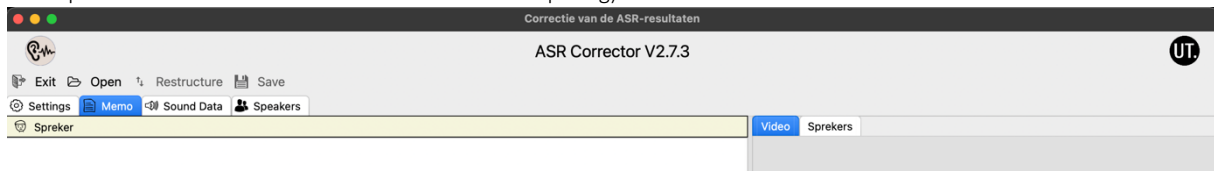


Handleiding ASRcorrector

Algemeen

De ASRcorrector-software is bedoeld voor het corrigeren van de resultaten van de eerste transcriptie, gedaan via Automatische Spraakherkenning (ASR). Daartoe worden de CTM-resultaten van een ASR door de software gelezen en omgezet in een speciaal XML-bestand. Dit XML-bestand bevat de automatisch gegenereerde transcripties, de door een menselijke operator geconverteerde transcripties, het herkende audio- of videobestand en de voortgangsstatus (het moment in het bestand waarop u aan het werk was net voordat u het bestand oploeg).



Figuur 1: Het beeld van de ASRcorrector als je de software opstart. De getoonde versie (V 2.7.3) bevat een mix van Nederlandse en Engelse titels/headers. Zodra de software "klaar" is zullen we het veranderen in echt Engels

Zodra de correcties zijn uitgevoerd, kunnen de resultaten worden geëxporteerd naar een speciaal tekstbestand voor de Geforceerde Uitlijning: een andere dienst bedoeld om de gesproken en geschreven tekst op elkaar af te stemmen.

Introductie

Met het gebruik van Deep Neural Networks (DNN) in moderne ASR is de spraakherkenning behoorlijk goed geworden. D.w.z. dat onze Nederlandse en Engelse herkenningssystemen meestal een betere dan 90% correcte herkenning van de gesproken woorden geven. Dit betekent dat het in een gewoon interview zinvol is om de ASR-resultaten te gebruiken als de basistekst die moet worden gecontroleerd en uiteindelijk gecorrigeerd.

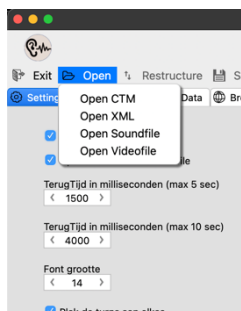
Bij een klassieke transcriptietaak duurt het tussen 6 en 8 uur om de gesproken inhoud woordelijk te transcriberen. Met de combinatie van ASR en correctie van de ASR-resultaten daalt dit echter tot 2 à 4 uur (uiteraard afhankelijk van de kwaliteit van de audio, de duidelijkheid en juistheid van de gesproken taal, enz. enz.)

Bovendien maakt de software het mogelijk om verschillende sprekers toe te voegen, een of meer woorden door een ander woord/woorden te vervangen, zinnen te maken, tekens voor onvoltooide zinsuiteinden toe te voegen en meer.

Start van de ASRcorrector

SHOW EXPERT MODE

De eerste keer dat je **ASRcorrector** opent, zal je iets zien zoals afbeelding 1 hierboven: een nogal leeg scherm. Om het CTM-bestand (het resultaat van de ASR-engine) in te lezen, ga je naar instellingen en klikt op "Toon expertmodus" (=Show expert mode). Als resultaat ziet men een aantal toevoegingen (zie fig.2).



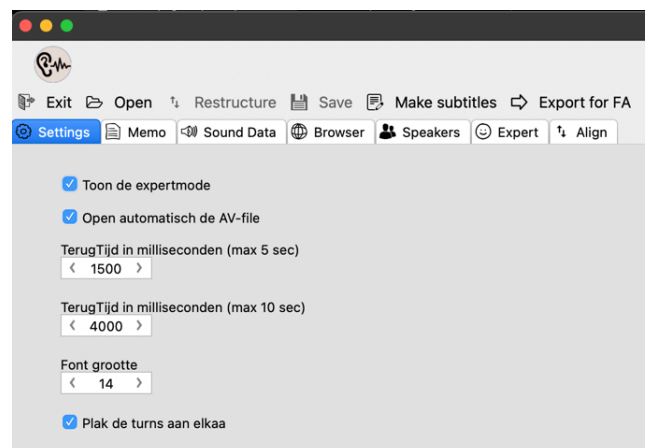
Figuur 1: Open de CTM-file

Verander de andere velden niet de eerste keer dat u ASRcorrector opent!

READ CTM-FILE

Als je nu op de **Open** knop klikt, kun je het CTM-bestand lezen.

Nadat je het ctm-bestand hebt ingelezen, open je of het audio- of het videobestand. Sla alles vervolgens op als een XML-bestand. Voor de correctiefase kun je nu het vinkje bij "Toon Expertmode" uitzetten.



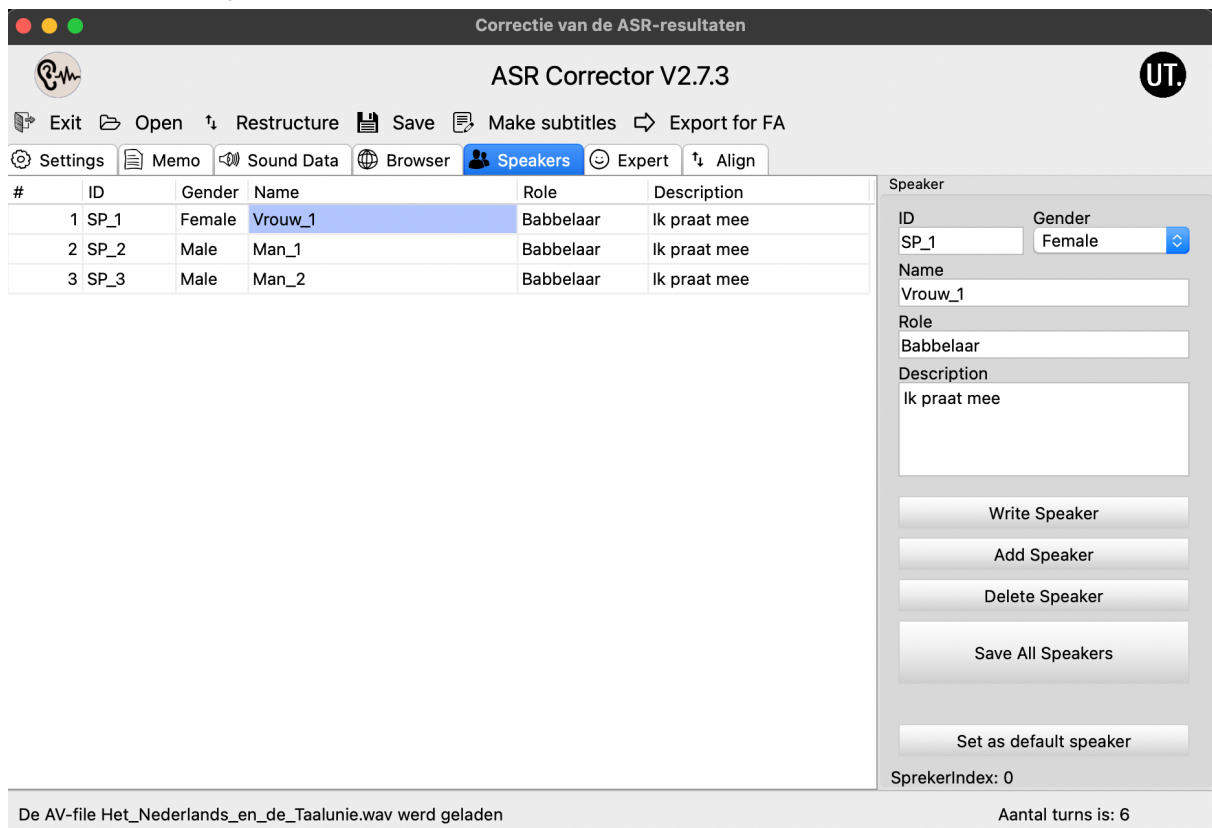
Figuur 2: Voor het lezen van het CTM-bestand, moet u klikken op "Show expertmode" in de instellingen tab

Aan de slag met de XML-file

Na de beginfase waarin u het CTM-bestand leest en samen met het audio/videobestand opslaat in het XML-bestand. Voor de correctiefase is dit **XML-bestand** het werkbestand.

TOEVOEGEN SPREKERS

Het volgende dat gedaan moet worden, is het toevoegen van de verschillende sprekers. Klik op het tabblad **Speakers** en je ziet een scherm zoals in figuur 4.



Figuur 4: Toevoeging van sprekers.

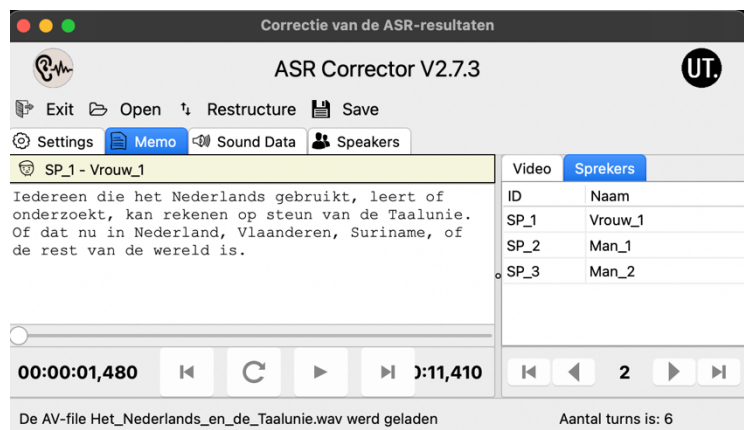
Om een nieuwe spreker toe te voegen, klik je op **Add Speaker**. Als ID krijg je SP_1, het geslacht wordt ingesteld op overig en de naam wordt ingesteld op male_x of female_x. Rol en Beschrijving kunnen worden gebruikt voor aanvullende informatie. Nadat je deze parameters hebt ingevuld, klik je op **Write Speaker**. De resultaten worden aan de linkerkant van het paneel toegevoegd. Als je alle sprekers hebt toegevoegd, klik je op **Save All Speakers**.

CORRECTIE

Klik op het tabblad Memo om de correctie van de ASR-resultaten te starten. In het veld aan de linkerkant van het paneel zie je de herkende tekst.

Corrigeer deze tekst.

Onder dit tekstveld zie je de audioknoppen, aan weerszijden begrensd door de begin- en eindtijd van dit fragment. De 4 knoppen in het midden zijn voor het afspelen/pauzeren van het geluidsbestand, begrensd door de begin- en eindtijd. Je kunt de balk ook gebruiken (vooral voor langere fragmenten) om het begin van het geluidsfragment in te stellen.



Figuur 5: het correctievenster.

Op het rechter paneel zie je 2 tabs: Video en Speakers. Als er een video is, zie je die video in de video-tab. Rechts-onderaan zie je de knoppen voor de chunks of beurten (fragmenten van de audio/video). Deze chunks worden ingesteld door het input/CTM-bestand. Indien je een *.ctm.spk bestand selecteert, worden de wisselingen van de sprekers als input gebruikt. Indien je een *.ctm file gebruikt, zal elke chunk 100 woorden bevatten.

De knoppen zijn -10, -1, +1 en +10 naar rechts of naar links.

SPREKER SELECTIE

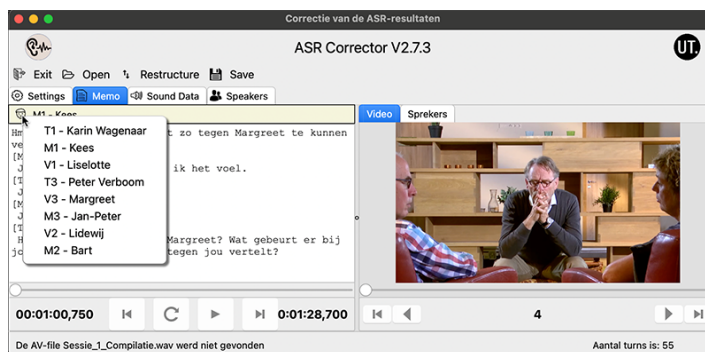
Er zijn twee manieren om een luidspreker in te stellen. De eerste manier is om de luidspreker voor de chunk in te stellen. Dit doe je door met rechts te klikken op het hoofd boven het tekstveld en dan de juiste spreker te kiezen.

Het komt echter vaak voor dat je ergens in het tekstveld nog een spreker moet toevoegen als iemand anders begint te spreken.

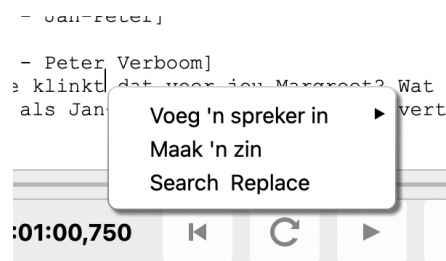
Ga daarvoor naar de plaats in de tekst waar je een nieuwe spreker wilt toevoegen. Klik met de rechtermuisknop en selecteer het bovenste item (Voeg een spreker toe of Voeg 'n spreker in). Selecteer vervolgens opnieuw de juiste spreker.

De software voegt de spreker toe met zijn ID en naam (bijvoorbeeld: [T3 - Peter Verboom]).

De harde returns hebben geen betekenis: ze zijn alleen toegevoegd om de leesbaarheid van de tekst te verbeteren.



Figuur 6: Toevoegen van de hoofdspreker van een turn.

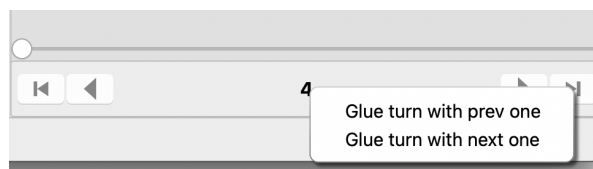


Figuur 7: Toevoegen van een andere spreker in het midden van de tekst van de beurt.

COMBINEREN TURNS

Het kan gebeuren dat je de beurt waaraan je werkt, wilt combineren met de vorige of volgende beurt. Dit gebeurt vooral wanneer de beurten dezelfde hoofdspreker bevatten en vrij klein zijn.

Klik daartoe met de rechtermuisknop op het nummer van de beurt in het rechter benedenvenster en selecteer prev of next. De afslagen worden dan aan elkaar geplakt, en de tekst wordt samengevoegd in één tekst. Voor een betere werkbaarheid moet u de afslagen niet te groot maken. Stukken van ongeveer 100 woorden zijn ideaal voor correctietaken.



Figuur 8: Lijmen van twee naast elkaar liggende bochten.

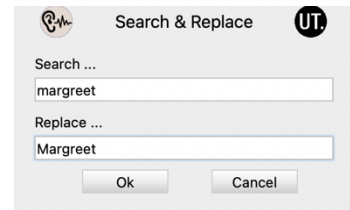
ZINNEN MAKEN

Zoals hierboven uitgelegd, geeft de ASR alleen woorden met kleine letters en zonder leestekens. Om het lezen van de ASR-resultaten beter/gemakkelijker te maken, hebben we een zinsdetector gemaakt die na elke pauze van 400 msec of langer een nieuwe zin begint. Het blijkt dat deze "400 sec grens" meestal vrij goed werkt. Er zijn echter sprekers en/of situaties waarin mensen aarzelen in het midden van een zin, waardoor de zinsgenerator te veel zinnen toevoegt, en anderzijds kunnen sprekers snel praten zodat de software geen verschil ziet tussen twee opeenvolgende uitingen en vermijdt er twee zinnen van te maken.

Gegeven een correct ASR-resultaat, kan men de tekst in twee zinnen splitsen door met de rechtermuisknop te klikken op de spatie tussen de twee zinnen en "Maak 'n zin" te selecteren.

ZOEK EN VERVANG WOORDEN

Door het ontbreken van alle hoofdletters moeten geografische en persoonsnamen vaak één voor één worden gewijzigd. Om dit proces te vereenvoudigen, hebben wij een eenvoudige zoek- en vervangroutine gemaakt die alle voorkomens van een woord of woordcombinatie vervangt door een ander woord of een andere woordcombinatie. Selecteer daartoe het (de) te vervangen woord(en), klik met de rechtermuisknop op de selectie, en selecteer **Search & Replace**. De software kopieert het woord in zowel het zoek- als het vervangveld. Als je op **Ok** klikt, wordt het oorspronkelijke woord in het hele document vervangen door het nieuwe woord.



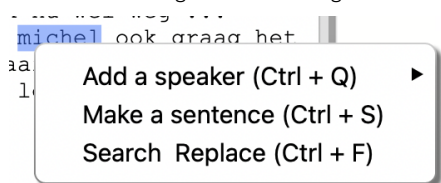
Things-To-Do

Het is moeilijk om een volledige reeks instructies te geven van dingen die noodzakelijkerwijs moeten worden gedaan. Wij kunnen echter wel enkele suggesties doen om ervoor te zorgen dat de transcripties zo goed en zo uniform mogelijk worden uitgevoerd.

NAMEN

Namen worden door de spraakherkenner niet als zodanig herkend en verschijnen dus als "normale woorden" in de transcriptie. Als bijvoorbeeld michel maar één keer wordt gezegd, dan kun je die michel gewoon vervangen door Michel. Als je echter denkt dat michel vaker voorkomt, dan is het misschien handiger om eerst michel te selecteren en dan op de rechtermuisknop te klikken.

Je ziet dan de volgende afbeelding:



Kies dan de knop "**Search Replace**", vervang michel door Michel en klik op OK.

Het programma zal nu alle michel's vervangen door Michel.

U kunt dit natuurlijk ook gebruiken om andere woorden en woordcombinaties te vervangen, zoals "zuid holland" door "Zuid-Holland". Het werkt goed, maar je moet wel oppassen met woorden die niet alleen namen zijn, maar ook gewone woorden. Denk bijvoorbeeld aan "oud-premier Kok" en "de kok in het restaurant". Hier kun je helaas niet zomaar "kok" vervangen door "Kok".

ZINNEN

Zorg ervoor dat je elke zin met een hoofdletter begint. En vergeet niet om een onvolledige zin te eindigen met **drie** puntjes aan het eind. Bijvoorbeeld: "en toen dacht ik...". Met die drie puntjes geef je aan dat de zin eindigt, maar dat er eigenlijk een deel van de tekst ontbreekt.

Je kunt er ook voor zorgen dat de zinnen niet te lang zijn. Dus ook al kun je er één lange zin van maken, het verdient de voorkeur om hem op te splitsen in twee kortere zinnen.

Voorbeeld:

"We zijn er nog elke dag mee bezig **en** dat wil ik de mensen ook meegeven dat de meeste mensen die hier nu naar gaan kijken het ook met mij eens zullen zijn dat het nooit af is."

Kan ook geschreven worden als:

"We zijn er nog elke dag mee bezig. **En** dat is ook wat ik de mensen wil meegeven, dat de meeste mensen die er nu naar gaan kijken het ook met mij eens zullen zijn dat het nooit af is."

NUMMERS

Veel getallen zijn erg lang (driehonderdzesenzestigduizend tweehonderdachtennegentig euro) en kunnen vaak als getal worden opgeschreven: €366298. Maar... dit hoeft niet altijd. Vooral bij kleinere getallen kun je er de voorkeur aan geven ze gewoon uit te schrijven. Probeer hier echter wel zo consequent mogelijk mee om te gaan.