

Forced Alignment: een krachtig hulpmiddel voor spraakonderzoek

Arjan van Hessen, Universiteit Twente

Louis ten Bosch, Radboud Universiteit, Nijmegen

In veel onderzoek aan spraak wordt gebruik gemaakt van Forced Alignment. Bij een Forced Alignment wordt de orthografische transcriptie van een bepaald audiofragment "opgelijnd" met dat fragment. Deze oplijning houdt in dat het akoestische begin en einde van elk woord zo precies mogelijk worden gezocht als ankerpunten in de audio. En dit oplijnen geldt niet alleen de woorden maar ook de eventuele stiltes voor, tussen en na de woorden. Als resultaat van de Forced Aligner weet je precies hoe lang woorden en stiltes duren. Deze kennis is van groot belang bij bijvoorbeeld onderzoek naar sprekervariatie, naar uitspraakvariatie, spreektempo, en voor het mogelijk maken van het semi-automatisch doorzoeken van audiobestanden via geschreven queries.

Bij het CLST in Nijmegen is in samenwerking met de Stichting Open Spraaktechnologie een aligner gebouwd waarin niet alleen woorden maar ook de spraakklanken in elk woord worden opgelijnd met een audiofile. De resultaten op woord- en foonniveau komen tegelijkertijd beschikbaar als twee tiers in een Praat textgrid file. Daarnaast is het mogelijk de aligner een eigen woordenboek mee te geven waarin bijvoorbeeld specifieke woorden kunnen worden voorzien van afwijkende uitspraakrealisaties. Dat maakt onderzoek aan uitspraakvarianten mogelijk.

In de presentatie gaan we in op de functionaliteit van de aligner in een aantal realistische toepassingen, en op de design filosofie van de forced alignment webservice.