Transcription Software Comparison Report

Mason Mohon
September 5, 2025

Overview

For this evaluation, I compared five transcription services —Amazon Transcribe, Microsoft, Otter.ai, and OpenAI Whisper —against the human-edited control transcript. The control included 120 segments, three speakers (Thomas Sowell, Phyllis Schlafly, Bill Hayes), and 8,088 total words. Each service's transcript was evaluated against this control for accuracy, completeness, and text similarity. This evaluation only measures accuracy and does not measure automation eligibility. Additionally, this evaluation does not evaluate speaker-matching.

How the Comparison Was Done

Evaluation started with a control transcript (transcribed and edited by Cecilia). Each service's transcript was then compared line-by-line and word-by-word against that control. The code calculated several simple measures:

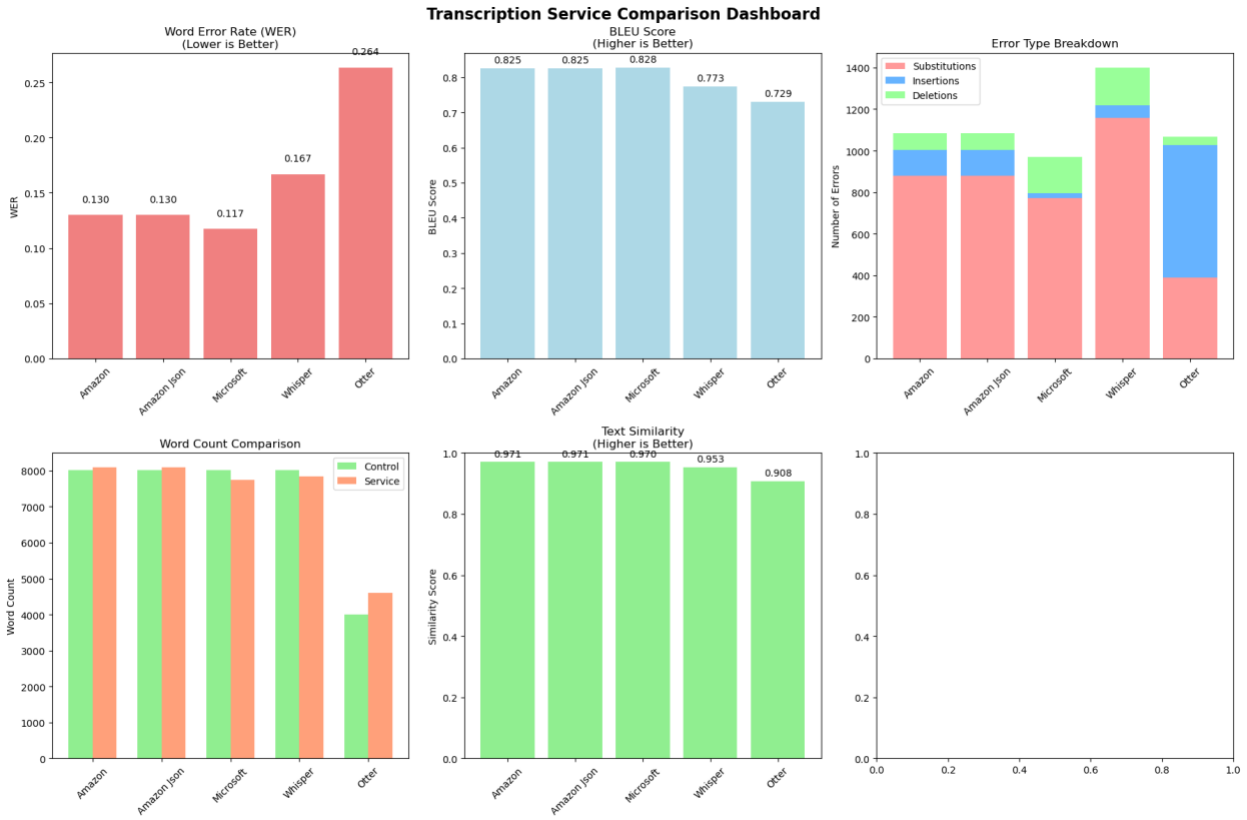
- Word Error Rate (WER): Counts how many words were wrong, missed, or extra. Lower is better.
- BLEU Score: Measures how closely the sentences match the control. Higher is better.
- Text Similarity: General "closeness" of the service transcript to the control, using pattern-matching. Higher is better.
- Error Types: Breaks down whether the mistakes were substitutions (wrong words), insertions (extra words), or deletions (missing words).
- Word Count Comparison: Shows if a service produced too few or too many words compared to the control.
- To keep it fair, I trimmed Otter.ai's transcript to the same length as the control, since it stopped short halfway through due to free trial use limitations.

The calculations were done automatically with Python. The code used standard methods like sequence matching (lining up word order), edit distance (how many fixes it takes to correct errors), and BLEU scoring (how well one sentence matches another).

Key Results

- Microsoft had the best overall accuracy, with the lowest WER (0.117) and the highest BLEU score (0.828).
- Amazon Transcribe performed very closely, both with a WER of 0.130 and a BLEU of 0.825.
- Whisper was somewhat less accurate (WER 0.167, BLEU 0.773) but still performed reasonably well.
- Otter.ai scored the lowest, with a WER of 0.264 and a noticeable transcript cut-off around 4,000 words.

The charts included in this report provide visual summaries of these results, showing error rates, word counts, similarity scores, and error breakdowns.



* Note: Amazon JSON text was the same as Amazon text, but stored in a different data format. This made not difference for performance.

Overall Rankings

- Microsoft – Best accuracy, best overall score.
- Amazon Transcribe – Very strong, nearly tied with Microsoft
- Whisper – Good, but behind Microsoft and Amazon.
- Otter.ai – Lowest performance due to missing content and higher errors.

Category Winners

- Overall Best: Microsoft
- Best for Accuracy: Microsoft
- Most Complete Transcript: Amazon

Conclusion

If accuracy is the top priority, Microsoft is the best choice. If we want a complete transcript that captures everything, Amazon Transcribe is a strong option. Whisper offers decent performance, but it may require some cleanup. Otter.ai, while easy to use, did not perform well in this test.

Although this evaluation did not test speaker-matching, Microsoft Word's transcription service includes effective and performant speaker matching. However, it is the least eligible for automation. Amazon Transcribe matches speakers effectively, and it can be integrated into an automated pipeline. However, this comes with AWS-related costs and is less user-friendly.

Appendix: Cost Overview by Service

Service	Pricing Details	Estimated Cost per Hour
Amazon Transcribe	$0.0004 per second (~$1.44 per hour) Minimum charge per request: 15 secs	~$1.44
Microsoft Azure Speech-to-Text	Standard: $0.50 per hour with high-volume commitment; pay-as-you-go may vary With an Office 365 Plan, 5 hours of transcription per month are included for free.	$0.50 (committed rate)
OpenAI Whisper API	$0.006 per minute ($0.36–$0.72 per hour)	~$0.36–$0.72
Otter.ai (Pro Plan)	$8.33 per user/month (annual billing) Includes 1,200 minutes/month (~20 hrs). Cost per hour ≈ $0.42	~$0.42/hour
Otter.ai (Business Plan)	$20/user/month (annual billing), 6,000 minutes/month (~100 hrs)	~$0.20/hour