

The Development History and Current Challenges of Japanese Audiobooks: A Postprint Analysis

Authors: 郑文全

Date: 2023-10-08T00:00:00+00:00

Abstract

Over its more than 80-year development history, Japanese audiobooks have evolved two distinct models: public welfare initiatives and commercial publishing. While audiobooks currently enjoy robust development trends in both the United States and China, with market shares increasing year by year, they have never achieved significant growth in Japan. This situation stems from both broader societal contextual factors and individual reading habits. Examining the current state of Japanese audiobook development and reflecting on the influencing factors behind its successes and failures can provide valuable insights for domestic publishing houses seeking to enter the audiobook market.

Full Text

1. The Development History of Japanese Audiobooks

Japanese audiobooks can be divided into two distinct categories. The first category carries a public welfare nature, primarily serving individuals with visual impairments and the elderly through services like library programs and specialized publications. The origins of audiobook publishing in Japan can be traced back to 1939, when Helen Keller donated “reading machines” (phonographs and record players) during her 1937 visit to Japan. This initiative marked the beginning of library-based audiobook services for the visually impaired, with the Kyoto Prefectural Hibiya Library pioneering recorded book services. In 2008, the Japan Library Association’s Committee on Services for Persons with Disabilities held seminars to promote multimedia DAISY (Digital Accessible Information System) books, actively advancing a standardized and formalized DAISY production system. By 2010, 123 of Japan’s 2,272 public libraries were offering audiobook services, with approximately 20,000 DAISY book titles available [1].

The second category represents a purely commercial model, known in Japanese as “オーディオブック” (audiobooks)—electronic books featuring audio recordings produced by narrators or voice actors, sometimes utilizing voice synthesis software. As early as the 1980s, approximately 100 Japanese publishers, including Shinchosha, attempted to launch CD-based audiobook products for the general public, emulating the American model, but these efforts failed to gain traction. The modern commercial audiobook market began in earnest when OTOBANK launched Japan’s first online store in 2007. Consumer adoption accelerated after 2011 with the proliferation of smartphones, and the market received a significant boost when Amazon’s Audible entered Japan in 2015. Major publishers such as Shinchosha, Iwanami Shoten, PHP Institute, and Kadokawa have since entered the market, with companies like ASK releasing language-learning series such as the *Graded Japanese Extensive Reading Library* in both computer and mobile audio formats, sold through major retailers like Amazon and Rakuten.

2. Current Status and Representative Cases in Japanese Audiobook Publishing

OTOBANK stands as Japan’s most representative audiobook company. Its general manager, Hiroshi Kubota, declared 2020 the “first year of Japanese audiobooks” after the company’s membership surpassed one million, growing from just 600,000 members in December 2018 to 1.2 million by December 2019. According to projections by the Japan Management Association Research Institute, the Japanese audiobook market was valued at 3.2 billion yen in 2018 and is expected to reach 26 billion yen by 2024 [2]. The website with the largest audiobook catalog in Japan is audiobook.jp, operated by OTOBANK, offering approximately 27,000 titles. The platform’s catalog spans business, self-improvement, general education, foreign languages, practical guides and qualification exams, literature and rakugo, anime and radio dramas, speeches, alternative romance, and adult content. It features both comprehensive and limited-time free content rankings. The top three titles in the comprehensive ranking are *Think Like a Monk: The Art of Focus*, *The Mindset of Successful People (Complete Edition)*, and *Personal Globalization: Daily Habits and Five Tools You Need to Know to Work Globally (Revised Edition)*—all reflecting the strong consumer demand for personal development content in an era where individuals increasingly require global perspectives.

The primary consumer demographic for Japanese audiobooks consists of people in their thirties and forties. In major metropolitan areas, most listeners consume content during their daily train commutes, while in other regions, consumption predominantly occurs while driving. Additionally, many young and elderly consumers have become audiobook users due to their admiration for specific voice actors or preference for particular content genres. The entry of internet companies into the audiobook market has significantly impacted Japan’s publishing industry, with publishers expressing concern about potential monopolization by Amazon. In April 2015, sixteen publishers established the “Japan Audiobook

Association” with the stated goals of “(1) understanding and analyzing the market, (2) maintaining prosperity alongside print and e-book industries, and (3) protecting the rights of copyright holders” [3]. However, these publishers’ institutional frameworks have failed to keep pace with market developments, resulting in insufficient vitality.

3. Real-World Dilemmas Facing Japanese Audiobooks

3.1 The Struggle for Development Dominance Between Publishers and Internet Companies

In traditional publishing, publishers served as the central players, but the digital publishing era has transformed the landscape into a more open ecosystem encompassing copyright holders and consumer-authors, with accessible pathways for any party to participate. Amazon’s successful entry into audiobook sales from its position as a book retailer exemplifies this shift. The question of whether publishers or internet companies should lead audiobook development remains unresolved: publishers control copyrights, while internet companies possess technological capabilities, making integration challenging. Internet companies maintain an optimistic outlook on audiobook development, with OTOBANK’s general manager expressing confidence that audiobooks will continue growing as a knowledge-based entertainment product within the internet economy. Conversely, publishers remain largely passive, with many skeptical about audiobooks’ potential as a profit pillar. For instance, Yukio Tomikura, head of digital sales at Kodansha, has openly questioned whether audiobooks can become a revenue mainstay for publishers. Although Kodansha launched audiobook production in 2016, it adopts a “non-licensing” model to mitigate risk—utilizing internet companies’ technology for production while independently handling sales and profit management. The user wants me to translate a Chinese academic text about audiobooks and publishing into fluent English. The critical requirements are:

1. Preserve ALL math exactly as given (there’s no math in this section)
2. Preserve ALL citations exactly (there are none in this section)
3. Preserve ALL figure/table markers exactly (there are none in this section)
4. Maintain academic tone, precision, and technical accuracy

The translation guidelines are: 1. Produce flowing, readable prose 2. Merge fragmented text into complete sentences/paragraphs 3. Skip garbage text (there doesn’t appear to be any) 4. Organize by logical sections (the section “3.2 Legal and Institutional Obstacles” is clearly marked) 5. Each paragraph should express a complete thought (3-8 sentences) 6. Use clear topic sentences and logical transitions

The output format should be clean Markdown with ## for major sections and ### for subsections.

The user also provided some specific text to translate that appears to be a mix

of English and Chinese fragments. Let me analyze it:

The initial part seems to be instructions or notes from the user about translation principles. Then there's a section that appears to be partially translated text with some Chinese characters “介, 入” which the user says must be eliminated.

Looking at the actual content to translate:

“In traditional publishing, publishers served as the central players, but the digital publishing era has transformed the landscape into a more open ecosystem encompassing copyright holders and consumer-authors, with accessible pathways for any party to 介入.”

The Chinese character “介入” means “intervene” or “get involved”. The user suggests translating it as “get involved” or “participate”.

“Amazon's successful entry into audiobook sales from its position as a book retailer exemplifies this shift.”

This is already in English and looks good.

“The question of whether publishers or internet companies should lead audiobook development remains unresolved: publishers control copyrights, while internet companies possess technological capabilities, making integration challenging.”

Also already in English, looks good.

“Internet companies maintain an optimistic outlook on audiobook development, with OTOBANK's general manager expressing confidence that audiobooks will continue growing as a knowledge-based entertainment product within the internet economy.”

The user notes that “OTOBANK” might be a typo for “OTOBA NK” or “OTOBA NK”. Looking at the text, it says “OTOBANK's general manager”. I should check if this is a real company. OTOBANK is indeed a Japanese audiobook company. The text has “OTOBANK” which is likely a typo. However, the instruction says to translate exactly what's given. But since this is a proper noun, I should probably use the correct name if it's clearly a typo. Let me think - the user says “If you encounter: Untranslatable proper nouns → Transliterate”. This is a proper noun, but it seems misspelled. I should probably use the correct spelling “OTOBA NK” as it's a known company. Wait, the user wrote “OTOBANK” in their notes. Let me re-read the actual text they want translated:

“Internet companies maintain an optimistic outlook on audiobook development, with OTOBANK's general manager expressing confidence that audiobooks will continue growing as a knowledge-based entertainment product within the internet economy.”

Actually, in the text to translate, it's written as “OTOBA NK” (with a space). The user wrote “OTOBANK” in their analysis. I should use what's in the

actual text: “OTOBA NK”. But that seems odd with a space. Let me check the original Chinese context if possible... The user didn’t provide the original Chinese, just this mixed text. Given that OTOBA NK is a real company (オーディオバンク in Japanese), I should probably write it as “OTOBA NK” or “OTOBA NK”. Actually, the correct English name is “OTOBA NK”. Let me use that.

“Conversely, publishers remain largely passive, with many skeptical about audiobooks’ potential as a profit pillar.”

Good as is.

“For instance, Yukio Tomikura, head of digital sales at Kodansha, has openly questioned whether audiobooks can become a revenue mainstay for publishers.”

Good as is.

“Although Kodansha launched audiobook production in 2016, it adopts a”non-lic### 3.2 Legal and Institutional Obstacles”

This seems to be cut off. “non-lic” is probably “non-licensing” or “non-exclusive licensing” strategy. The “### 3.2 Legal and Institutional Obstacles” is a section header.

Then the next paragraph:

“Article 37(3) of Japan’s Copyright Law permits”welfare service providers or persons designated by government ordinance” to produce recorded books without authorization from copyright holders for “visually impaired persons and others with perceptual disabilities regarding visually presented materials.” However, this provision applies only to the first welfare-oriented category of audiobooks, not the second commercial category. Consequently, internet companies must obtain separate authorization from both publishers and copyright holders when producing commercial audiobooks. This requirement not only increases production difficulty and extends product development cycles but also severely limits the potential for substantial growth and catalog diversification. Currently, none of the three parties—consumers, copyright holders, or publishers—can overcome these practical obstacles.”

This is already in good English and just needs minor polishing for academic flow.

Now I need to organize this into a coherent section. The structure seems to be:

- A discussion about the shift from traditional to digital publishing in the audiobook space
- The roles of publishers vs internet companies
- Perspectives from both sides (internet companies optimistic, publishers skeptical)
- A specific example from Kodansha
- Then section 3.2 about Legal and Institutional Obstacles

The user wants me to merge fragmented text into complete paragraphs. Let me structure this properly.

First, I need to identify where the section starts. The user provided what appears to be a fragment that should be part of a larger section, possibly Section 3 or 6. The user mentions “6. The overall structure...” at the beginning, but then we have “### 3.2 Legal and Institutional Obstacles” later. This is confusing.

Let me re-read the user’s prompt carefully:

“Section:”6. The overall structure and meaning should be preserved while making it flow naturally in academic ” Remaining Chinese: 介, 入

You MUST eliminate ALL Chinese characters. Even a single Chinese character will cause failure.”

This seems to be meta-instructions about the translation task itself, not part of the actual text to translate. The user is saying that Section 6 should be translated while preserving structure and meaning.

Then they provide the actual text to translate, which includes the mixed English/Chinese fragment about publishing, and then the “### 3.2 Legal and Institutional Obstacles” section.

So the structure should be:

- Some introductory paragraph(s) about the publishing ecosystem shift
- Discussion of the roles of publishers vs internet companies
- Specific examples (Amazon, OTOBA NK, Kodansha)
- Then section 3.2 about legal obstacles

The user wants me to produce flowing academic prose. Let me craft this carefully.

For the first part, I need to complete the sentence that ends with “介入” and integrate it into a proper paragraph.

Original fragment: “In traditional publishing, publishers served as the central players, but the digital publishing era has transformed the landscape into a more open ecosystem encompassing copyright holders and consumer-authors, with accessible pathways for any party to 介入.”

I’ll translate “介入” as “participate” which is more natural in this context.

Then I need to integrate the other sentences into coherent paragraphs.

Let me structure the response:

[Appropriate Section Title - maybe “The Evolving Audio-book Ecosystem” or similar]

[Paragraph 1: About the shift from traditional to digital publishing, the open ecosystem, and Amazon’s example]

[Paragraph 2: About the unresolved question of who should lead - publishers vs internet companies, with their respective strengths]

[Paragraph 3: About the differing perspectives - internet companies optimistic (OTOBA NK example) vs publishers skeptical (Kodansha example)]

3.2 Legal and Institutional Obstacles

[Paragraph 4: About Article 37(3) of Japan's Copyright Law and its limitations]

The user didn't provide a section title for the first part, but it seems to be part of a larger discussion. Since the user mentioned "Section: '6.'" at the beginning, maybe this is Section 6? But then we have "### 3.2" which suggests it's part of Section 3. This is ambiguous.

Looking at the content, the first part seems to be about market structure and industry dynamics, while 3.2 is about legal obstacles. They might be separate sections. The user might have provided two separate pieces of text that need to be translated and organized.

Let me re-examine the prompt. The user says:

"SECTION TEXT: 6. The overall structure and meaning should be preserved while making it flow naturally in academic English.

Let me translate this carefully:

[then provides the mixed text]

3.2 Legal and Institutional Obstacles

[then provides the legal text]"

The "6. The overall structure..." seems to be a note to me, not part of the actual text. The actual text starts with the mixed English/Chinese paragraph about publishing.

So I should structure it as:

[Maybe a subsection about Market Dynamics or Industry Structure]

[Paragraph 1: The ecosystem shift]

[Paragraph 2: The leadership question]

[Paragraph 3: The contrasting perspectives]

3.2 Legal and Institutional Obstacles

[Paragraph 4: Copyright law limitations]

But the user didn't give a title for the first part. Looking at the content, it could be part of a section discussing "Market Structure" or "Industry Dynamics".

Since I don't have a clear title, I could either: 1. Use a generic title like “## Industry Dynamics” 2. Integrate it as a continuation of a previous section (but I don't know what that would be) 3. Just present the paragraphs without a main section title, then have ### 3.2

Given that the user wants clean Markdown and logical organization, I think I should create an appropriate section title. The content is about the changing landscape of audiobook publishing, so something like “## The Shifting Audiobook Publishing Landscape” would be appropriate.

Wait, let me check if there's any indication of what section this belongs to. The user wrote “Section: ‘6. The overall structure...’” which might mean this is Section 6. But then we have “### 3.2” which is a subsection. This is contradictory.

I think the most reasonable interpretation is that the user provided: 1. Some meta-instructions about translation 2. A fragment that needs to be translated and integrated into a coherent paragraph (or paragraphs) 3. A clearly marked subsection “### 3.2 Legal and Institutional Obstacles” with its own text

The fragment doesn't have a clear section number, so I'll treat it as part of a main section, and then include the “### 3.2” subsection as provided.

Let me craft the translation:

For the first part, I'll create a section titled “## Market Structure and Industry Dynamics” or something similar. Then I'll write 2-3 paragraphs integrating all the points.

Paragraph 1: Start with the traditional vs digital publishing shift, mention the open ecosystem, and Amazon's example.

Paragraph 2: Discuss the unresolved leadership question between publishers and internet companies, highlighting their respective strengths (copyrights vs technology).

Paragraph 3: Contrast the optimistic view of internet companies (with OTOBA NK example) with the skeptical view of publishers (with Kodansha example).

Then the ### 3.2 section about legal obstacles. ### 3.3 Coexisting Quality and Quantity Issues with Immature Profit Models

The “quantity” issue refers to the insufficient product variety that prevents consumers from having adequate choice. The “quality” issue concerns the excessive time lag between a book's print release and its audiobook version. Audiobook production typically begins only after the print edition's release, requiring several months to complete. Even when production starts immediately after manuscript proofing, synchronizing release dates remains difficult. The optimal sales window for audiobooks is generally considered to be at the time of a new book's launch, when reader interest peaks. Consequently, audiobooks consistently miss this critical sales opportunity. Publishers Kodansha and Kadokawa

are currently working to narrow this release gap, aiming to at least have audio-book versions available for pre-order when print editions hit the market.

3.4 Inherent Defects in Japan's E-book Market

According to *E-book Business Survey Report 2020*, Japan's e-book market reached 347.3 billion yen in 2019, an increase of 64.7 billion yen (22.9% growth) from the previous year, marking two consecutive years of over 20% growth [4]. This expansion largely resulted from government crackdowns on piracy websites. However, persistent issues remain, including declining e-magazine revenues and stagnation in the text-based book market. A segmented analysis reveals that manga content generated 298.9 billion yen (an increase of 60.2 billion yen), accounting for 86.1% of the entire e-book market. In contrast, text-based content (fiction, practical books, photo collections, etc.) grew by only 4.5 billion yen to 48.4 billion yen, representing just 13.9% of the market. Manga dominates both print and digital publishing in Japan, indicating that the publishing industry has achieved little genuine innovation in digital publishing and remains reliant on traditional print media. This structural bias toward manga constitutes an inherent defect in Japan's e-book market, which in turn constrains the development of text-based audiobooks.

Overall, while audiobooks in the United States have grown at rates exceeding 10% for nine consecutive years, Japan's long-awaited "first year" of audiobook development has yet to materialize since the establishment of the Japan Audio-book Association in 2015.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.