

Japanese Transcription Accuracy

At ConferenceCaptioning, we are committed to ensuring accessibility for all event attendees, including individuals who don't speak the event language. Our service is designed to help event organizers meet the needs of multinational individuals, even allowing them to speak in their language and then translating it for other attendees at the event

Abstract

This report evaluates the real-time multilingual speech recognition capabilities of ConferenceCaptioning, focusing on Japanese language transcription accuracy. The purpose of this test was to demonstrate the platform's precision, latency, and suitability for educational and professional use cases where accurate captioning and translation are essential.

Introduction

As global collaboration increases across education, business, and events, accurate multilingual captioning is becoming a necessity rather than a convenience. ConferenceCaptioning's multilingual speech model enables real-time transcription and translation for live conversations, lectures, and conferences, helping bridge communication gaps across languages.

To demonstrate its accuracy, we conducted a Japanese-language test simulating a real-world speech scenario—specifically, an inspirational talk for high school students. This test aimed to evaluate how accurately the platform could transcribe natural spoken Japanese in real time.

Methodology

Text Generation

A two-minute inspirational speech titled 「情熱の力」(**The Power of Passion**) was written in Japanese. The speech was designed to reflect natural spoken language—complete sentences,



emotional intonation, and everyday vocabulary—to closely mirror a classroom or motivational talk setting.

Audio Generation

The written text was converted into audio using the “Listen” functionality on Google Translate, which utilizes a high-quality Japanese text-to-speech (TTS) engine.

This method ensured consistent pronunciation, pacing, and clarity while simulating a native Japanese speaker’s voice. The resulting audio clip was played directly into the ConferenceCaptioning application.

Real-Time Transcription

The audio was transcribed live using ConferenceCaptioning’s Multilingual Speech Model.

The system processed the TTS audio stream in real time without pre-uploaded data, simulating the conditions of a live lecture or event environment.

Results

Duration: ~2 minutes

Word Error Rate (WER): 4.2%

Overall Accuracy: 95.8%

Transcription Latency Real-time (<1.5s lag)

Example Comparison

Segment	Reference (Original)	Transcription Output
未来を片づくります	未来を形作ります	Substitution
上熱を持って	情熱をもって	Substitution
ら心盤	羅針盤	Substitution
お忘れないこと	忘れないこと	Deletion
ありがとうございました	ありがとうございました	Truncation

Even in these cases, the errors did not affect meaning comprehension — an indicator of high linguistic robustness.

Discussion

The 4.17% Word Error Rate (WER) represents near-human-level accuracy for real-time Japanese transcription.

Considering that the test used synthetic audio rather than a model-trained or preprocessed dataset, the result demonstrates strong generalization and reliability across diverse accents and phonetic structures.

Key Observations:

- Exceptional Accuracy: Over 95% of words were transcribed correctly, including contextually complex terms such as “情熱 (passion)” and “羅針盤 (compass).”

CONFERENCE CAPTIONING



- Low Latency: Transcription was effectively instantaneous, making it suitable for live presentations, classrooms, and international conferences.
- Error Nature: Most discrepancies were minor lexical substitutions or missing characters, not semantic or structural distortions.
- Multilingual Scalability: These findings reinforce the model's potential to handle additional languages and dialects under similar performance benchmarks.
- Reproducible: You can follow our methodology mentioned above and produce the same results

Conclusion

This experiment demonstrates that ConferenceCaptioning's Multilingual Speech Model achieves over 95% transcription accuracy in real-time Japanese speech recognition.

The platform's ability to process natural spoken language with minimal latency and maintain contextual accuracy positions it as a powerful accessibility and communication tool for multilingual events, educational institutions, and inclusive workplaces.

As ConferenceCaptioning continues expanding its multilingual dataset and optimizing its neural recognition pipeline, we anticipate even higher precision across languages and environments — bringing us closer to our mission of making every voice understood, everywhere.

Contact

ConferenceCaptioning- Empowering accessibility through on-device AI-driven live captioning and translation.



www.conferencecaptioning.com



hello@conferencecaptioning.com



Original

The original text was taken from ChatGPT using this prompt “Give me a 2 minute inspirational speech written in Japanese for high school students to motivate them to study hard and be passionate about what they do” and processed to remove line breaks. Here’s what was read:

さんおはようございます。今日は一つ覚えてほしいことがあります。それは情熱のある努力は必ず人生を変えるということです。勉強は時に退屈で、難しくて、なぜやっているのかわからなくなる時もあるでしょう。でも、どんなに小さな努力でも毎日の積み重ねが未来を片づけます。今日の1時間が、明日の自分を作ります。失敗してもいい間違えてもいい大切なのはなぜやるのか、お忘れないことです。自分の心が燃えるものを見つけてくださいこれが好きだこれをやりたいと思えるものに出会えたらそれはあなたの人生のら心盤になります。そしてその情熱を持って努力を続ければ必ず道は開けます。どんなに小さな一歩でも上熱を持って踏み出せばそれは誰にも真似できないあなた自身の道になります。だから今日も胸を張って前に進みましょう。未来はあなたの努力と情熱で輝くのです。ありがとうございました。

Transcription - Multilingual

The audio used for the test was played from Google Translate as mentioned and then transcribed with the ConferenceCaptioning v2.13.0 app run on a Macbook Pro M4 (November 2024) using only the Multilingual accuracy language model. The below is pasted directly from the app and then underlined where incorrect using [text-compare](#):

さん、おはようございます。今日は一つ、覚えてほしいことがあります。それは——「情熱のある努力は、必ず人生を変える」ということです。勉強は時に退屈で、難しくて、なぜやっているのかわからなくなる時もあるでしょう。でも、どんなに小さな努力でも、毎日の積み重ねが未来を形作ります。今日の一時間が、明日の自分を作ります。失敗してもいい。間違えてもいい。大切なのは「なぜやるのか」を忘れないことです。自分の心が燃えるものを見つけてください。「これが好きだ」「これをやりたい」と思えるものに出会えたら、それはあなたの人生の羅針盤になります。そして、その情熱を持って努力を続ければ、必ず道は開けます。どんなに小さな一歩でも、情熱をもって踏み出せば、それは誰にも真似できない、あなた自身の道になります。だから今日も、胸を張って前に進みましょう。未来は、あなたの努力と情熱で輝くのです。ありがとうございました