

## Appendix A: Technical Validation

We assessed the accuracy of the transcripts by comparing a subset of the transcription data from GCP to a small sample of manually transcribed data from 2008 to 2022. More specifically, we took a random 1% from 50 episodes in the dataset. We then manually transcribed the words spoken in the audio. Then, we calculated Word Error Rate (WER), Match Error Rate (MER), Word Information Lost (WIL), Word Information Preserved (WIP), and Character Error Rate (CER). This information can be found in Table 1.

Table 1. Validation metrics of speech-to-text transcriptions of audio data to manual human coding.

metric	GCP		Whisper	
	M	SD	M	SD
WER	2.99%	2.46%	27.5%	7.83%
MER	2.96%	2.31%	26.8%	7.54%
WIL	4.22%	3.38%	42.99%	11.24%
WIP	95.77%	3.38%	57.00%	11.23%
CER	2.01%	1.82%	9.79%	3.91%

*Note:* Validation metrics of speech-to-text transcriptions of audio data. M = Mean. SD = Standard Deviation. Word Error Rate (WER) is a metric that quantifies the difference between the reference transcription (i.e., the manually transcribed text) and the output of the speech-to-text system. It measures the percentage of words in the reference transcription incorrectly recognized or omitted in the system's output. Match Error Rate (MER) is a metric that assesses the similarity between the reference transcription and the system's output by considering not only word errors but also substitutions, insertions, and deletions. It provides a more comprehensive view of transcription accuracy than WER. Word Information Lost (WIL) represents the amount of essential information missing or incorrect in the system's output compared to the reference transcription. It quantifies the loss of meaningful content in terms of words. Word Information Preserved (WIP) represents the proportion of meaningful content in the reference transcription that is correctly recognized and retained in the system's output. It quantifies the system's ability to preserve the original information in the speech. (WIP and WIL should sum to 1 as these are complementary metrics.) Character Error Rate (CER) is a metric that measures the accuracy of character-level transcription. It calculates the percentage of characters in the reference transcription that differ from the characters in the system's output, including substitutions, insertions, and deletions ([Morris et al, 2004](#)). For WER, MER, WIL, CER, a lower number is more desirable. For WIP, a higher number is more desirable.