

Handleiding FA-converter

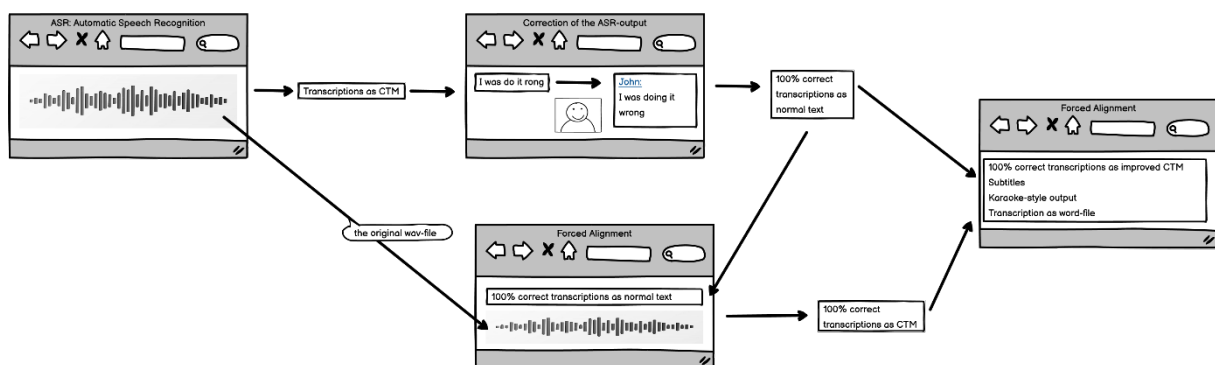
Voor verkrijgen:

- ondertiteling: .srt file
- karaoke: .html file
- transcript, .doc file

Samenvatting

Deze handleiding voor het programma FA converter (Forced Alignment converter) is bedoeld om gecorrigeerde ASR transcripten die via het programma *Forced Alignment 2* op de server van de Universiteit van Nijmegen zijn opgelijnd (<https://webservices.cls.ru.nl>), uiteindelijk in correcte ondertiteling, 'karaoke' -functie en goed leesbaar transcript om te zetten.

Het schema ziet er als volgt uit:



Figuur 1: Links hebben we de originele invoer van de geluidsfile (wav) die door de ASR-engine omgezet wordt in transcripties. Deze transcripties (die nooit 100% correct zullen zijn) worden vervolgens door mensen gecorrigeerd, van leestekens voorzien en er wordt aangegeven wie wanneer sprak. De tekstuele resultaten worden vervolgens, samen met de originele wav-file weer naar de server in Nijmegen gestuurd alwaar ze opgelijnd worden. Het resultaat is weer een CTM-file maar nu eentje die 100% correct is. Om de uiteindelijk resultaten te krijgen wordt deze CTM-file gecombineerd met de door de gebruiker gecorrigeerde tekst. Op deze manier worden ondertitels, karaoke-file en een leesbaar word-file gemaakt waarin bij de eerste twee nog steeds de link met de oorspronkelijke audio bewaard is gebleven. Maar de ondertitels en Karaoke-file bevatten nu hoofdletters en leestekens.

Invoer

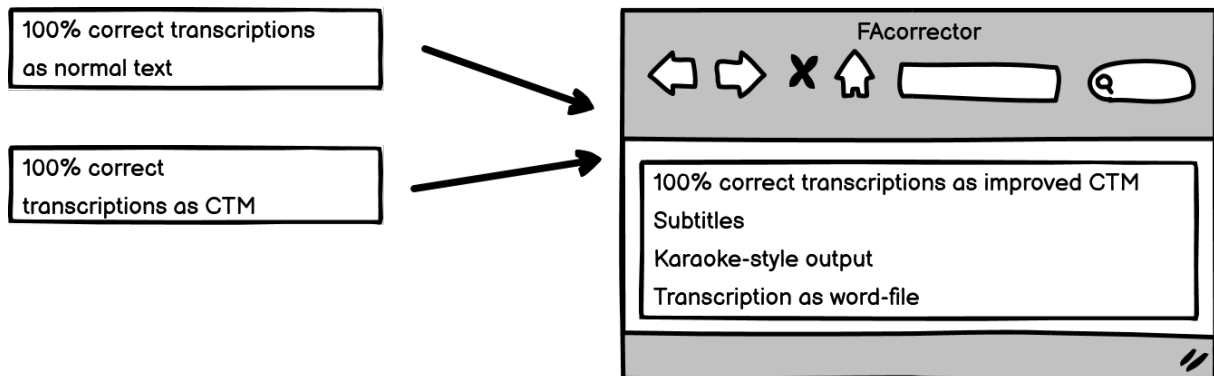
Open het open source programma [FAConverter](https://speechandtech.eu), download van: <https://speechandtech.eu>

De invoer bestaat uit achtereenvolgens twee tekstfiles die in dezelfde folder staan met de wav file, allen met dezelfde *root name*.

1. Klik op 'Open' uit hoofdmenu en klik op 'CTM / SPK';
2. Selecteer de nieuwe CTM-file op je computer, na oplijnen via het programma [Forced Alignment 2](#),
3. Klik nogmaals op 'Open' uit hoofdmenu en klik op 'Tekst'
4. Selecteer de correcte TXT-file, gecorrigeerd met [ASR Corrector](#) (check voor nieuwste versie)

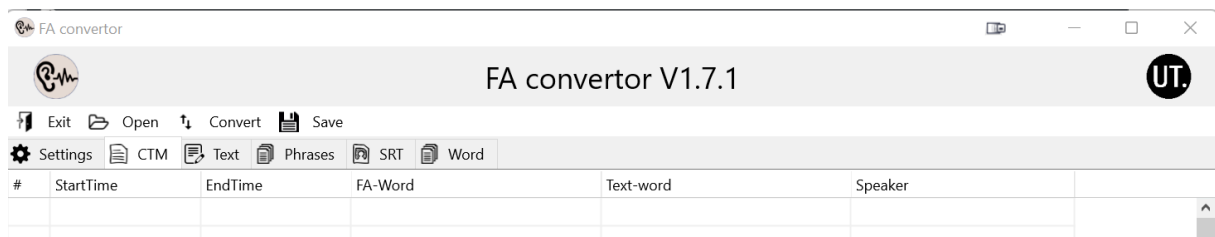
Uitleg:

De CTM-file bevat alle woorden maar zonder leestekens, hoofdletters of sprekers maar heeft wel van elk woord de starttijd en de duur. De tweede (een txt-file) heeft wel leestekens, correct geschreven woorden en wie er aan het woord is, maar mist de verbinding met de geluidsgolf.



Figuur 2: Invoer voor FAconvertor. de CTM-file en de correcte tekst.

Door deze twee tekstfiles te combineren, kunnen we een ondertitelings-file, een Karaoke-file en een leesbaar transcript in een MS-Word file maken waarin we alle goede zaken combineren.



Figuur 3: Openingsscherm van FAconvertor. Via Open kun je de twee files (CTM en TXT) invoeren.

Convert

Wanneer de .CTM en de .TXT files zijn ingeladen, klik op 'Convert'.

Uitleg:

De twee tekstfiles die inhoudelijk dezelfde woorden bevatten, worden dan gelezen en er wordt intern eerst een nieuwe CTM-file gemaakt die wel leestekens, woorden met hoofdletter en sprekers bevat. Die wordt vervolgens gebruikt om het uiteindelijke resultaat te maken. Na klikken op 'Convert' zie je zo iets als het volgende:

#	StartTime	EndTime	FA-Word	Text-word	Speaker
1	0	830	SIL	SIL	
2	830	5560	??Hij	Hij	[Speaker: SP_3 - Jan van de Pavert]
3	5560	5770	SIL	SIL	
4	5770	5950	doet	doet	
5	5950	5980	SIL	SIL	
6	5980	6930	??het!	het!	
7	6930	6960	SIL	SIL	
8	6960	7030	We	We	
9	7030	7300	zitten	zitten	
10	7300	7410	in	in	

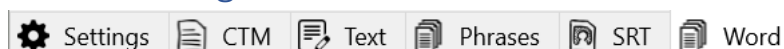
CTM-file: E:\Documenten\Projecten\OH\Sanneke\Pavert\Pavert_20090527_1.ctm
Text-file: E:\Documenten\Projecten\OH\Sanneke\Pavert\Pavert_20090527_1.txt
Word-file: E:\Documenten\Projecten\OH\Sanneke\Pavert\Pavert_20090527_1_word.txt

Figuur 4: Na de twee files geladen te hebben en op Convert geklikt te hebben, zie je het volgende scherm. Index, start en eindtijd van ieder woord. Dan de twee woorden: eerst het FA-word dat uit de CTM-file komt (geen leestekens, altijd klein geschreven) en daarna het Tekst-word uit de nette tekstfile. Deze bevatten wel komma's, punten, vraagtekens en uitroepetekens. Ook zijn ze soms met een hoofdletter geschreven. In de laatste kolom staat de spreker.

Opslaan

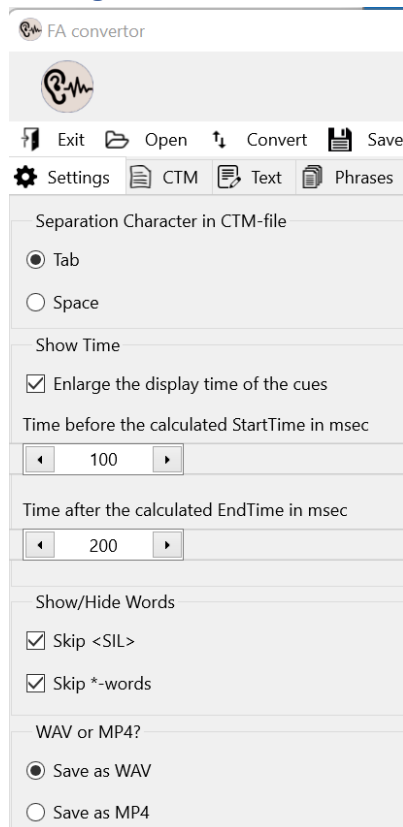
1. Klik dan op 'Save'.
2. Kies waar je de srt (ondertitelingsfile) wilt opslaan. Default is de directory folder van waaruit je de bestanden hebt ingelezen. Tegelijk met de srt-file, worden ook een html-file voor afspelen in karaoke-stijl en een goed leesbaar transcript in een semi-Word file (**-word.txt**) gedownload. Dit laatste document kun je bewerken in een tekstverwerkingsprogramma, bijvoorbeeld om het te voorzien van inhoudelijke metadata voor de context van het document. In principe ben je nu klaar om het materiaal veilig te deponeren bij [DANS-KNAW](#).

Wat is er nog meer te zien?



In de 6 tabs staan verschillende velden met settings, input en output data

Settings

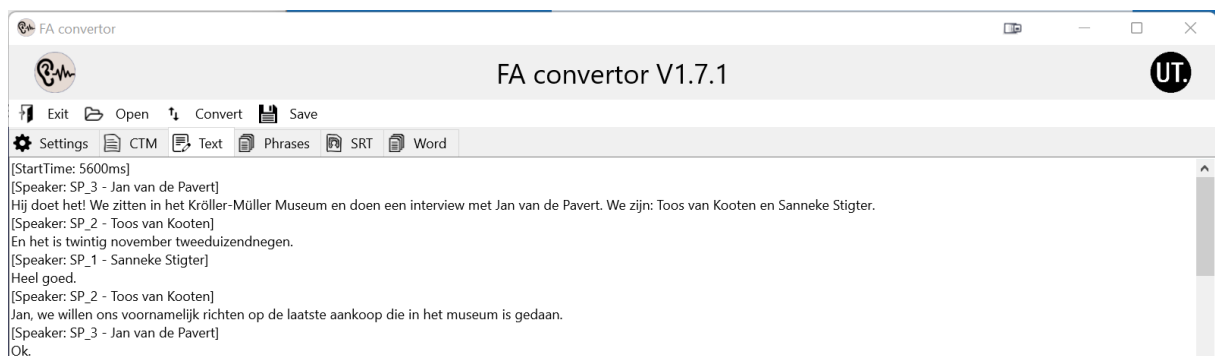


Van belang zijn voorlopig de optie Tab of Space. Sommige CTM-file scheiden de kolommen met een spatie en anderen juist weer met een Tab. En als laatste of je voor de Karaoke-file de wav-file of een mp4-file wilt gebruiken.

CTM

Geeft de resultaten van de twee files (CTM en TXT) en laat het uiteindelijke resultaat zien.

Tekst



Dit is het resultaat van de tekst invoer-file. We zien dat de woorden met een hoofdletterwoord beginnen, komma's, punten, vraagtekens en uitroepetekens bevatten en dat bepaalde woorden ook met een hoofdletter geschreven worden. Deze tekst nogmaals, wordt gebruikt voor de generatie van de uiteindelijke resultaatfiles.

Phrases

Een overzicht (voor de generatie van de ondertitelingsfiles) van de zinnen die er instaan. Elke zin eindigt met een punt, uitroepeteken of vraagteken.

SRT

Dit laat de gegenereerde ondertitelingfile zien (*.srt).

Word



Hier zien we de platte txt-file die een leesbaarder versie geeft dan de invoer tekstfile. De file kan makkelijk in MSword geopend worden en omgezet worden in een mooie en bewaarbare docx-file.

Vragen

Voor vragen en opmerkingen, mail mij maar:

A.J.vanHessen@utwente.nl