声合成で読み上げたものは今のところ多くはありませんが、今後は増えてくる可能性があります。

音声合成については前々回の sanbi i-com No.188 で簡単に触れましたが、今回はその補足です。

1. タイトル数を増やす鍵は、やはり音声合成

オーディオブックには、目で読む書籍(紙と電子)と比べて不利な点が少なくとも二つあります。

(1) タイトル数が少ない

(2) 制作に時間とコストがかかる

人間が読むか、機械で読む(音声合成する)かに関わらず、オーディオブックには読み上げるための元の本のテキストが必要です(※)。このため、オーディオブックのタイトル数は、目で読む書籍に比べてどうしても少なくなってしまう。

※例外として、講演の録音をオーディオブックのみで販売(元の本なし)というケースがあり得ますが、出版市場全体の膨大なタイトル数から見れば微々たる数でしょう。

タイトルが多いか少ないかの感じ方は人によりけりとはいえ、少ないと思われたままだと下表のような悪循環が定着しかねません。

表中の略語: D=需要側、S=供給側、AB=オーディオブック

①	タイトルが少ない	←
②	D: 聴きたい AB がない →AB に魅力を感じない。興味を持たない。	
③	AB 市場が伸びない	
④	S: <u>市場が小さい</u> ので、 <u>コストをかけて</u> タイトルを投入しても、売れて儲かる見込みが持てない	
⑤	タイトルが投入されず、増えない →①へ戻る	→

④でタイトル投入を阻害している要因を、青字の

「市場が小さいから」と見てしまうと、鶏が先か卵が先かのような堂々巡りの話(タイトルが少ないから市場が小さい⇔市場が小さいからタイトルが少ない)になってしまいます。そうではなく、赤字の「コストをかけて」の方に着目すれば、「コストがかかるから損益分岐点が高くなり、投入できるタイトルに限られる」という単純な話になります。④でかかるコストを抑えられるならば、損益分岐点が下がり、投入できるタイトルを増やすことができます。

人間が読むオーディオブックの制作にかかるコストや時間の例として [AI Comes to Audiobooks](#) という記事で見つけた米国での数字をご紹介します。

●ナレーターへの支払い

有名タレントを起用した場合の相場は、朗読1時間当たり\$1000以上。1冊の朗読時間の平均は8時間なので、1冊平均 \$8000以上となります。

●録音スタジオ

朗読時間の倍以上のスタジオ時間が必要。8時間の本なら、16時間以上が必要となります。

音声合成ならば、音声合成ソフト/サービスの利用コストがかかるとはいえ、ナレーター代もスタジオ代も不要になるので、総じて相当の時間/コストの抑制が可能になるでしょう。タイトル数を増やす、ひいては今後オーディオブック市場を成長させる鍵は、やはり音声合成にあると言えます。

2. Audible は現状、人間読みのみだが・・・

AI を使った音声合成技術の近年のめざましい性能向上を受けて、オーディオブック先進国である米国ならば機械読みのオーディオブックもかなり出回っているのではないかと思いきや、「オーディオブックといえば人間読みである」は、今も基本的には変わっていないようです。

その一番の理由は、米国のオーディオブック販売で50%のシェアを持つと言われる最大手、Amazon の Audible が人間読みに制限しているからだと言われています。

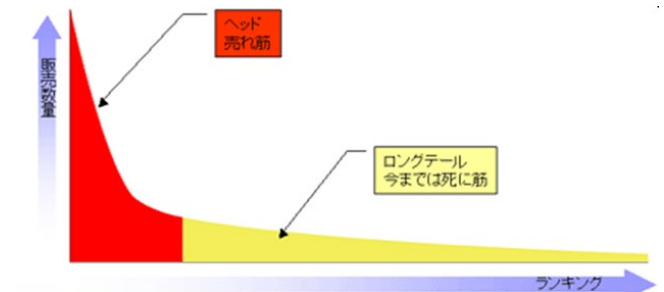
Audible によれば「読者(聴者)はストーリーだけではなくパフォーマンス(誰がどう読むか)で選んで購入している。その期待に応えるには人間読みが必須」とのことです。Audible 以外の多くのオーディオブック販売所にはこのような制限はないにも関わらず、一強とも言える最大手のルールなので、市場全体への影響力が大きいと見られます。

Audible が今後もこの制限を続けるのかについては、先述の [AI Comes to Audiobooks](#) の記事の中に「制限が取り払われるのは時間の問題」との見方が紹介されています。

筆者も同感です。Amazon と言えば、ロングテール戦略で成功している企業の代表格です。人間読みへの制限は、オーディオブック制作のハードルを高くしてしまい、タイトル投入をあきらめさせるケースを増やしてしまいます。

これではロングテールの切り捨てであり、Amazon

の基本に反するので、いずれ制限は取り払われるだろうと予想します。



また、同記事を読み進みますと、Speechkit という TTS (Text to Speech)の会社に、Audible が一社を間に挟んで融資…云々と書かれています。これは Audible 自身が音声合成の可能性を有望視している現れです。

なお、日本市場の場合は、Audible 一強ではなく、audiobook.jp と Audible が二強と呼ばれており、少なくとも audiobook.jp の方は既に機械読みのオーディオブックも販売しています。市場規模では日米で大きな差がついていますが、二強の一角が既に機械読み OK であることから、音声合成の活用は、案外日本の方が早く進んでいくかもしれません。

以上
(第 190 回: 2022 年 4 月 27 日)