PfeifferReport

Premiere Pro:
The Productivity Impact
of *Speech to Text*

About this Benchmark Project

This report presents the findings of a market-specific benchmarking project conducted by Pfeiffer Consulting for Adobe. The main aim of the research was to **document the efficiency and productivity gains of the *Speech to Text* feature set** of Adobe Premiere Pro, compared to other methods for creating captions in videos.

Benchmarks were executed using *Pfeiffer Consulting's Methodology for Productivity Benchmarking*, which has been fine-tuned over more than a decade, and measures the time experienced operators take to execute specific tasks. Please refer to the Methodology section on the last page of this document for more information

About Speech-to-Text

Market research confirms that **adding captions to video content significantly increases engagement.** But adding captions to a video can be a daunting and time-consuming task. ***Speech to Text*, introduced with the 15.4 release of Premiere Pro can automate this process**, by generating captions from one of the audio tracks of a video, using Adobe Sensei technology. Currently, 13 languages are supported.

The benchmarks conducted for this project document the productivity-gains of the feature set. We compared using *Speech to Text* on real-world footage of varying length, measuring not only the time it takes to manually create captions, but also when using automatically generated captions from several online transcription services. **On average, using *Speech to Text* was almost five times faster than other methods in these benchmarks.** (See chart below.)

Executive Summary

▷ The *Speech to Text* feature set introduced in Premiere Pro (v.15.4) **automates the creation of captions** using Adobe Sensei machine learning technology.

▷ In benchmarks conducted for this research, ***Speech to Text* produced significantly better results** than several online transcription services benchmarked, producing fewer misspellings and more coherent sentence analysis.

▷ On average, ***Speech to Text* provided a 187% productivity increase over using online transcriptions** in our benchmarks, based on six individual workflow scenarios tested.

▷ Based on all benchmarks, including manual as well as online transcriptions, ***Speech to Text* was on average almost five times faster** than other methods, and **can save hours of work** in the captions workflow.

Speech to Text Productivity: Average of All Benchmarks

With *Speech to Text* (20.48%)

Without *Speech to Text* (100%)

Chart based on the average of **18 different workflow scenarios**. A total of **185 individual benchmark measures** were taken. Reference value: Average time when working without *Speech to Text*. **Shorter is better.**

How *Speech to Text* Impacts the Premiere Pro Captions Workflow

Key Aspects of *Speech to Text*

Speech to Text **represents a major extension of the Premiere Pro captions workflow**, which itself had been significantly upgraded with the introduction of a new dedicated workspace in version 15.0. It gives user the option to transcribe text contained in one of the audio tracks; Premiere Pro uploads the audio, and processes it using Adobe Sensei-based machine learning technology. Within a few minutes the transcription is processed and can be turned into captions that line up perfectly with the audio track. *Speech to Text* **currently supports 13 languages**, including Hindi, Russian, Japanese and Chinese.
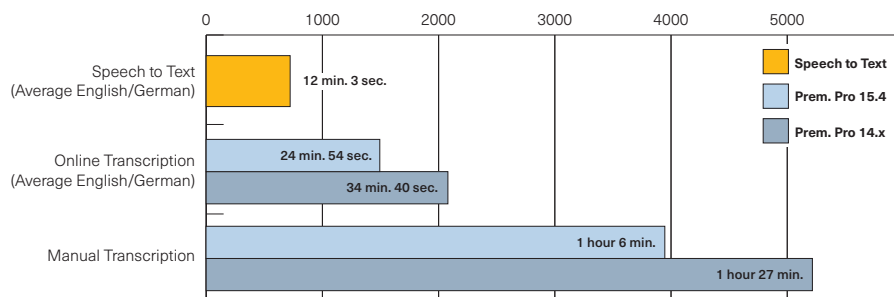
The Productivity Advantage of *Speech to Text*

Creating captions from scratch can be extremely tedious and time-consuming—manually creating captions for a 5-minute tech podcast took on average over an hour in our benchmarks. (See chart below.)

Using online transcription services to automate the creation of captions speeds things up, but *Speech to Text* **nevertheless has a significant productivity advantage**, specifically because of the quality of the transcriptions produced, which results in significantly less misspellings or problems of precise caption alignment with the audio track: in our benchmarks, using two popular online transcription services, **finalizing the caption track once the transcripts were imported took two to three times longer than when using *Speech to Text*** because of more precise caption timing and speech analysis.

Major Points

▷ *Speech to Text* **speeds up the creation of captions significantly,** and **works in 13 languages**, including Japanese, Chinese, Korean and Hindi, among others.

▷ Compared to online transcription services, *Speech to Text* **produced significantly better results** in terms of transcription and sentence analysis in our benchmarks, leading to higher productivity.

▷ The new *Captions* workspace introduced in Premiere Pro 15.0 **improves productivity of the captions workflow over older versions** of the program.

Speech to Text Benchmarks: 5-Minute Video (Average English/German)
Time-scale in seconds. All data are the average of 3 individual benchmarks
Shorter is better.

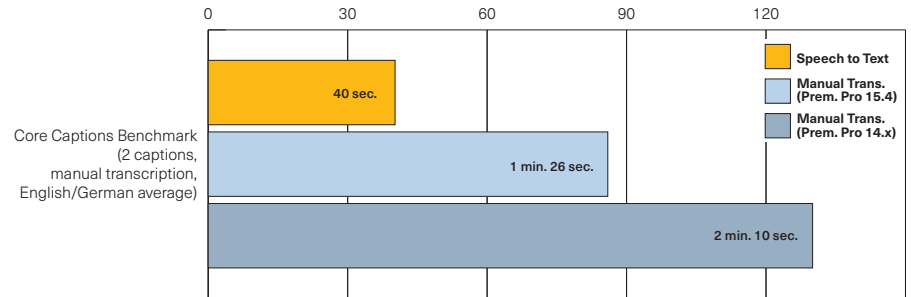|	Value
Speech to Text (Average English/German)	12 min. 3 sec.
Online Transcription (Average English/German)	24 min. 54 sec.
	34 min. 40 sec.
Manual Transcription	1 hour 6 min.
	1 hour 27 min.

Legend: Speech to Text / Prem. Pro 15.4 / Prem. Pro 14.x

For our benchmarks we tested Speech to Text *using real-world video podcasts in two languages, English and German, comparing it to two different popular transcription services.* **Results produced by** Speech to Text **were significantly better** *in terms of transcription, but also in the separation of captions into coherent sentences,* **resulting in significantly shorter time for correcting the automatically produced captions.** *The benchmarks also* **underline the productivity advantages of the new captions workflow** *introduced with Premiere Pro 15.0 that speeds up creating and fine-tuning captions significantly.*

The Impact of *Speech to Text* on Workflow Productivity: Key Benchmark Figures
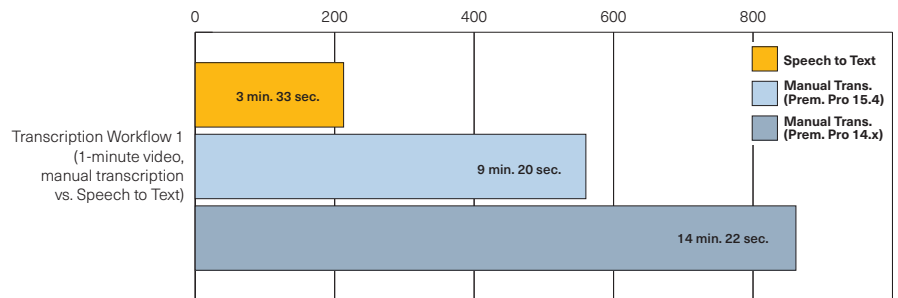
Core Captions Benchmark (Average English/German): Even for very simple operations, *Speech to Text* provides significant productivity advantages. Simply creating, typing and fine-tuning the alignement of two captions took over two minutes using the old captions workflow, **compared to 40 seconds with *Speech to Text*.** Using the new *Captions* workspace introduced in Premiere Pro 15.0 provided an over 50% productivity advantage over the previous workflow.

Speech to Text Benchmarks: Core Captions Benchmark (English/German)
Time-scale in seconds. All data are the average of 3 individual benchmarks
Shorter is better.

Core Captions Benchmark (2 captions, manual transcription, English/German average)
- Speech to Text: 40 sec.
- Manual Trans. (Prem. Pro 15.4): 1 min. 26 sec.
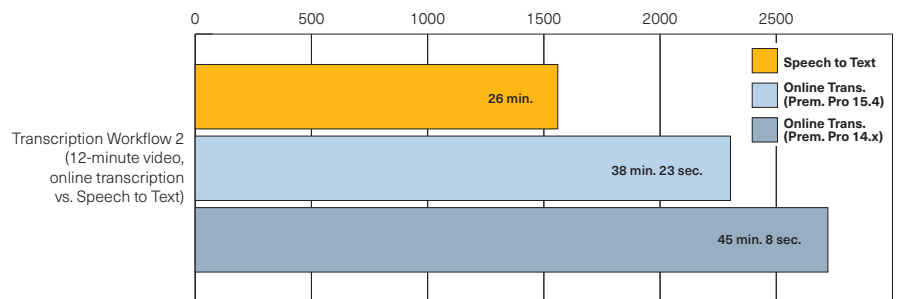- Manual Trans. (Prem. Pro 14.x): 2 min. 10 sec.

Transcription Workflow 1 (1-Minute video): Creating and fine-tuning captions for a 1-minute video manually took almost 15 minutes using the previous caption workflow, compared to just over 9 minutes using the new captions workflow. Using *Speech to Text*, the operation, including correcting and fine-tuning the automatic transcription, **took just three and a half minutes.**

Speech to Text Benchmarks: Transcription Workflow 1 (1-Minute Video)
Time-scale in seconds. All data are the average of 3 individual benchmarks
Shorter is better.

Transcription Workflow 1 (1-minute video, manual transcription vs. Speech to Text)
- Speech to Text: 3 min. 33 sec.
- Manual Trans. (Prem. Pro 15.4): 9 min. 20 sec.
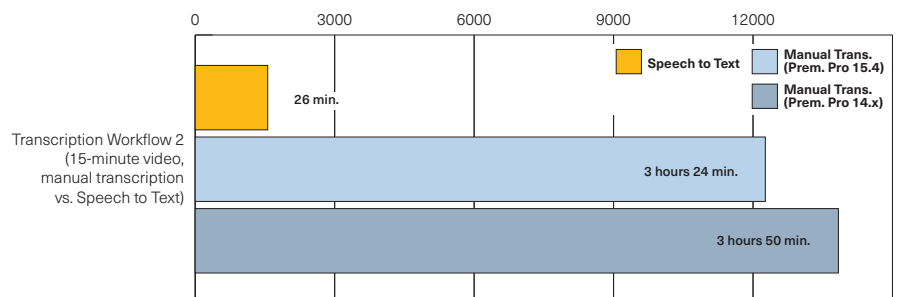- Manual Trans. (Prem. Pro 14.x): 14 min. 22 sec.

Transcription Workflow 2 (12-Minute Video, Online Transcription): For our longform benchmark, we used **a fast-paced 12-minute video tech podcast**, comparing *Speech to Text* to online transcription services. Finalizing the captions using the imported transcriptions took just over 38 minutes using the new captions workflow—and 45 minutes with the older version. *Speech to Text* saved **12 and almost 20 minutes respectively**.

Speech to Text Benchmarks: Transcription Workflow 2 (12-Minute Video, Online)
Time-scale in seconds. All data are the average of 3 individual benchmarks
Shorter is better.

Transcription Workflow 2 (12-minute video, online transcription vs. Speech to Text)
- Speech to Text: 26 min.
- Online Trans. (Prem. Pro 15.4): 38 min. 23 sec.
- Online Trans. (Prem. Pro 14.x): 45 min. 8 sec.

Transcription Workflow 2 (12-Minute Video, Manual Transcription): If one relies on manually creating captions for videos, *Speech to Text* can literally save hours: Manually transcribing the same 12-minute video tech podcast used for the previous benchmark **took between three and four hours** depending on the captions workflow used. **Creating and fine-tuning the captions took only 26 minutes using *Speech to Text.***

Speech to Text Benchmarks: Transcription Workflow 2 (12-Minute Video, Manual)
Time-scale in seconds. All data are the average of 3 individual benchmarks
Shorter is better.

Transcription Workflow 2 (15-minute video, manual transcription vs. Speech to Text)
- Speech to Text: 26 min.
- Manual Trans. (Prem. Pro 15.4): 3 hours 24 min.
- Manual Trans. (Prem. Pro 14.x): 3 hours 50 min.

Methodology

This benchmark project was commissioned by Adobe and independently executed by Pfeiffer Consulting.

All the productivity measures presented in this document are based on real-world workflow examples, designed and executed by professionals with many years of experience with these applications and workflows.

How we measure productivity

The basic approach is simple: in order to assess productivity gains that a program or solution may (or may not) bring, we start by analyzing the minimum number of steps necessary to achieve a given result in each of the applications or workflows that have to be compared.

Once this list of actions has been clearly established, we start to execute the operation or workflow in each solution, with the help of seasoned professionals who have long-standing experience in the field and with the solutions that are tested.

Every set of steps is **executed three times**, the average of the three measures is used as final result.

About Pfeiffer Consulting

Pfeiffer Consulting is an independent technology research and benchmarking operation focused on the needs of publishing, digital content production, and new media professionals.

For more information, please contact
research@pfeifferreport.com

Pfeiffer Report