

DIGITAL PRODUCTION

MAGAZIN FÜR DIGITALE MEDIENPRODUKTION

JULI | AUGUST 04:2021



Social Media

Tipps für Akquise,
Inspiration und Effizienz

Neue Tools

Cinema 4D S24, Assimilate
Live FX, Nuke 13, Flame 2022

Cooler Helfer

SyncSketch, Frame.io, KI
für Transkription und Audio

Schöne Bilder

Pixars Luca, Foo Fighters,
Circus in Town und mehr

Transcript Tools: Untertitel- und Übersetzungstools

Da Social-Media-Content überall konsumiert wird, werden Untertitel immer wichtiger. Adobe Premiere hat hier gerade einen Riesenschritt voran gemacht und das Transcript Tool in die aktuelle 2021er-Version integriert (siehe Artikel im Heft 03:21). Das Subtitle Tool ist nun intuitiv nutzbar. Doch es gibt auch viele andere Tools, die online transkribieren und sogar Texte übersetzen. Wir haben einige von ihnen getestet und verglichen und mit verschiedenen Sprachaufnahmen gefüttert, natürlich in Deutsch und Englisch und ein bisschen Französisch, Bayerisch und Pfälzisch.

von Nils Calles und Bela Beier



Nun gestehen wir erst mal, dass dieser Test nicht ganz uneigennützig ist. Alle Interviews der diesjährigen FMX fanden via Zoom und Ähnlichem statt, und die Tipperei für 20 Interviews (die ihr alle im Lauf der nächsten Ausgaben lesen könnt) motiviert, schnellere Alternativen zu finden. Aber auch wenn es scheinbar viele selbsternannte „Beste Softwares“ gibt, auf was schauen wir und was muss man bedenken, bevor man der knappen Deadline zustimmt?

Warum testen wir?

Weil Untertitel nicht nur für Redaktion, sondern auch für Delivery wichtig sind – hin und wieder wollen wir einfach eine Übersetzung/Transkription, zum Beispiel um ohne Ton zu schauen (in der Bahn, wenn es laut ist oder wenn die Review Session sich zieht) und natürlich für alle, die mit der Akustik Schwierigkeiten haben.

Natürlich könnte man das Video-File einfach auf Youtube hochladen und deren automatische Übersetzung und Transkription aktivieren, aber der ganze Prozess ist weder elegant noch effizient.

Was testen wir?

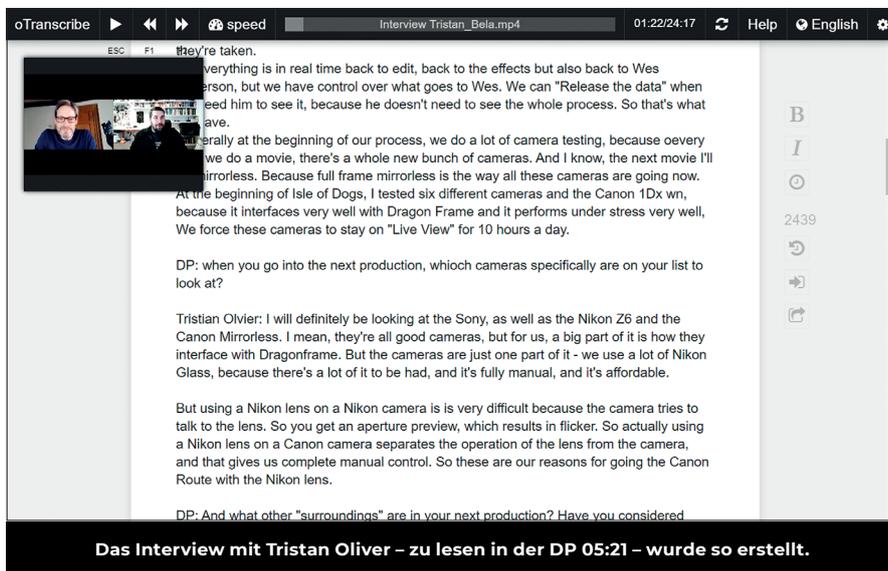
Nun gibt es natürlich viele Ansatzpunkte für Transkriptionen. Wir haben zwei davon herausgepickt: einmal der TV-Ablauf, in dem ein Redakteur Untertitel abgibt für den Beitrag. Diesen Workflow hat Nils inklusive der diversen Schnittstellen getestet. Der andere ist der reine Textworkflow, in dem aus einem

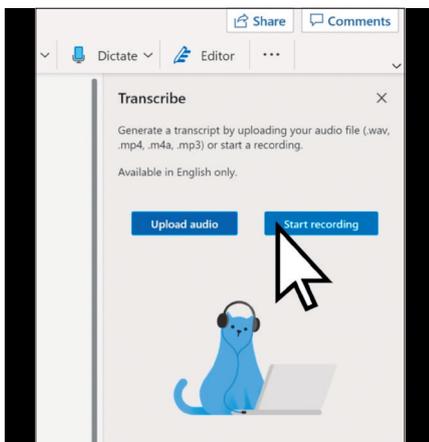
Gespräch – okay, heutzutage aus einem Zoom Call oder einer anderen Art von Online-Meeting – ein lesbarer Artikel wird, der den Inhalt zusammenfasst. Ihr kennt das aus diesem Magazin.

Hier sind auch die zwei Unterschiede: Für einen Textworkflow braucht es gute Sprechererkennung, saubere Formatierung und die Möglichkeit, spezielle Wörter zu trainieren. Für einen TV-Workflow kommt aber noch eine Menge hinzu: der Export in alle möglichen Untertitelformate, eine Frame-genaue Steuerung der Untertitel und idealerweise ein Export für das bevorzugte Schnitt-Tool. TLDR-Resultat: Simon Says macht genau das.

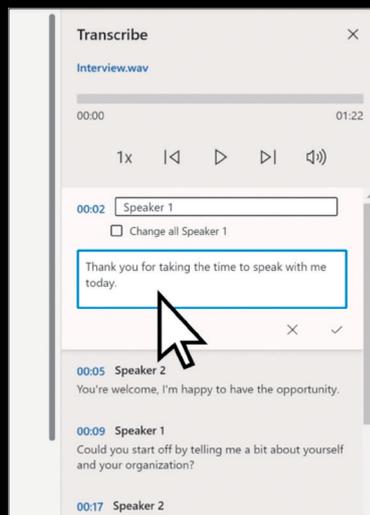
Die Referenz: oTranscribe

Als Basis legen wir den Zeitaufwand eines VFX-Redakteurs mit akzeptabler Tippgeschwindigkeit und noch schlimmerer Tippfehlerdichte in oTranscribe als Messlatte an. Wer es noch nicht kennt: oTranscribe (otranscribe.com) ist eine Webapp, in der man (kostenlos und ohne Registrierung) ein Video- oder Audio-File hochladen kann und dieses per Tastatursteuerung während des Tippens anhält, vor/zurückspulen und auch – je nach Tippgeschwindigkeit – in der Abpielgeschwindigkeit verändern kann. Per Shortcut kann man Timestamps einfügen sowie Basisformatierungen vornehmen.

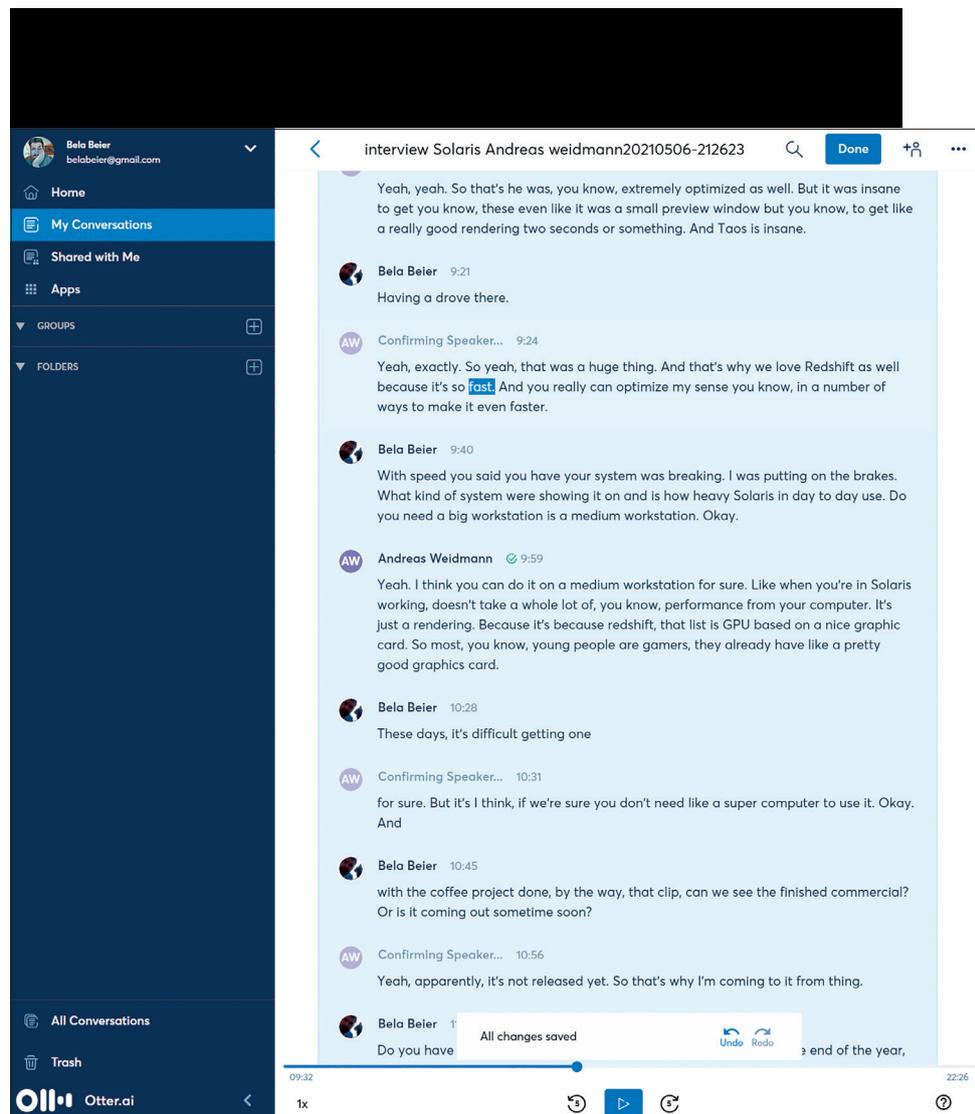




Hochladen, transkribieren. Mittlerweile auch in anderen Sprachen verfügbar, die dann angegeben werden müssen.



Sprecher können identifiziert werden, und das klappt – insbesondere bei Stereoaufnahmen – auch recht gut.



Das Interview mit Andreas Weidmann zum Thema Solaris lest ihr in der nächsten Ausgabe.

Alles nichts, was das Rad neu erfindet oder Nobelpreise gewinnt, aber nach einem Jahrzehnt Redaktionserfahrung das Tool der Wahl, da es einfach ist, sinnvoll angeordnet, keine Installation benötigt und dazu noch kostenlos verfügbar ist. Den Testtext abzutippen, hat hier 3 Minuten gedauert (was ein bisschen unfair ist, weil ich den Text ja bereits kannte). Als Daumenregel braucht man für 10 Minuten Sprachinterview – mit Sortieren und Schleifen – ungefähr eine Stunde.

Die nächste Stufe: Microsoft Office

Wer für seine Office-Tätigkeiten auf Microsoft Office setzt – insbesondere in der 365-Online-Variante – der hat vielleicht ein

Tool parat, weiß aber nicht, dass es vorhanden ist. Als zweite Option hinter „Dictate“ verbirgt sich ein Transkriptionsdienst. Wie viele Minuten Transkription enthalten sind, haben wir nicht herausgefunden – da müsst ihr in eurer Nutzerverwaltung nachsehen, es hängt vom speziellen Vertrag und Abo ab. Aber sobald man das weiß, ist es sehr einfach: Datei hochladen, kurz warten und dann den Text in ein Word-Dokument einfügen. Unser Urteil: Es ist nicht wirklich gut, aber besser als nichts. Bei idealer Tonaufzeichnung sind die Ergebnisse für Englisch und Deutsch akzeptabel, aber bei Rauschen im Ton-File setzt es aus. Auch mit Dialekten kommt Office nicht so gut klar, was wir als Diskriminierung traditioneller Lebensentwürfe sehen und verzweifelt

das hochdeutsche Wort für „Schumpa“ suchen.

Otter.ai

Einer der ersten Kandidaten, der in den Listen auftaucht ist Otter.ai. Der Dienst spezialisiert sich auf Ton- und Online-Konferenzen. Wie wir herausgefunden haben: Die Zoom-eigene Transkription ist nichts anderes als Otter. Aber auch für beispielsweise Google Meet gibt es ein Plug-in, und da es für Firmen gedacht ist, kann man es bequem mit dem Google-Kalender sowie Microsoft und Dropbox verknüpfen. Ein weiterer Vorteil ist, dass Otter sehr günstig ist: Für 13 US-Dollar bekommt man 6.000 Minuten zum Abhören pro Monat. Das lohnt dann schon.

Interview Anton Knoblach



Bild: Bernd Bundert

Zum Thema Transcript in Premiere haben wir Anton Knoblach, dem neuen Product Marketing Manager bei Adobe, ein paar Fragen gestellt.

DP: Hallo Anton, schön dass du dir für uns Zeit nimmst. Erzähl uns doch erst mal etwas über dich, deinen Background und deine Aufgaben bei Adobe.

Anton Knoblach: Ich beschäftige mich mit unserem Video-Segment im deutschsprachigen Raum. Dort bin ich die Verbindung zwischen unseren Entwicklern und den Leuten, die unsere Software tagtäglich einsetzen. Vor meiner Zeit bei Adobe war ich selbst Filmemacher und habe zwei Medienfirmen mitgegründet.

DP: Bevor nun jeder Leser hektisch die Buttons sucht: Wie weit ist Adobe bei der Entwicklung und Einführung der automatischen Transkription?

Anton Knoblach: Wir haben das Speech-to-Text-Feature letzten Herbst vorgestellt und bieten seitdem die Möglichkeit, es vorab in einer geschlossenen Beta zu testen. Wer möchte, kann sich hier anmelden (is.gd/adobe_premiere_speechtotext) und die Spracherkennung nach erfolgreicher Registrierung sogar in der Release-Version von Premiere Pro ausprobieren. Der Transkriptions-Button sitzt dann direkt im Text-Panel des neuen Untertitel-Workflows.

Einen ständigen Einblick in die Features von morgen gibt es außerdem in der öffentlichen Beta – die könnt ihr einfach in der Creative-Cloud-App parallel zum Release installieren. Ein Highlight der öffentlichen Beta ist aktuell Multithreading in After Effects, das wir Stück für Stück weiter ausbauen.

DP: Was waren die Hauptprobleme der Editoren bei der Transkription?

Anton Knoblach: Untertitel sind heute überall wichtig, manuell zu transkribieren ist allerdings ziemlich mühsam. Natürlich gibt es da Workarounds, doch unser Ziel war, den gesamten Workflow von der Transkription bis zum Export innerhalb von Premiere Pro abzudecken.

DP: In der Kooperation mit anderen Tools und Plattformen (Youtube & Co.): Welche Anforderungen werden hier gestellt für Untertitel und Transkriptionen? Und welche Dateiformate werden im Alltag eingesetzt?

Anton Knoblach: Uns ist vor allem wichtig, dass der Workflow möglichst einfach funktioniert, Profis aber trotzdem alle Möglichkeiten zur Konfiguration haben. Eine Option ist, Untertitel im eigenen Design einzubrennen. Geschlossene Untertitel wiederum können im Exportfenster konfiguriert und als SRT-File neben das Video gelegt werden.

DP: Wie ist das Feedback der Beta-Tester bis jetzt und welche Hürden seht ihr noch, bis es marktreif ist?

Anton Knoblach: Das Feedback ist bisher extrem positiv, auch für die deutsche Transkription. Trotzdem freuen wir uns weiterhin über Feedback, am besten direkt auf UserVoice. Wir möchten, dass das Feature zum Start fertig ist, und werden bis zum Release noch einiges verbessern.

DP: Und wenn wir jetzt 10 Jahre in die Zukunft gucken und raten: Welche zusätzlichen Möglichkeiten wird Premiere 2032 mit Transkription und Text/Sprache/Bild bieten?

Anton Knoblach: Natürlich fragen wir uns: Wenn eine bestimmte Infrastruktur existiert, was ist damit alles möglich? Die Verknüpfung von Editing und Text eröffnet viele spannende neue Möglichkeiten, über die wir nachdenken können. Automatische Übersetzung und textbasierter Schnitt sind zwei davon. Am Ende geht es immer darum, lästige Aufgaben einfacher zu machen und den sich ändernden Anforderungen gerecht zu werden.

Otter-Resultate

Was dabei rauskommt, ist gut verwendbar. Man muss zwar noch mal drüberschauen, was man aber bei jedem der Testkandidaten muss. Sprechererkennung, manche Fachwörter und der Umgang mit Rauschen und Nebengeräuschen ist akzeptabel. Auch der Online-Editor für die fertigen Transkripte lässt keine Wünsche offen und hat ein elegantes, einfaches Interface.

Ebenfalls toll: Für jeden Text versucht Otter schon beim Transkribieren, Schlagwörter zu erzeugen. Wer Sachen ins Web stellt, wird das zu schätzen wissen. Die Plug-ins für Zoom und Meet funktionieren identisch, und das Handling der Transkripte kann per Gruppen, Freigaben und Shares gesteuert werden. Warum ist also der Transkriptionstest nicht mit dem Satz „Verwendet Otter!“ abgeschlossen?

Die Anbindung an Videosysteme gibt es nicht. Es wird Text exportiert, und sonst nichts. Außerdem gibt es Otter nur für Englisch. Aber wer zum Beispiel eine Firma hat, in der Transkripte von Meetings Standard sind, der sollte sich das fröhliche Meeressäuger definitiv mal anschauen.

Premiere Pro

Den neuen Captions-Workflow haben wir ja schon im letzten Heft kurz vorgestellt. Nun ist das Tool auch aus der Beta-Version in die normale Version gewandert. Um zu transkribieren, wechselt man am einfachsten in den neuen Untertitel-Workspace. Dann wählt man die zu transkribierende Timeline aus und klickt auf den Transkribieren-Tab im Textfenster.

Nachdem die Sprache ausgewählt und das Transkribieren gestartet wurde, wird automatisch ein Audio-File gerendert und in die Adobe Cloud hochgeladen. Der transkribierte Text erscheint nach kurzer Wartezeit im Textfenster und kann dort direkt editiert werden. Die Timeline läuft dabei synchron zu der optisch hervorgehobenen Textstelle. Auch verschiedene Sprecher können erkannt werden.

Bei deutlich gesprochenen Texten funktioniert das Transkribieren sehr gut und auch mit englischen Fachausdrücken in einem deutschen Text kommt das Plug-in recht gut klar. Nachdem die Fehler im Transkript korrigiert sind, können die Untertitel erstellt werden. Die Navigation beim Texteditieren ist noch nicht so optimal und wird durch viele nötige Mausklicks ausgebremst. Wie es effektiver geht, sieht man bei Simon Says. Diese haben nun eine eigene Spur bekom-

Ob dieser in Hindi gesprochene Text richtig transkribiert wurde, wissen wir nicht – der Untertitel-Font fehlt, aber sonst scheint es zu passen.

men, und der Look wird über das Essential Graphics Panel gesteuert. Dazu legt man am Besten im Vorfeld schon eine Vorlage im Panel an, die man bei der Untertitelerstellung direkt auswählen kann.

Auf diese Art lässt sich die Typogestaltung sehr einfach modifizieren, und man kann schnell zwischen verschiedenen Layouts wechseln. Und hier steht die ganze Palette der Textgestaltung des Essential Graphics Panels wie Fonts, Schatten, Border und Hintergründe zur Verfügung, sodass die Titel auch an das Kunden-CI angepasst werden können. Die Untertitel landen dann schon sinnvoll in Blöcke verteilt auf der Untertitelspur oben in der Timeline. Die Blöcke lassen sich nach Bedarf dann ganz einfach zusammenfügen oder trennen. Der Text kann über das 3-Punkte-Menü oben rechts im Textfenster als reiner Text oder als SRT-File ausgegeben werden. Dieses kann dann auch zum Generieren von einblendbaren Untertiteln zum Beispiel auf Youtube verwendet werden. An Exportformaten wird alles geboten, was es an Untertitelformaten gibt. Das kann länderspezifisch sehr unterschiedlich sein. Für Sender sollte man natürlich ein Format wie MXF wählen, was die Untertitel-dateien auch verarbeiten kann.

Aktuell stehen in Premiere für die Transkription 12 Sprachen zur Auswahl: Englisch, Englisch (UK), vereinfachtes Chinesisch (Mandarin), traditionelles Chinesisch (Kantonesisch), Spanisch, Deutsch, Französisch, Japanisch, Portugiesisch, Koreanisch, Italienisch, Russisch.

Simon Says

Dieses sehr vielseitige Tool basiert auf einer Online-Plattform. Es transkribiert in über 100 Sprachen und kann außerdem im zweiten Schritt automatisch übersetzen. Es ver-

fügt über sehr viele Schnittstellen und Exportformate (siehe Kasten). Nachdem man einen Account erstellt hat, bekommt man 15 Freiminuten zum Testen. Das Transkribieren kostet im Gratis-Account dann 15 US-Dollar pro Stunde oder oder 25 Cent pro Minute.

Der Preis variiert dann in den verschiedenen Abo-Modellen. Zum Transkribieren können die Clips direkt per Drag-and-drop im Browserfenster hochgeladen werden. Es gibt aber auch eine direkte Verbindung zu den Cloud-Diensten Dropbox, Google Drive und

The screenshot shows the Simon Says web interface. On the left, a 'Keyboard Shortcuts' panel lists various controls: 'ctrl + space' for Play/Pause, 'ctrl + z' for Undo changes, 'ctrl + y' for Redo changes, 'ctrl + ←' for Go back 5 seconds, 'ctrl + →' for Go forward 5 seconds, 'ctrl + ↑' for Go to the previous transcript row, 'ctrl + ↓' for Go to the next transcript row, 'ctrl + .' for Increase playback speed, 'ctrl + ,' for Decrease playback speed, 'shift + enter' for Create new paragraph line, 'enter' for Create a new row, 'ctrl + mouse-click' for Goes to that timecode point in the audio / video, and 'ctrl + s' for Save changes. The main area shows a transcript with a list of actions: '1. Edit the text and add speaker names. It auto-saves every 2 mins.', '2. Bookmark and annotate key parts of the transcript.', '3. sync timecode.', '4. Translate the transcript. Export to the Visual Subtitle Editor, Export to your video editing app, Word, subtitles, and many other formats.', and '5. invite team members.' Below the transcript, there is a list of video clips with their timecodes and descriptions of the scenes.

Die Tastatursteuerung ist recht durchdacht. Während der Cursor im Textfenster verbleibt, kann man mit gedrückter Steuerungstaste durch das Video und den Text navigieren. Mit Strg+Mausklick in den Text springt das Video direkt an die Textstelle.

The screenshot shows the Ledcave Team Edit interface. It features a multi-track editing workspace with three video tracks labeled 'DJI_0401.MP4', 'DJI_0408.MP4', and 'DJI_0400.MP4'. Each track has a search bar and a list of clips with their timecodes and descriptions. A central text box explains: 'Es werden bis zu 2 Clips als Source-Monitore geladen. Nachdem man die gewünschte Textpassage ausgewählt hat, kann diese mit einem Mausclick in den Record-Monitor geschnitten werden. Dort kann die Reihenfolge der Schnitte dann noch einfach per Drag-and-drop verschoben werden.' On the right, there is a 'Frame Rate' panel showing 'Resolution' (Height: 2160, Width: 3840) and 'Start Timecode' (00:00:00:00). The interface is dark-themed and includes various editing tools and a search bar.

Es werden bis zu 2 Clips als Source-Monitore geladen. Nachdem man die gewünschte Textpassage ausgewählt hat, kann diese mit einem Mausclick in den Record-Monitor geschnitten werden. Dort kann die Reihenfolge der Schnitte dann noch einfach per Drag-and-drop verschoben werden.

Box. Das erleichtert die Zusammenarbeit, denn Projekte können auch geteilt und im Team bearbeitet werden. Nach dem Upload muss man noch die Sprache auswählen und kann dann die Transkription starten. Ist das Transkript fertig, bekommt man das in einer Mail mitgeteilt. Der Text kann dann direkt im Browser editiert werden. Auch hier wird der hervorgehobene Text mit dem Videofenster synchronisiert.

Natürlich funktioniert auch hier die Transkription nicht perfekt. Doch der Text lässt sich in Verbindung mit dem Videofenster und der Tastatursteuerung gut editieren. Es gibt hierbei sogar eine Version History, man kann also auch eine ältere Version wiederherstellen. Nach der Korrektur kann man einige Optionen nutzen, um den Text für andere Anwendungen zu exportieren. Dazu gehören MS Word und Text und diverse Untertitelformate sowie XML- und EDL-Formate für die gängigen Schnittsysteme. Was den Workflow etwas ausbremst, ist, dass Text- und Word-Dateien immer als ZIP heruntergeladen werden, aber in der gezippten Datei noch nicht mal der Name übernommen wird. Was Simons Says außerdem kann, ist, den Text automatisch in über 100 Sprachen zu übersetzen. Der Clou ist allerdings der neue Assemble-Modus. Man kann hier durch Markieren von Text das Video online im Browser schneiden.

Das Video-Editieren geht genau wie das Text-Editieren im Team interaktiv in Echtzeit, ähnlich wie in Google Docs. Dazu muss nur der Link geteilt oder eine Einladung verschickt werden. Wer im Team editieren will, muss sich zumindest einen kostenlosen Account einrichten. Der fertige Rough-Cut kann

Simon Says Exportformate

Es sind möglich: Microsoft Word, Text-Dokumente, Adobe Premiere Pro, Adobe Audition, Final Cut Pro X, Final Cut Pro 7, Grass Valley Edius, Avid Pro Tools, Avid Media Composer via Text Locators, ScriptSync und AF, Blackmagic Designs DaVinci Resolve (Mac only) oder über XML, SAMI Subtitles/Captions, SRT Subtitles/Captions, Avid STL/SubCap Subtitles/Captions, WebVTT Subtitles/Captions, CSV/Microsoft Excel files, JSON, MAXQDA, Burned-in Subtitles/Timecodes

Wer lieber direkt in seiner Software arbeitet, findet Plug-ins und Integrationen für Adobe Premiere und Audition, Final Cut Pro, Avid und Resolve (Mac only).

als Schnittliste im EDL- oder XML-Format exportiert werden. Das funktioniert auch mit dem Gratis-Account. Kosten entstehen nur für Transkription und Übersetzung.

Die beste Integration von Simon Says gibt es in Adobe Premiere. Die kostenlose Erweiterung ist direkt in den Workspace integriert und kann als Fenster im Workspace angedockt werden. Um ein Transkript zu erstellen, legt man mit dem Plus-Symbol ein neues Projekt an und importiert die Clips oder die Sequenz, indem man diese im Projektfenster auswählt und auf „Import“ klickt. Dann werden die Clips automatisch mit dem Media Encoder in ein passendes Low-Res-Format transcodiert und hochgeladen. Ist dies geschehen, muss die Sprache ausgewählt werden, und nach einem Klick auf den

Assemble Edit für Simon Says

Das Webinterface eignet sich ideal zur Zusammenarbeit zwischen Kameramann, Redakteur und Editor.

Dabei ist folgender Workflow für das Remote-Editing denkbar: Irgendwo auf diesem Planeten dreht ein Kameramann eine wichtige Rede von einem ebenso wichtigen Menschen. Die relevanten Passagen sollen von einem Redakteur im Sender ausgesucht werden und dann vor Ort von einem Editor zu einem kurzen Nachrichten-Clip zusammengeschnitten werden.

Wenn man mit der Premiere-Erweiterung arbeitet, ist es recht einfach, da die Proxies automatisch über den Media Encoder erstellt werden. Dabei benutzt Simon Says eine Auflösung von 480 x 360 Pixeln mit 0,3 Mbit und AAC Audio in 48 kHz mit 128 Kbit in Mono. Doch es funktioniert eigentlich mit jeder Schnittsoftware, die EDL- oder XML-Formate importieren kann.

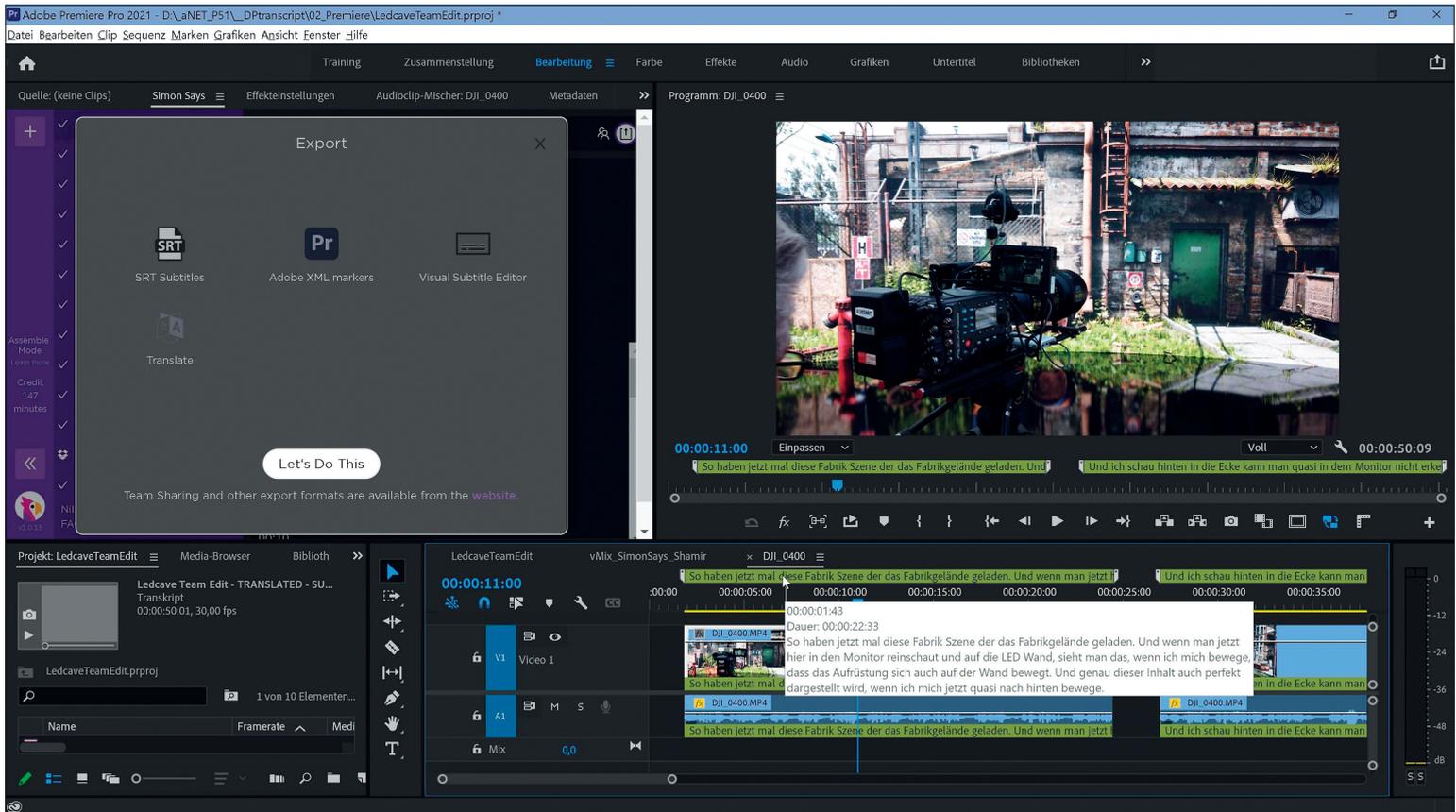
Zunächst wird das Rohmaterial in ein möglichst kleines datensparendes Videoformat konvertiert. Ob man das mit dem Adobe Media Encoder, dem kostenlosen Handbrake oder als Batch Encoding zum Beispiel in Kyno macht, ist eigentlich egal. Allerdings sollte die Audioqualität im Vordergrund stehen, da diese ausschlaggebend für eine akkurate Transkription ist. Die Dateinamen sollten idealerweise beibehalten werden. Die Videoclips werden dann ins Portal hochgeladen und transkribiert.

Der Redakteur bekommt den Teamlink und kann den Text editieren und den gewünschten Text im Assemble Mode aneinanderschneiden. Die EDL wird dann ins Schnittsystem vor Ort geladen und die Original-Clips relinkt. Danach kann der in Broadcast-Qualität rausgerenderte Zusammenschnitt per File-Transfer zum Sender geschickt werden.

Wichtig ist, dass vor dem XML-Export die Projekt-Settings angepasst werden. Da man ja mit niedrig aufgelösten Proxies arbeitet, muss das Projekt am Ende für einen 4K-Edit auch auf 4K eingestellt werden, da sonst das Bild an den Rändern abgeschnitten wird.

Nach einem Klick auf die Uhr öffnen sich die Projekt-Settings. Angezeigt wird hier natürlich die Bildgröße der hochgeladenen Proxies. Da die Original-Clips in 4K vorliegen, müssten wir hier logischerweise 3840 x 2160 Pixel einstellen.

Transcribe-Button erscheinen der Clip und der Text nach einigen Minuten im Fenster. Dort können das Video direkt abgespielt und der Text editiert werden. Parallel dazu erscheint das Projekt auch im Browser-Interface. Leider funktionieren die praktischen Tastaturkommandos im Erweiterungsfenster nicht so wie im Browser, sodass hier das Editieren wesentlich mehr Mausklicks erfordert. Deshalb macht es mehr Sinn, den Text im Browser zu editieren. Dieser wird dann auch im Fenster aktualisiert, indem man kurz ein anderes Projekt anklickt und dann wieder das aktuelle.



Hier sieht man, wie die mit Simon Says generierten Marker in Premiere aussehen.

Man kann aber in Premiere dann über den Upload-Button oben rechts direkt SRT-Subtitles, XML-Marker exportieren oder den Visual Subtitle Editor oder die Übersetzung starten. Da die Übersetzung textbasiert funktioniert, geht diese wirklich

sehr schnell. Danach kann daraus wieder ein SRT-File erstellt werden, das in Verbindung mit dem neuen Subtitle Workflow in Premiere auch direkt wieder in die entsprechende Timeline gezogen werden kann. Die im Interview (siehe Kasten) erwähnte

datensichere Transkription ist allerdings etwas teurer. Sie kostet 100 US-Dollar für 5 und 2.000 US-Dollar für 100 Stunden. Das sollte aber bei den entsprechenden Kunden kein Problem sein. Da das Tool im Browser läuft, funktioniert es auch auf dem Handy

trint < Interview_mit_Bela_Beier_und_Claus_Pfeifer.m4a [Edit](#) B

Find & Replace Undo Redo Highlight Strike Add Marker Comment Translate Add to Story Export Share

Speaker 2 00:03:36
Arche über die Tür zu. Wird beide Katze nicht reinkommt.

Speaker 1 00:03:39
Avni Tütü dann wird das Geschrei richtig laut. Naja. Also hier diese ganzen stolz sachen von emo Kavallerist und wie sie alle heißen. Zusammen mit ihnen den modest. Die Ka ist ja ja eigentlich auch. Aber die würden ja im prinzip angenommen. Ich kann letztes des Kamera ist Decay an mein Auto ist Joker oder sowas anknüpfen. Dann wäre ja das im Prinzip eine komplette er ist kommt ne komplette Filmpartner, weil ich schon fast mit der ich wirklich von

⏪ ▶ ⏩
Playback speed x1.0
Timecode [Adjust](#)
Total duration: 00:22:43
Highlights duration: 00:00:00
Paragraphs verified ✔ 0 / 43 (0%)

Der Editor ist schick – aber wer keinen von Experten aufgenommenen Ton hat, und das Vokabular trainiert, wird wenig Erleichterung verspüren.

oder Tablet. Auf meinem Samsung Note 20 Ultra kann ich durchaus arbeiten. Dabei ist zu beachten, dass das Videofenster nur sichtbar ist, wenn das Gerät quer gehalten wird. Auf meinem betagten iPad Air der ersten Generation funktioniert es allerdings nur sehr hakelig, und in der eigentlich für das iPhone programmierten App kann ich den Text nicht editieren.

Abschließend ist zu sagen, dass die Transkription in Simon Says nicht ganz so gut funktioniert wie in Premiere. Ideal wäre es deshalb, die kostenlose Transkription mit Premiere direkt zu machen und dann eine bezahlte Übersetzung mit SRT-Export mit Simon Says. Doch leider habe ich bisher keine Möglichkeit gefunden, Text und SRT in Simon Says zu importieren. Und Copy-and-paste in den Text-Editor ist dann doch etwas zu mühsam.

Trint

Ein weiterer Anbieter, der nicht zwangsläufig in den Videobereich zielt, ist Trint ([Trint.com](https://trint.com)). Für 55 Euro pro Monat kann man 7 Files transkribieren, wer unbegrenzte Transkriptionen will, muss 65 Euro zahlen. Übersetzung ist dabei enthalten. Die Menge wird mit einem nicht genauer definierten „Fair Use“ beschrieben für den „unbegrenzten“ Umfang. Wie viel das ist, wissen wir nicht – hier also aufpassen!

Im Test hat die Transkription eines 22-Minuten-Interviews gerade mal 11 Minuten gedauert – auf Deutsch. Ein Zitat aus dem aufgenommenen Zoom Call ohne Nebengeräusche: „Auch theoretisch extrem viel kann ich Courtaige von Bruno die Katze grade fünf Minuten hat und dann immer noch so gut die eigentliche Arche über die Tür zu. Avni Tütü dann wird das Geschrei richtig laut. Najä. Also hier diese ganzen stolz sachen von emo Kavallerist. Wo ist das? Sydney oder Melbourne machen, marodieren.“ So viel dazu.

Sonix

Der Letzte, den wir praktisch getestet haben, hört auf den Namen Sonix (<https://sonix.ai>) und war lustigerweise einer der Ersten, den wir schon vor dem Test kannten. Die Transkription kostet 10 US-Dollar pro Stunde, und die automatisierte Übersetzung hat denselben Preis, wird aber mit „Automated translations are provided 'as is', are not perfect, and may require additional editing“ angeboten. Wer mehr braucht, kann den Stundenpreis auf 5 US-Dollar reduzieren und dafür 22 US-Dollar Gebühr pro User und Monat zahlen.

Nichtsdestotrotz funktioniert alles – insbesondere bei Englisch – überraschend flüssig, und auch Fachbegriffe wie zum Beispiel

Interview Shamir Allibhai



Ich habe mit Shamir Allibhai, dem Gründer von Simon Says, ein Interview geführt und ihm einige Fragen gestellt. Hier sind die interessantesten Antworten:

DP: Wenn ich in einem Transkript ein Wort korrigiert habe, merkt sich die AI-Software die Korrektur und korrigiert dann beim nächsten Mal automatisch?

Shamir: Nein, das funktioniert leider nicht. Aber man kann sich eigene Dictionaries anlegen, in denen die Worte und Namen eingetragen werden, die in einem speziellen Kontext oder Interview verwendet werden. Das sagt dann der AI: Wenn du dieses Phonem hörst, dann nimm diese Schreibweise.

DP: Gibt es Unterschiede zwischen den exportierbaren XML-Formaten für die unterschiedlichen Schnittprogramme?

Shamir: Ja, wir exportieren Custom XML-Formate für jede Schnittsoftware. Denn jede Software kann andere Informationen verarbeiten. Die Final-Cut-7-XML ist zum Beispiel über 10 Jahre alt und unterstützt kein Multicam. FCP-10-XML unterstützen Multicam, und im Prinzip kann man in Simon Says Assemble eine Multicam Schnittliste ausgeben. Das unterstützt die Adobe-XML leider noch nicht.

DP: Was ist mit der On Premise Transcription, wie funktioniert das? Das ist ja für Firmen gedacht, bei denen wegen der Geheimhaltung keine Daten ins Internet übertragen werden dürfen?

Shamir: Da wird einfach die komplette Software heruntergeladen und die AI passiert auf dem lokalen Rechner. Man kann dann alle Internetverbindungen kappen und es funktioniert trotzdem.

DP: Sind es dann nicht sehr große Datenbanken, die heruntergeladen werden müssen, um die AI zu füttern?

Shamir: Das kommt auf die benötigten Sprachen an. Wir empfehlen mindestens 20 Gbyte freien Speicherplatz, was aber auf einem Videoschnitt-Rechner nicht wirklich ins Gewicht fällt.

DP: Was ist in Zukunft an Extensions geplant?

Shamir: Wir haben ja schon eine Extension für Resolve auf dem Mac. Die Extension für die PC-Version ist gerade in Arbeit, und dann werden wir sehen, was wir uns als Nächstes vornehmen.

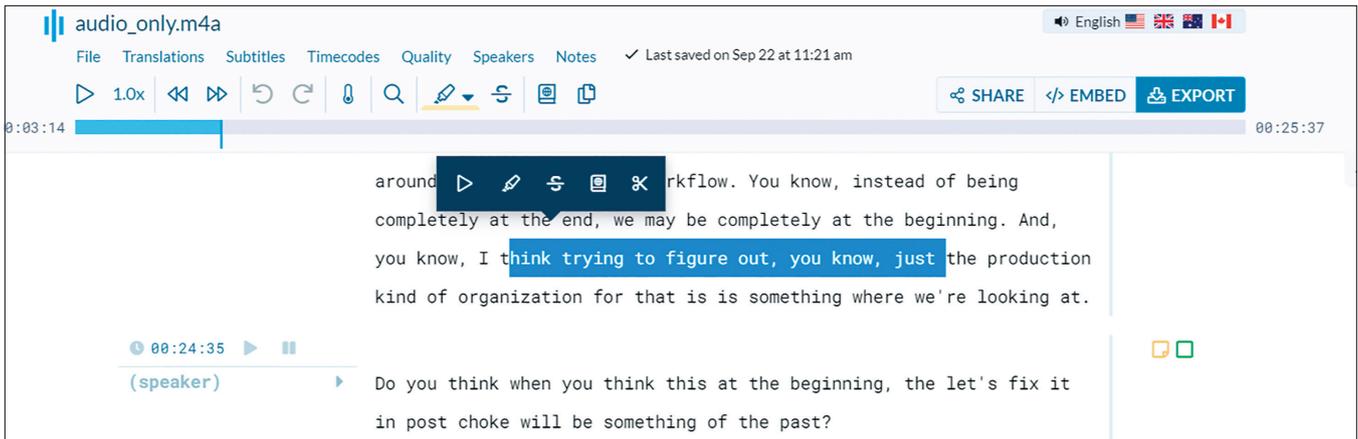
Maya, Bifrost, Emitter und Ähnliches wurden im Test-Text problemlos erkannt. Integration via Zapier ist auch möglich und so der Einbau in Pipelines. Dazu gibt es eine wohl dokumentierte API – ein Argument, wenn der TD fit ist. Zoom und eine Chrome-Extension gibt es außerdem. Was noch praktisch ist: Ein „Transcriptionist Directory“ erlaubt, in Stresszeiten einen Menschen für das Clean-up der Texte hinzubuchen, der dann im System fit ist. Im Fazit nicht ganz günstig, aber voller Features, die für Nutzer mit großen Umfang sinnvoll sind.

Fazit

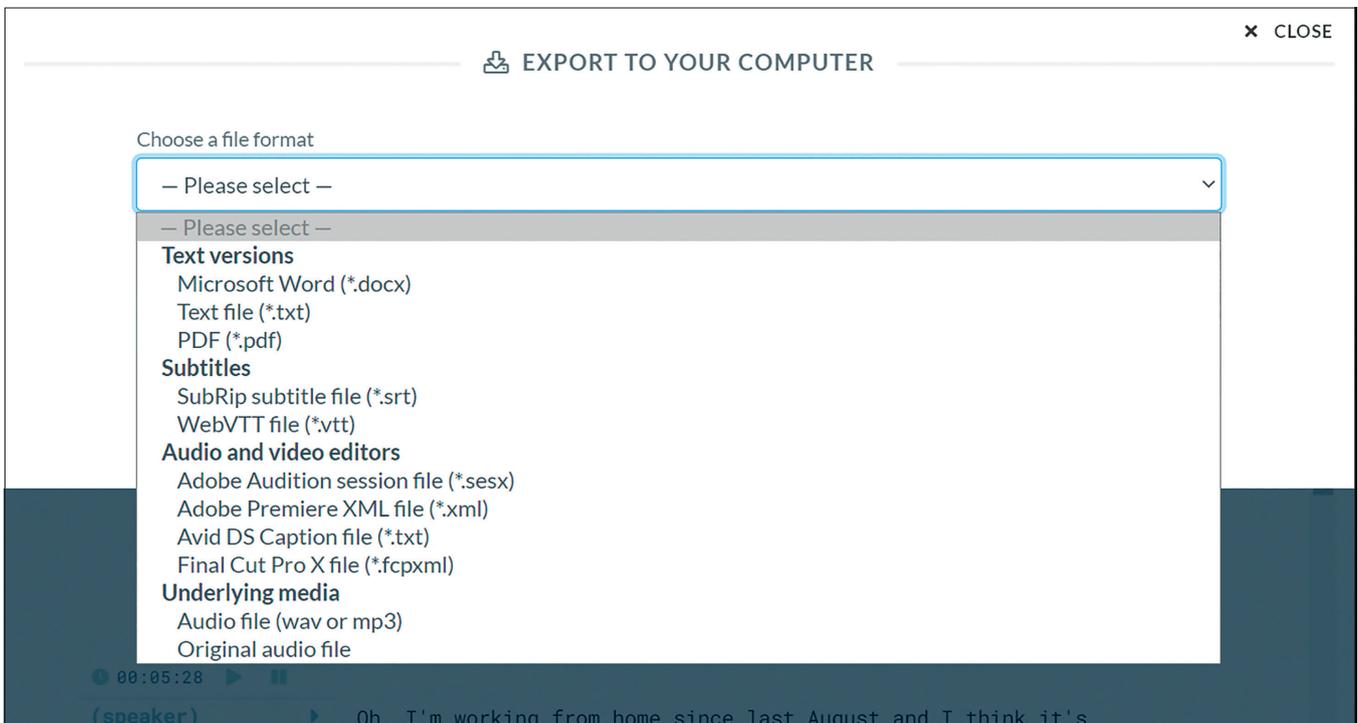
Natürlich ist die Frage, wer welches Tool für welche Pipeline braucht. Wer nur einmal

etwas Kurzes machen will, wird mit Sonix gut bedient sein, und wer sich auf englische Sprache beschränken kann, ist mit Otter.ai gut versorgt. In einem Video-Workflow hat Simon Says klar die Nase vorn – zumindest bis wir alle die Premiere-eigenen Tools verwenden können.

Eine Warnung möchten wir aber noch aussprechen: Es gibt weitere Dienste, die aber stellenweise nur auf andere aufsetzen und dann dementsprechend teuer werden können. So wie die Zoom-Transkription nichts anderes als ein Interface für Otter.ai ist, haben wir auch Dutzende von angeblich spezialisierten Diensten gefunden, die nichts tun, als ein paar Prozent vom Preis abzuschöpfen – hier lohnt also der Vergleich. Alle Anbieter haben Test-Accounts – verwendet



Der Sonix Editor ist kompakt, aber bietet viele Features – die Steuerung des Textes ist im Umgang flüssig und auch komplett per Shortcuts steuerbar. Die Notiz und „Als erledigt markieren“-Funktion auf der rechten Seite haben wir so bei keinem Konkurrenten gefunden.



Hier sind wieder die Untertitelooptionen – die wichtigsten Formate werden mit nur einem Klick unterstützt. Im Premium-Paket kann man Farben, Fonts und Position anpassen sowie Burn-in aktivieren und alles am Timecode ausrichten.

diese, bevor ihr in einer Abo-Falle hängt. Bei den kostenlosen Diensten ist für Videos der zugegebene holprige Workflow über Youtube nicht schlecht, und wenn der Ton wirklich miserabel ist, dann hilft nur noch oTranscribe und ein menschliches Ohr, bevorzugt mit guten Kopfhörern.

Aber hier steht die Entwicklung nicht still. Sogar in unserem Testzeitraum haben

die Dienste dazugelernt, und wenn wir in die Glaskugel schauen, wird es in ein paar Jahren mit einem Klick erledigt sein. Bis dahin hoffen wir, dass euch unser Test hilft. Und wer sich die diversen Testschnipsel ansehen will, kann sich die Clips kostenlos auf digitalproduction.com herunterladen, um Nils und Bela beim pfälzern/bayern zu sehen und zu hören. Mit Untertiteln, natürlich. > ei



Nils Calles betreibt eine Medienagentur im Musikpark 2 im Kreativviertel Jungbusch in Mannheim, wo er schwerpunktmäßig Trailer, Motion Graphics, Imagefilme, Magazinsendungen im virtuellen Studio, Tutorials und 360-Grad-Webseiten produziert. Des Weiteren ist er Producer für LinkedIn Learning und Trainer und Consultant für Workflows und Software im Postproduktionsbereich. Er ist außerdem seit über 25 Jahren für verschiedene Produktionen und Sender weltweit als Cutter, Producer und Compositor auf Messen und im Ü-Wagen unterwegs und hat die Seite <https://postproduction-tutorials.net/> mit ausgewählten Video-Tutorials online gestellt. Hier hat er auch, passend zu diesem Artikel, einige gute Beiträge und Links gesammelt.

Bild: Thommy Mardo



AMD
THREADRIPPER
PRO

AMD Ryzen™ Threadripper™ PRO Processors: Creation Without Limits

Say goodbye to bottlenecks. Say goodbye to slowdowns. Say goodbye to limits.

With AMD Ryzen™ Threadripper™ PRO processors, enjoy more power, performance, and possibility than you ever thought possible.

Offering up to 64 cores and 128 threads, a total L3 Cache of up to 256 Mb, and a Max Boost Clock of up to 4.2GHz, the only limit is your imagination.

Learn more: (<https://why-ryzen.com/uk/threadripper-pro>)