

	aTrain	Voice AI	NoScribe
Overview	- free, can be used offline - audio data does not have to be uploaded to internet - tool can be used without registration - transcription in 57 languages possible	- free - tool can be used online - transcription in 57 languages possible	- free, can be used offline - audio data does not have to be uploaded to the internet - tool can be used without registration - transcription in 99 languages possible
Download	- download via Microsoft store or website of Uni Graz Mac: https://business-analytics.uni-graz.at/de/forschung/atrain/download/ Microsoft: https://apps.microsoft.com/detail/atrain/9N15Q44SZNS2	- reachable via Academic Cloud Academic Cloud: https://academiccloud.de/services/VoiceAI/	- Download via website (Links for Mac and Windows) https://github.com/kaixx/noScribe
Handling	- easy to use, well-arranged desktop - clean desktop - upload data → processing → result	- easy to use, well-arranged desktop - clean desktop - upload data → choose text format → processing → result	- easy to use, well-arranged desktop - upload data → processing → result
Transcription speed	- around three times of interview length (if used on up-to-date-device)	- faster transcription if speakers are easier to understand - average: 25 words/minute (just to clear it up, there were totally different lengths with different videos, I built an average –	- around two to five times of interview length (depending on the device the tool is used on)

		one example had 32 words/minute another: 19 words/minute)	
Quality of transcription	- problems with speaker identification - problems in recognition what is said - transcription with timestamps possible - no recognition of language changes - pauses are not dealt with - timestamps can be integrated in MAXQDA	- detailed, most reliable transcription - recognizes word duplications - no recognition of changes in language - sometimes (but not always) mistakes endings (e.g: “you” instead of “you’re”) - pauses can be calculated by looking at timestamps - transcriptions are available in different formats (normal text, SRT, VTT) - transcription with timestamps possible	- problems with speaker identification - reliable transcription - transcription with time-stamps possible - no recognition of language changes - pauses can be considered -
Speaker recognition	- speaker recognition available - automatic or manual speaker recognition - difficulties concerning overlapping speech → nearly no transcription if audio contains overlapping speech	- no speaker identification - can deal with overlapping speech	- speaker recognition available - automatic or manual speaker recognition - recognizes overlapping speech (extra tool to observe possible overlapping speech)
Criticism /summary	+ easy to handle + fast transcription - no precise transcription → giving a simple idea of what is said - problems with speaker recognition	+ most reliable transcription + option to translate transcriptions into English language + access to previous transcriptions in the program - <i>biggest disadvantage</i>: transcription speed and no speaker recognition	+ easy to handle + fast and reliable transcription + menu items to specify individually what you are looking for in the transcription

	→good to get a simple overview	- big data is difficult to upload (takes long or software crashes) - slowest tool	- partly flawed transcription
--	--------------------------------	--	---

→ no matter which AI-tool, it can only be used as a support not as reliable source for transcription → own work is needed!

→ the better the audio quality and figure of speech, the better the transcription

	aTrain	Voice AI	NoScribe
Überblick	- kostenlos - Nutzung offline möglich - Dateien müssen nicht ins Internet hochgeladen werden - Nutzung ohne Registrierung möglich - Transkription in 57 Sprachen	- kostenlos - Nutzung online möglich - Transkription in 57 Sprachen	- kostenlos - Nutzung offline möglich - Dateien müssen nicht ins Internet hochgeladen werden - Transkription in 99 Sprachen
Download	- Download über Microsoft Store oder Webseite der Uni Graz Mac: https://business-analytics.uni-graz.at/de/forschung/atrain/download/ Microsoft: https://apps.microsoft.com/detail/atrain/9N15Q44SZNS	- erreichbar über Academic Cloud → Zugriff über „Tools“ Academic Cloud: https://academiccloud.de/services/ Voice AI: https://voice-ai.academiccloud.de	- Download über Webseite (Links für Mac und Windows) https://github.com/kaixxx/noScribe
Bedienung	- einfache Handhabung - selbsterklärende Nutzoberfläche - Hochladen der Daten → Datenverarbeitung → Ergebnisse	- einfache Handhabung - selbsterklärende Nutzoberfläche - Hochladen der Daten → Auswahl des Textformats → Datenverarbeitung → Ergebnisse	- einfache Handhabung - selbsterklärende Nutzoberfläche - Hochladen der Daten → Datenverarbeitung → Ergebnisse
Geschwindigkeit	- etwa dreifache Interviewdauer (auf aktuellen Geräten)	- schnellere Spracherkennung, wenn Gesprochenes leichter zu verstehen ist - Durchschnitt: 25,5 Wörter pro Minute	- etwa zwei- bis Fünffache Interviewdauer (kommt auf Leistung des Geräts an)

		(um es klarzustellen; verschiedene Proben haben sehr verschiedene Ergebnisse gebracht, ich habe einen Durchschnitt gebildet Ein Probevideo hatte 19 Wörter/Minute Anderes Video 32 Wörter/Minute	
Qualität der Transkription	- Probleme mit der Sprecheridentifikation - Probleme mit allgemeiner Spracherkennung - Transkription mit Zeitmarken möglich - kein Erkennen von Sprachwechseln - keine Berücksichtigung von Pausen - Zeitmarken können bei MAXQDA integriert werden	- detaillierte und verlässlichste Transkription - erkennt Dopplungen - erkennt teilweise Endungen (z.B.: „you“ anstatt „you’re“) - Pausen können berechnet werden über Zeitmarken - Transkription in verschiedenen Dateiformaten erhältlich (normaler Text, SRT, VTT) - Transkription mit Zeitmarken möglich	- Probleme mit der Sprecheridentifikation - Zuverlässige Spracherkennung - Transkription mit Zeitmarken möglich - Kein Erkennen von Sprachwechseln - Berücksichtigung von Pausen möglich
Sprechererkennung	- automatische oder manuelle Sprechererkennung - Schwierigkeiten bei überlappenden Aussagen	- keine Sprechererkennung - guter Umgang mit überlappenden Aussagen	- automatische oder manuelle Sprechererkennung - erkennt auch überlappende Aussagen (extra Tool als Auswahlmöglichkeit)
Kritik/Fazit	+ einfache Benutzung	+ zuverlässigste Transkription + Möglichkeit, Transkribiertes zusätzlich von Originalsprache ins Englische zu übersetzen	+ einfache Benutzung + relativ präzise Transkription + schnelle Ergebnisse

	- keine präzise Transkription --> liefert nur eine Idee dessen, was gesagt wurde - fehlerhafte Sprechererkennung → nützlich für ersten Überblick	+ Zugriff auf frühere Transkriptionen innerhalb des Programms - <i>größter Nachteil</i>: keine Sprechererkennung und Dauer der Transkription - Probleme bei großen Dateien (Upload-Probleme)	+ Auswahlmöglichkeiten zur individuellen Anpassung dessen, was gewünscht ist - fehlerhafte Sprechererkennung
--	---	--	---

→ unabhängig davon, welches Tool genutzt wird, ist eine eigenständige Kontrolle und Überarbeitung der Transkriptionen notwendig

→ je besser die Audioqualität und das Sprachbild der Sprecher*innen, desto besser ist auch die Transkription