Soniox Speech-to-Text Benchmarks

Soniox Inc, March 2025
support@soniox.com
https://soniox.com

Abstract

Soniox conducted a comprehensive evaluation of the accuracy of various speech recognition providers in the industry. The benchmarking results are summarized as follows:

- **Providers evaluated:** Soniox, OpenAI, Google, AWS, Azure, NVIDIA, Deepgram, AssemblyAI, Speechmatics, and ElevenLabs.
- **Languages evaluated:** 60 languages.
- **Evaluation datasets:** Real-world datasets of YouTube videos for each language, covering diverse acoustic conditions, speaking styles, accents, topics, and speaker variations.
- **Ground truth transcriptions:** Transcribed and double-reviewed by humans, then normalized to ensure a fair evaluation across providers.
- **Processing modes evaluated:** Asynchronous (file/batch) transcription.
- **Results:** Soniox achieved the highest speech recognition accuracy across most languages by a significant margin.

Results

Evaluated Providers and Languages

To assess the accuracy of speech recognition providers across multiple languages, we conducted a rigorous benchmarking study using **Word Error Rate (WER)** and **Character Error Rate (CER)** as the primary evaluation metrics. These industry-standard metrics provide a quantitative measure of transcription accuracy, with lower values indicating superior performance. CER was used for the following languages: Korean, Chinese, Japanese, and Thai.

Our evaluation was based on **45 to 70 minutes of real-world audio per language**, sourced from YouTube to ensure a diverse and challenging dataset. The selected samples encompass **various acoustic conditions, speaking styles, accents, and topics**, providing a robust assessment of model performance in real-world scenarios. **Ground truth transcriptions** were carefully transcribed and double-reviewed by humans, then normalized to ensure consistency and fairness across all providers.

The table below presents accuracy results (WER and CER) for each speech recognition provider across the evaluated languages. An empty cell in the table indicates that the provider does not support that language.

Comparative Performance Analysis

In addition to tabulated results, we present visual comparisons of Soniox's performance against other providers. These plots illustrate accuracy differences between Soniox and competitors such as OpenAI, Google, AWS, and others across all languages.

Methodology

Evaluation Process

1. **Dataset Selection:** For each language, real-world YouTube videos were chosen to reflect diverse speech patterns, varying acoustic conditions, and multiple speaker types.
2. **Transcription & Ground Truth Creation:** All dataset transcriptions were manually created, double-reviewed by humans, and normalized to provide a consistent reference for evaluation. The normalization process included removing punctuation and ignoring capitalization; otherwise, the ground truth transcription remained unchanged.
3. **Model Integration:** Each provider's API was carefully integrated according to official documentation, ensuring a fair and accurate comparison.
4. **Evaluation Metrics:**
    - **WER (Word Error Rate):** Measures transcription errors at the word level.
    - **CER (Character Error Rate):** Used for logographic or non-space-separated languages to provide a finer-grained accuracy measurement.
5. **Processing Mode:** All models were evaluated in **asynchronous (file/batch) transcription mode** to ensure consistency in testing.

Models Evaluated

Provider	Model Evaluated
Soniox	`stt-async-preview`
OpenAI	`Whisper large-v3`
Google	`long` (for supported languages), `chirp_2` (for other languages)
AWS	`Best/Default`
Azure	`Best/Default`

Provider	Model Evaluated
NVIDIA	`conformer-{lang}-asr-offline-asr-bls-ensemble` (for supported languages), `parakeet-1.1b-unified-ml-cs-universal-multi-asr-offline-asr-bls-ensemble` (for other languages)
Deepgram	`nova-3` (for English), `nova-2` (for other languages)
AssemblyAI	`best` (for supported languages), `nano` (for other languages)
Speechmatics	`enhanced`
ElevenLabs	`scribe_v1`

This evaluation provides a transparent and rigorous comparison of speech recognition performance across industry-leading providers for 60 languages.

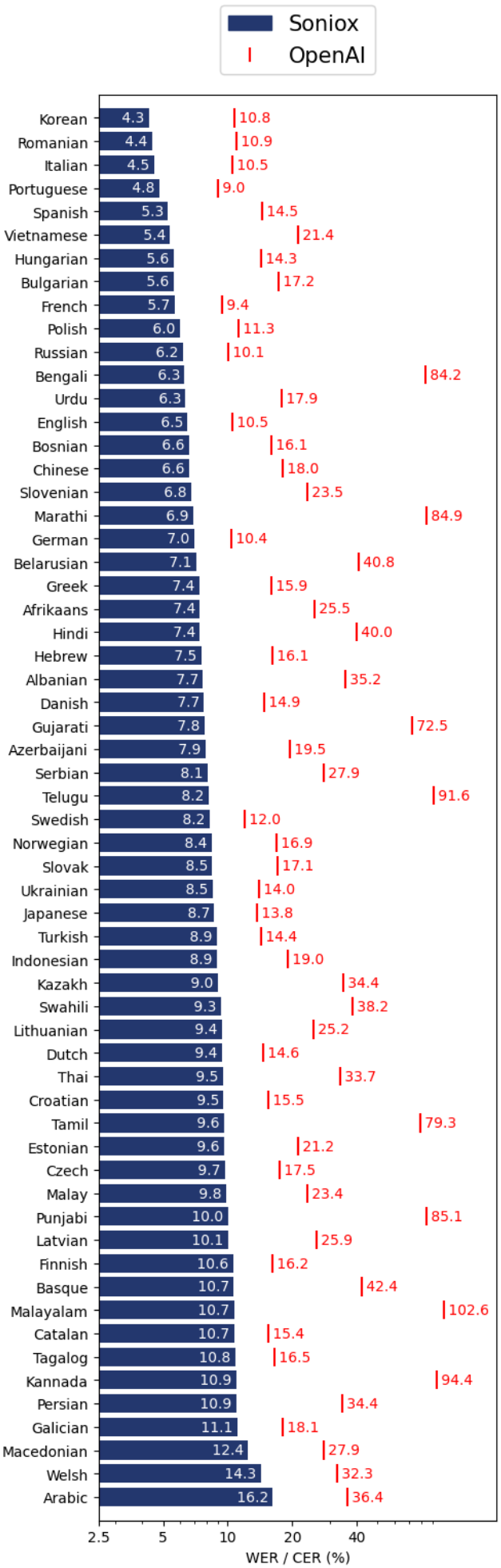
Table: Evaluated Providers and Languages

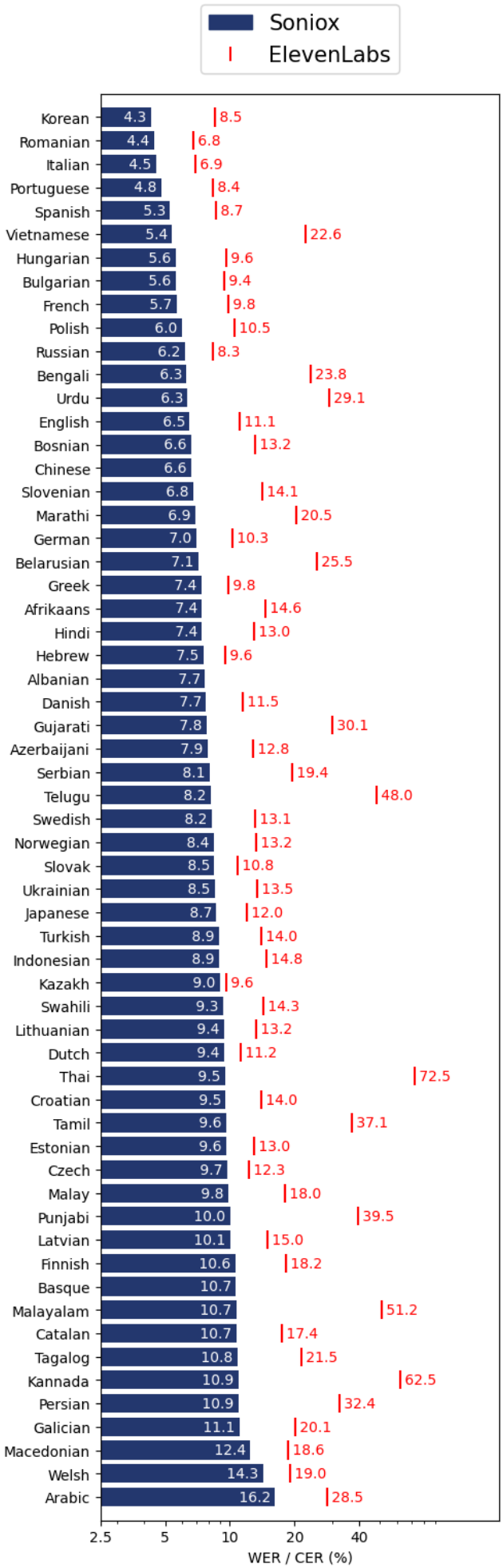
Language	Soniox	OpenAI	Google	AWS	Azure	Speech-matics	AssemblyAI	Deepgram	NVIDIA	ElevenLabs
Afrikaans	7.4	25.5	7.4	**6.5**	12.7		64.0			14.6
Albanian	**7.7**	35.2	10.3		7.9		91.1			
Arabic	**16.2**	36.4	28.8	42.2	37.1	22.9	55.6		44.5	28.5
Azerbaijani	**7.9**	19.5	12.6	13.5	19.9		49.5			12.8
Basque	10.7	42.4	15.1	11.8	18.2	**7.2**	77.4			
Belarusian	**7.1**	40.8	10.1	9.8		10.2	74.1			25.5
Bengali	**6.3**	84.2	48.9	17.7	40.7	27.7	101.1			23.8
Bosnian	**6.6**	16.1	12.3	15.5	19.6		34.2			13.2
Bulgarian	**5.6**	17.2	15.3	10.8	13.9	9.2	36.2	24.8		9.4
Catalan	**10.7**	15.4	21.7	17.8	24.6	13.9	34.7	19.3		17.4
Chinese	**6.6**	18.0	54.1	16.8	10.1	14.2	19.9	94.4	17.8	
Croatian	**9.5**	15.5	14.1	13.7	21.1	13.4	38.2			14.0
Czech	**9.7**	17.5	18.2	17.3	17.9	14.4	38.2	22.1	98.7	12.3
Danish	**7.7**	14.9	18.0	13.6	15.0	10.7	35.6	16.5	93.1	11.5
Dutch	**9.4**	14.6	23.2	12.0	14.4	11.6	10.6	12.1	22.0	11.2
English	**6.5**	10.5	18.2	11.4	9.4	9.3	9.3	9.9	10.1	11.1
Estonian	**9.6**	21.2	13.4	15.4	15.1	12.3	62.1	32.9		13.0
Finnish	**10.6**	16.2	20.3	17.9	19.9	15.2	15.8	15.8		18.2
French	**5.7**	9.4	15.3	10.9	12.4	8.3	8.3	10.7	16.1	9.8
Galician	**11.1**	18.1	18.8	16.6	21.9	16.0	48.5			20.1
German	**7.0**	10.4	16.1	13.0	10.8	9.0	8.1	12.5	16.5	10.3
Greek	**7.4**	15.9	10.3	17.1	13.2	9.5	31.1	17.4		9.8
Gujarati	**7.8**	72.5	22.4	17.2	62.8		101.7			30.1
Hebrew	**7.5**	16.1	12.1	11.2	15.8	10.6	34.4		87.2	9.6
Hindi	**7.4**	40.0	20.0	18.9	18.3	14.0	20.6	25.2	17.4	13.0
Hungarian	**5.6**	14.3	25.2	14.7	13.6	9.5	33.9	24.7		9.6
Indonesian	**8.9**	19.0	14.5	16.0	17.0	15.5	28.7	17.1		14.8
Italian	**4.5**	10.5	9.6	9.9	9.0	7.4	7.6	10.8	12.2	6.9
Japanese	**8.7**	13.8	14.2	16.2	14.0	10.3	14.8	11.7	21.9	12.0
Kannada	**10.9**	94.4	56.6	31.4	58.8		100.4			62.5
Kazakh	9.0	34.4	67.8	**7.4**	13.9		81.5			9.6
Korean	**4.3**	10.8	11.3	13.3	11.2	8.1	11.4	12.8	13.8	8.5
Latvian	**10.1**	25.9	13.9	17.2	16.2	14.2	60.7	53.7		15.0
Lithuanian	**9.4**	25.2	13.6	16.4	15.7	12.4	62.0	36.6		13.2
Macedonian	**12.4**	27.9	28.2	12.5	24.3		57.1			18.6
Malay	**9.8**	23.4	16.8	25.0	20.7	20.3	30.6	27.2		18.0
Malayalam	**10.7**	102.6	42.2	31.0	59.0		102.8			51.2
Marathi	**6.9**	84.9	43.4	15.3	40.0	24.5	98.5			20.5
Norwegian	**8.4**	16.9	20.8	15.1	14.7	11.8	37.2	17.5	90.5	13.2
Persian	**10.9**	34.4	21.7	22.3	40.1	24.2	67.5			32.4
Polish	**6.0**	11.3	13.6	14.7	11.1	7.7	10.3	10.3	97.4	10.5
Portuguese	**4.8**	9.0	34.1	13.4	10.7	8.1	8.0	13.4	18.3	8.4
Punjabi	**10.0**	85.1	19.9	27.5	48.1		100.9			39.5
Romanian	**4.4**	10.9	11.3	8.5	9.6	6.8	25.5	27.5		6.8
Russian	**6.2**	10.1	16.7	14.3	16.5	11.7	11.1	8.0	16.7	8.3
Serbian	**8.1**	27.9	11.3	*	*		36.6			19.4
Slovak	**8.5**	17.1	12.3	14.3	16.2	11.8	41.4	23.6		10.8
Slovenian	**6.8**	23.5	10.8	15.6	16.7	11.8	55.6			14.1

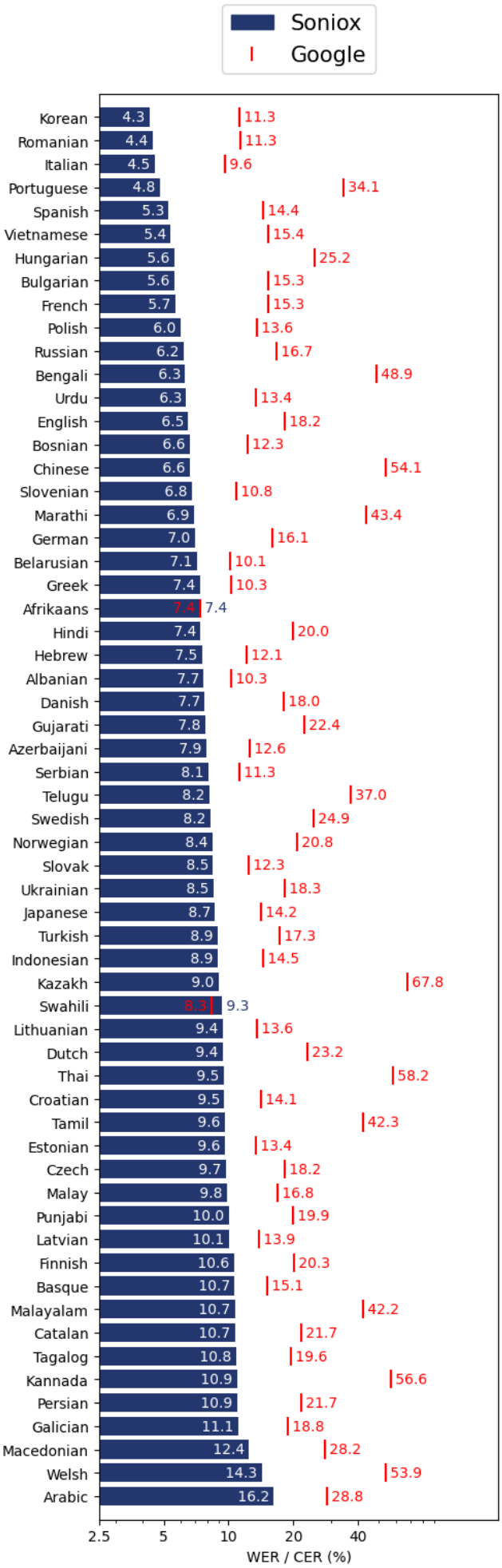
Spanish	**5.3**	14.5	14.4	11.4	9.5	7.1	7.1	10.1	14.5	8.7
Swahili	9.3	38.2	8.3	**7.8**	10.3	11.4	84.5			14.3
Swedish	**8.2**	12.0	24.9	15.3	14.1	11.2	28.4	13.6	91.4	13.1
Tagalog	**10.8**	16.5	19.6	17.2	16.9		34.5			21.5
Tamil	**9.6**	79.3	42.3	41.2	47.2	37.6	88.5			37.1
Telugu	**8.2**	91.6	37.0	26.9	51.7		99.9			48.0
Thai	**9.5**	33.7	58.2	14.2	15.5	13.3	85.6	48.6	92.5	72.5
Turkish	**8.9**	14.4	17.3	21.1	18.0	13.3	11.6	15.4	92.0	14.0
Ukrainian	**8.5**	14.0	18.3	15.8	19.1	15.7	13.3	14.1		13.5
Urdu	**6.3**	17.9	13.4		19.3	33.9	42.6			29.1
Vietnamese	**5.4**	21.4	15.4	11.2	11.2	8.9	12.1	14.2		22.6
Welsh	**14.3**	32.3	53.9	19.1	28.7	19.7	75.8			19.0

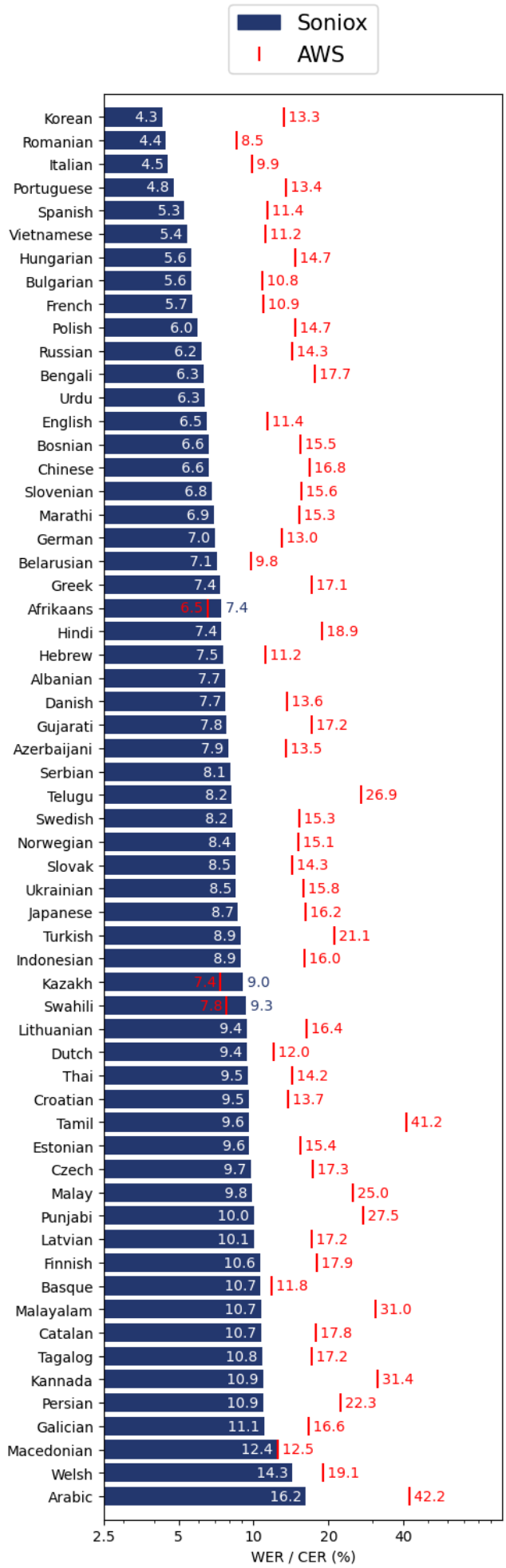
Notes:

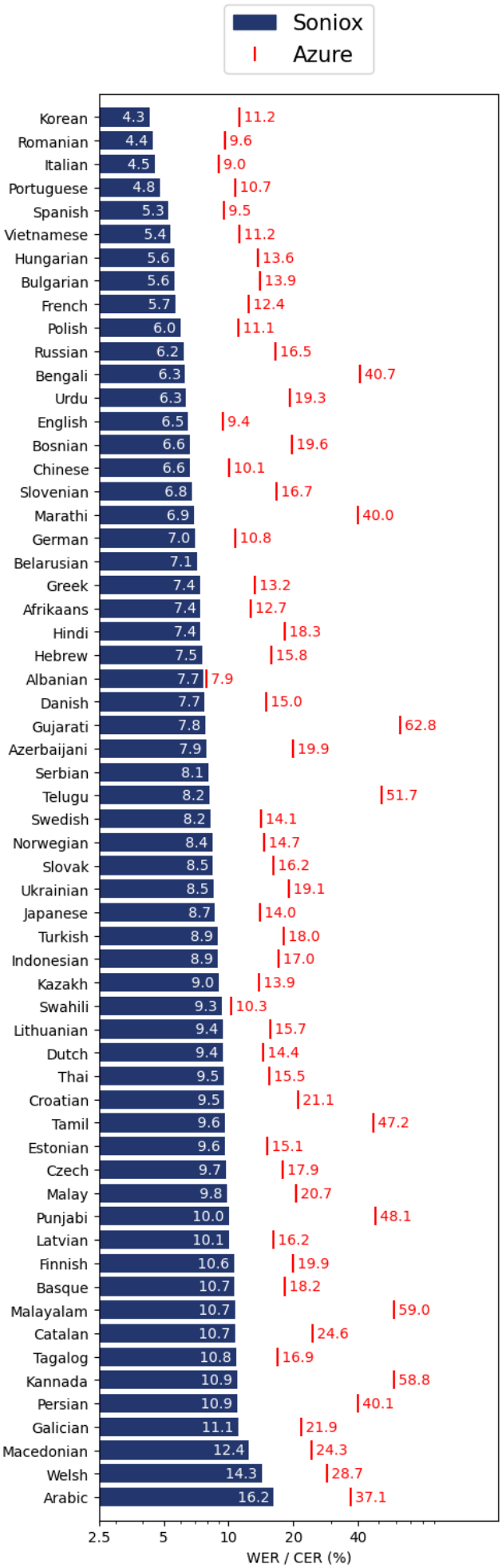
- * For the Serbian language, AWS and Azure transcribe in Cyrillic, not Latin, whereas the ground truth transcripts are in Latin; therefore, the WER results are omitted.
- We manually reviewed the returned transcripts of provider-language pairs with high WER (e.g., AssemblyAI on Bengali) and found that the high WER was due to the provider's STT model generating low-quality transcripts rather than other technical reasons.

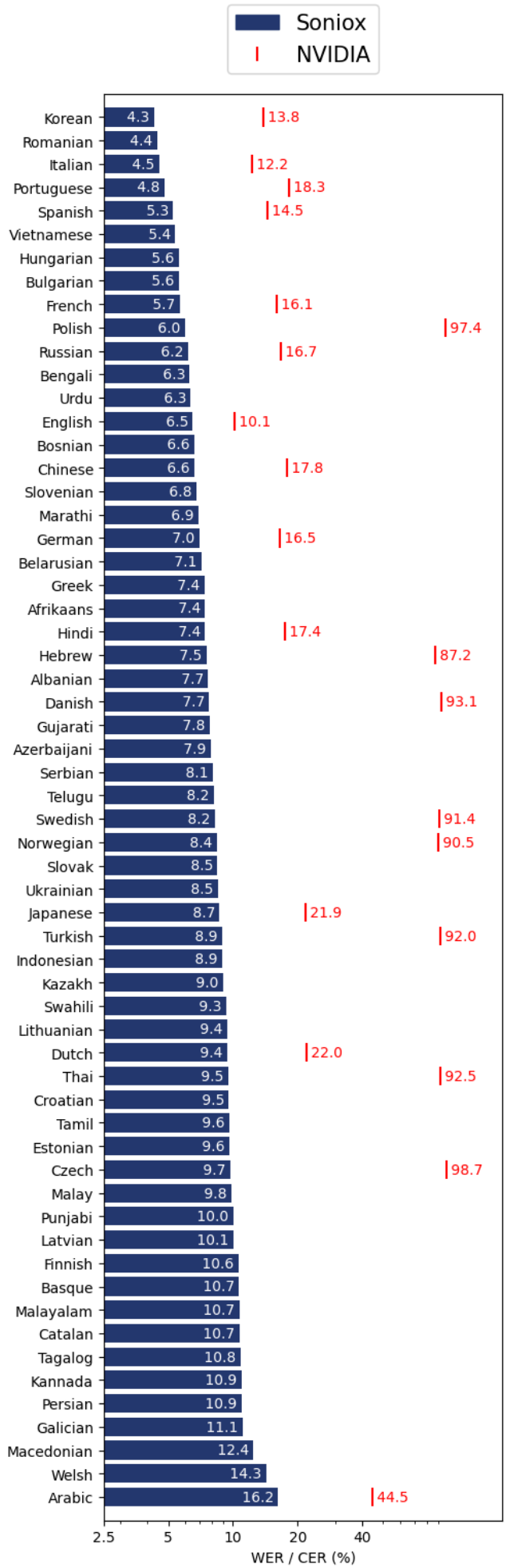

Soniox
OpenAI
Korean 4.3 10.8
Romanian 4.4 10.9
Italian 4.5 10.5
Portuguese 4.8 9.0
Spanish 5.3 14.5
Vietnamese 5.4 21.4
Hungarian 5.6 14.3
Bulgarian 5.6 17.2
French 5.7 9.4
Polish 6.0 11.3
Russian 6.2 10.1
Bengali 6.3 84.2
Urdu 6.3 17.9
English 6.5 10.5
Bosnian 6.6 16.1
Chinese 6.6 18.0
Slovenian 6.8 23.5
Marathi 6.9 84.9
German 7.0 10.4
Belarusian 7.1 40.8
Greek 7.4 15.9
Afrikaans 7.4 25.5
Hindi 7.4 40.0
Hebrew 7.5 16.1
Albanian 7.7 35.2
Danish 7.7 14.9
Gujarati 7.8 72.5
Azerbaijani 7.9 19.5
Serbian 8.1 27.9
Telugu 8.2 91.6
Swedish 8.2 12.0
Norwegian 8.4 16.9
Slovak 8.5 17.1
Ukrainian 8.5 14.0
Japanese 8.7 13.8
Turkish 8.9 14.4
Indonesian 8.9 19.0
Kazakh 9.0 34.4
Swahili 9.3 38.2
Lithuanian 9.4 25.2
Dutch 9.4 14.6
Thai 9.5 33.7
Croatian 9.5 15.5
Tamil 9.6 79.3
Estonian 9.6 21.2
Czech 9.7 17.5
Malay 9.8 23.4
Punjabi 10.0 85.1
Latvian 10.1 25.9
Finnish 10.6 16.2
Basque 10.7 42.4
Malayalam 10.7 102.6
Catalan 10.7 15.4
Tagalog 10.8 16.5
Kannada 10.9 94.4
Persian 10.9 34.4
Galician 11.1 18.1
Macedonian 12.4 27.9
Welsh 14.3 32.3
Arabic 16.2 36.4
2.5 5 10 20 40
WER / CER (%)

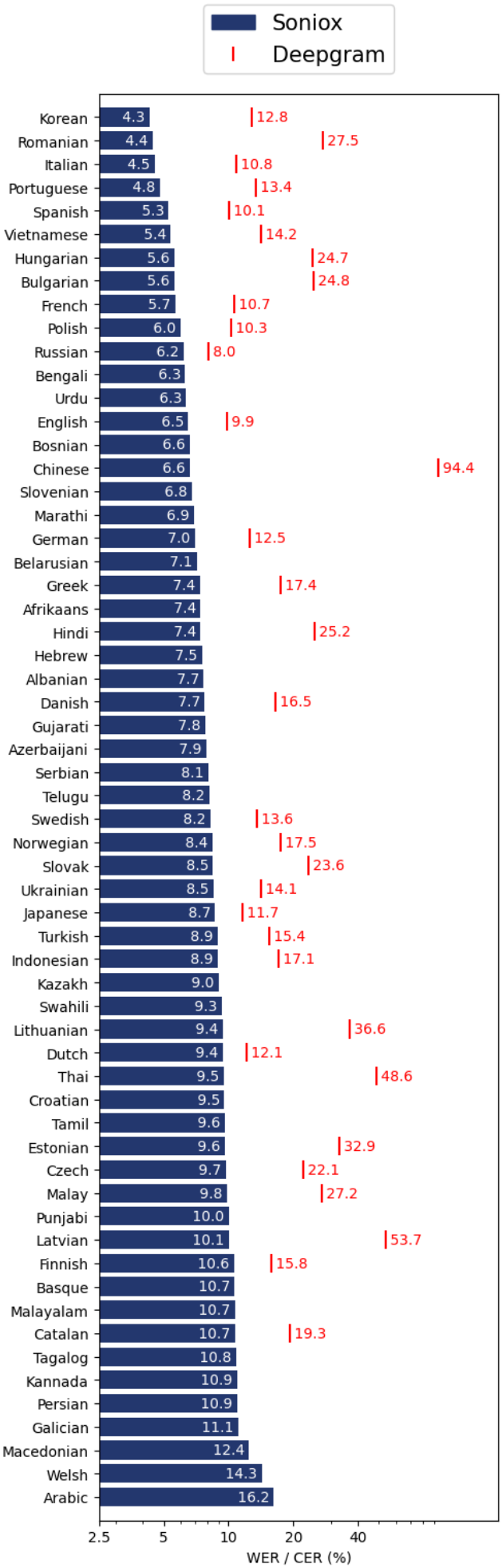

Soniox
ElevenLabs
Korean 4.3 8.5
Romanian 4.4 6.8
Italian 4.5 6.9
Portuguese 4.8 8.4
Spanish 5.3 8.7
Vietnamese 5.4 22.6
Hungarian 5.6 9.6
Bulgarian 5.6 9.4
French 5.7 9.8
Polish 6.0 10.5
Russian 6.2 8.3
Bengali 6.3 23.8
Urdu 6.3 29.1
English 6.5 11.1
Bosnian 6.6 13.2
Chinese 6.6
Slovenian 6.8 14.1
Marathi 6.9 20.5
German 7.0 10.3
Belarusian 7.1 25.5
Greek 7.4 9.8
Afrikaans 7.4 14.6
Hindi 7.4 13.0
Hebrew 7.5 9.6
Albanian 7.7
Danish 7.7 11.5
Gujarati 7.8 30.1
Azerbaijani 7.9 12.8
Serbian 8.1 19.4
Telugu 8.2 48.0
Swedish 8.2 13.1
Norwegian 8.4 13.2
Slovak 8.5 10.8
Ukrainian 8.5 13.5
Japanese 8.7 12.0
Turkish 8.9 14.0
Indonesian 8.9 14.8
Kazakh 9.0 9.6
Swahili 9.3 14.3
Lithuanian 9.4 13.2
Dutch 9.4 11.2
Thai 9.5 72.5
Croatian 9.5 14.0
Tamil 9.6 37.1
Estonian 9.6 13.0
Czech 9.7 12.3
Malay 9.8 18.0
Punjabi 10.0 39.5
Latvian 10.1 15.0
Finnish 10.6 18.2
Basque 10.7
Malayalam 10.7 51.2
Catalan 10.7 17.4
Tagalog 10.8 21.5
Kannada 10.9 62.5
Persian 10.9 32.4
Galician 11.1 20.1
Macedonian 12.4 18.6
Welsh 14.3 19.0
Arabic 16.2 28.5
2.5 5 10 20 40
WER / CER (%)

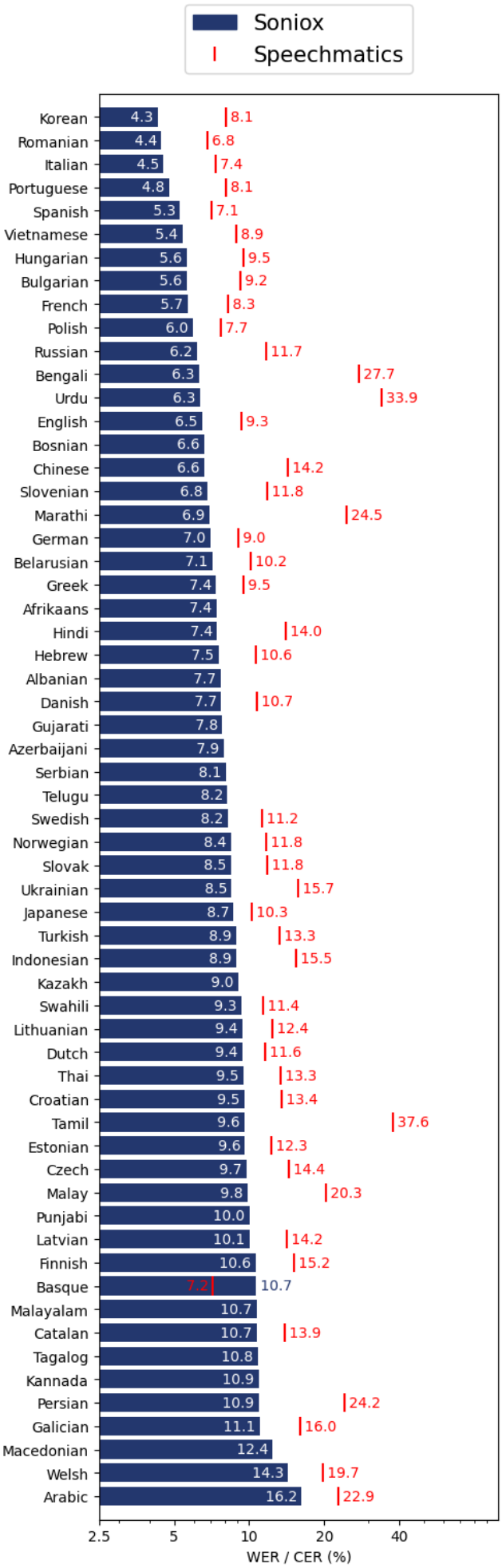

Soniox
Google
Korean 4.3 11.3
Romanian 4.4 11.3
Italian 4.5 9.6
Portuguese 4.8 34.1
Spanish 5.3 14.4
Vietnamese 5.4 15.4
Hungarian 5.6 25.2
Bulgarian 5.6 15.3
French 5.7 15.3
Polish 6.0 13.6
Russian 6.2 16.7
Bengali 6.3 48.9
Urdu 6.3 13.4
English 6.5 18.2
Bosnian 6.6 12.3
Chinese 6.6 54.1
Slovenian 6.8 10.8
Marathi 6.9 43.4
German 7.0 16.1
Belarusian 7.1 10.1
Greek 7.4 10.3
Afrikaans 7.4 7.4
Hindi 7.4 20.0
Hebrew 7.5 12.1
Albanian 7.7 10.3
Danish 7.7 18.0
Gujarati 7.8 22.4
Azerbaijani 7.9 12.6
Serbian 8.1 11.3
Telugu 8.2 37.0
Swedish 8.2 24.9
Norwegian 8.4 20.8
Slovak 8.5 12.3
Ukrainian 8.5 18.3
Japanese 8.7 14.2
Turkish 8.9 17.3
Indonesian 8.9 14.5
Kazakh 9.0 67.8
Swahili 8.3 9.3
Lithuanian 9.4 13.6
Dutch 9.4 23.2
Thai 9.5 58.2
Croatian 9.5 14.1
Tamil 9.6 42.3
Estonian 9.6 13.4
Czech 9.7 18.2
Malay 9.8 16.8
Punjabi 10.0 19.9
Latvian 10.1 13.9
Finnish 10.6 20.3
Basque 10.7 15.1
Malayalam 10.7 42.2
Catalan 10.7 21.7
Tagalog 10.8 19.6
Kannada 10.9 56.6
Persian 10.9 21.7
Galician 11.1 18.8
Macedonian 12.4 28.2
Welsh 14.3 53.9
Arabic 16.2 28.8
2.5 5 10 20 40
WER / CER (%)


Soniox
AWS
Korean 4.3 13.3
Romanian 4.4 8.5
Italian 4.5 9.9
Portuguese 4.8 13.4
Spanish 5.3 11.4
Vietnamese 5.4 11.2
Hungarian 5.6 14.7
Bulgarian 5.6 10.8
French 5.7 10.9
Polish 6.0 14.7
Russian 6.2 14.3
Bengali 6.3 17.7
Urdu 6.3
English 6.5 11.4
Bosnian 6.6 15.5
Chinese 6.6 16.8
Slovenian 6.8 15.6
Marathi 6.9 15.3
German 7.0 13.0
Belarusian 7.1 9.8
Greek 7.4 17.1
Afrikaans 6.5 7.4
Hindi 7.4 18.9
Hebrew 7.5 11.2
Albanian 7.7
Danish 7.7 13.6
Gujarati 7.8 17.2
Azerbaijani 7.9 13.5
Serbian 8.1
Telugu 8.2 26.9
Swedish 8.2 15.3
Norwegian 8.4 15.1
Slovak 8.5 14.3
Ukrainian 8.5 15.8
Japanese 8.7 16.2
Turkish 8.9 21.1
Indonesian 8.9 16.0
Kazakh 7.4 9.0
Swahili 7.8 9.3
Lithuanian 9.4 16.4
Dutch 9.4 12.0
Thai 9.5 14.2
Croatian 9.5 13.7
Tamil 9.6 41.2
Estonian 9.6 15.4
Czech 9.7 17.3
Malay 9.8 25.0
Punjabi 10.0 27.5
Latvian 10.1 17.2
Finnish 10.6 17.9
Basque 10.7 11.8
Malayalam 10.7 31.0
Catalan 10.7 17.8
Tagalog 10.8 17.2
Kannada 10.9 31.4
Persian 10.9 22.3
Galician 11.1 16.6
Macedonian 12.4 12.5
Welsh 14.3 19.1
Arabic 16.2 42.2
2.5 5 10 20 40
WER / CER (%)


Soniox
Azure
Korean 4.3 11.2
Romanian 4.4 9.6
Italian 4.5 9.0
Portuguese 4.8 10.7
Spanish 5.3 9.5
Vietnamese 5.4 11.2
Hungarian 5.6 13.6
Bulgarian 5.6 13.9
French 5.7 12.4
Polish 6.0 11.1
Russian 6.2 16.5
Bengali 6.3 40.7
Urdu 6.3 19.3
English 6.5 9.4
Bosnian 6.6 19.6
Chinese 6.6 10.1
Slovenian 6.8 16.7
Marathi 6.9 40.0
German 7.0 10.8
Belarusian 7.1
Greek 7.4 13.2
Afrikaans 7.4 12.7
Hindi 7.4 18.3
Hebrew 7.5 15.8
Albanian 7.7 7.9
Danish 7.7 15.0
Gujarati 7.8 62.8
Azerbaijani 7.9 19.9
Serbian 8.1
Telugu 8.2 51.7
Swedish 8.2 14.1
Norwegian 8.4 14.7
Slovak 8.5 16.2
Ukrainian 8.5 19.1
Japanese 8.7 14.0
Turkish 8.9 18.0
Indonesian 8.9 17.0
Kazakh 9.0 13.9
Swahili 9.3 10.3
Lithuanian 9.4 15.7
Dutch 9.4 14.4
Thai 9.5 15.5
Croatian 9.5 21.1
Tamil 9.6 47.2
Estonian 9.6 15.1
Czech 9.7 17.9
Malay 9.8 20.7
Punjabi 10.0 48.1
Latvian 10.1 16.2
Finnish 10.6 19.9
Basque 10.7 18.2
Malayalam 10.7 59.0
Catalan 10.7 24.6
Tagalog 10.8 16.9
Kannada 10.9 58.8
Persian 10.9 40.1
Galician 11.1 21.9
Macedonian 12.4 24.3
Welsh 14.3 28.7
Arabic 16.2 37.1
2.5 5 10 20 40
WER / CER (%)


Soniox
NVIDIA
Korean 4.3 13.8
Romanian 4.4
Italian 4.5 12.2
Portuguese 4.8 18.3
Spanish 5.3 14.5
Vietnamese 5.4
Hungarian 5.6
Bulgarian 5.6
French 5.7 16.1
Polish 6.0 97.4
Russian 6.2 16.7
Bengali 6.3
Urdu 6.3
English 6.5 10.1
Bosnian 6.6
Chinese 6.6 17.8
Slovenian 6.8
Marathi 6.9
German 7.0 16.5
Belarusian 7.1
Greek 7.4
Afrikaans 7.4
Hindi 7.4 17.4
Hebrew 7.5 87.2
Albanian 7.7
Danish 7.7 93.1
Gujarati 7.8
Azerbaijani 7.9
Serbian 8.1
Telugu 8.2
Swedish 8.2 91.4
Norwegian 8.4 90.5
Slovak 8.5
Ukrainian 8.5
Japanese 8.7 21.9
Turkish 8.9 92.0
Indonesian 8.9
Kazakh 9.0
Swahili 9.3
Lithuanian 9.4
Dutch 9.4 22.0
Thai 9.5 92.5
Croatian 9.5
Tamil 9.6
Estonian 9.6
Czech 9.7 98.7
Malay 9.8
Punjabi 10.0
Latvian 10.1
Finnish 10.6
Basque 10.7
Malayalam 10.7
Catalan 10.7
Tagalog 10.8
Kannada 10.9
Persian 10.9
Galician 11.1
Macedonian 12.4
Welsh 14.3
Arabic 16.2 44.5
2.5 5 10 20 40
WER / CER (%)


Soniox
Deepgram
Korean 4.3 12.8
Romanian 4.4 27.5
Italian 4.5 10.8
Portuguese 4.8 13.4
Spanish 5.3 10.1
Vietnamese 5.4 14.2
Hungarian 5.6 24.7
Bulgarian 5.6 24.8
French 5.7 10.7
Polish 6.0 10.3
Russian 6.2 8.0
Bengali 6.3
Urdu 6.3
English 6.5 9.9
Bosnian 6.6
Chinese 6.6 94.4
Slovenian 6.8
Marathi 6.9
German 7.0 12.5
Belarusian 7.1
Greek 7.4 17.4
Afrikaans 7.4
Hindi 7.4 25.2
Hebrew 7.5
Albanian 7.7
Danish 7.7 16.5
Gujarati 7.8
Azerbaijani 7.9
Serbian 8.1
Telugu 8.2
Swedish 8.2 13.6
Norwegian 8.4 17.5
Slovak 8.5 23.6
Ukrainian 8.5 14.1
Japanese 8.7 11.7
Turkish 8.9 15.4
Indonesian 8.9 17.1
Kazakh 9.0
Swahili 9.3
Lithuanian 9.4 36.6
Dutch 9.4 12.1
Thai 9.5 48.6
Croatian 9.5
Tamil 9.6
Estonian 9.6 32.9
Czech 9.7 22.1
Malay 9.8 27.2
Punjabi 10.0
Latvian 10.1 53.7
Finnish 10.6 15.8
Basque 10.7
Malayalam 10.7
Catalan 10.7 19.3
Tagalog 10.8
Kannada 10.9
Persian 10.9
Galician 11.1
Macedonian 12.4
Welsh 14.3
Arabic 16.2
2.5 5 10 20 40
WER / CER (%)
Soniox
AssemblyAI
Korean 4.3 11.4
Romanian 4.4 25.5
Italian 4.5 7.6
Portuguese 4.8 8.0
Spanish 5.3 7.1
Vietnamese 5.4 12.1
Hungarian 5.6 33.9
Bulgarian 5.6 36.2
French 5.7 8.3
Polish 6.0 10.3
Russian 6.2 11.1
Bengali 6.3 101.1
Urdu 6.3 42.6
English 6.5 9.3
Bosnian 6.6 34.2
Chinese 6.6 19.9
Slovenian 6.8 55.6
Marathi 6.9 98.5
German 7.0 8.1
Belarusian 7.1 74.1
Greek 7.4 31.1
Afrikaans 7.4 64.0
Hindi 7.4 20.6
Hebrew 7.5 34.4
Albanian 7.7 91.1
Danish 7.7 35.6
Gujarati 7.8 101.7
Azerbaijani 7.9 49.5
Serbian 8.1 36.6
Telugu 8.2 99.9
Swedish 8.2 28.4
Norwegian 8.4 37.2
Slovak 8.5 41.4
Ukrainian 8.5 13.3
Japanese 8.7 14.8
Turkish 8.9 11.6
Indonesian 8.9 28.7
Kazakh 9.0 81.5
Swahili 9.3 84.5
Lithuanian 9.4 62.0
Dutch 9.4 10.6
Thai 9.5 85.6
Croatian 9.5 38.2
Tamil 9.6 88.5
Estonian 9.6 62.1
Czech 9.7 38.2
Malay 9.8 30.6
Punjabi 10.0 100.9
Latvian 10.1 60.7
Finnish 10.6 15.8
Basque 10.7 77.4
Malayalam 10.7 102.8
Catalan 10.7 34.7
Tagalog 10.8 34.5
Kannada 10.9 100.4
Persian 10.9 67.5
Galician 11.1 48.5
Macedonian 12.4 57.1
Welsh 14.3 75.8
Arabic 16.2 55.6
2.5 5 10 20 40
WER / CER (%)


Soniox
Speechmatics
Korean 4.3 8.1
Romanian 4.4 6.8
Italian 4.5 7.4
Portuguese 4.8 8.1
Spanish 5.3 7.1
Vietnamese 5.4 8.9
Hungarian 5.6 9.5
Bulgarian 5.6 9.2
French 5.7 8.3
Polish 6.0 7.7
Russian 6.2 11.7
Bengali 6.3 27.7
Urdu 6.3 33.9
English 6.5 9.3
Bosnian 6.6
Chinese 6.6 14.2
Slovenian 6.8 11.8
Marathi 6.9 24.5
German 7.0 9.0
Belarusian 7.1 10.2
Greek 7.4 9.5
Afrikaans 7.4
Hindi 7.4 14.0
Hebrew 7.5 10.6
Albanian 7.7
Danish 7.7 10.7
Gujarati 7.8
Azerbaijani 7.9
Serbian 8.1
Telugu 8.2
Swedish 8.2 11.2
Norwegian 8.4 11.8
Slovak 8.5 11.8
Ukrainian 8.5 15.7
Japanese 8.7 10.3
Turkish 8.9 13.3
Indonesian 8.9 15.5
Kazakh 9.0
Swahili 9.3 11.4
Lithuanian 9.4 12.4
Dutch 9.4 11.6
Thai 9.5 13.3
Croatian 9.5 13.4
Tamil 9.6 37.6
Estonian 9.6 12.3
Czech 9.7 14.4
Malay 9.8 20.3
Punjabi 10.0
Latvian 10.1 14.2
Finnish 10.6 15.2
Basque 7.2 10.7
Malayalam 10.7
Catalan 10.7 13.9
Tagalog 10.8
Kannada 10.9
Persian 10.9 24.2
Galician 11.1 16.0
Macedonian 12.4
Welsh 14.3 19.7
Arabic 16.2 22.9
2.5
5
10
20
40
WER / CER (%)