

Spraaksynthese met kinderstemmen

*Arthur Dirksen
Fluency, Amsterdam*

Kinderen die zelf niet spreken gebruiken daarvoor een communicatiehulpmiddel met spraaksynthese. Maar vaak moeten ze zich behelpen met een computerstem die is ingesproken door een volwassen spreker. In samenwerking met rdgKompagne is Fluency daarom begonnen met de productie van een reeks kinderstemmen, ingesproken door jongens en meisjes van 8-12 jaar uit verschillende regio's.

Een probleem dat eerst opgelost moest worden is dat het tekstcorpus dat we laten inspreken voor een volwassen stem voor deze nog jonge sprekers veel te ingewikkeld is, zowel wat betreft de woordenschat als wat betreft de zinscomplexiteit. Daarom hebben we een nieuw corpus samengesteld, dat bestaat uit 300 vrij eenvoudige zinnen van gemiddeld 8,3 woorden. De woordenschat is zoveel mogelijk afgestemd op gebruik in een communicatiehulpmiddel.

Het inspreken van de 300 zinnetjes kost ongeveer 3 uur en levert een kwartier spraak op. Dit is erg weinig voor de unit-selectie synthese waar we mee werken: nieuwe zinnen worden gemaakt door geschikte fragmenten uit de opgenomen spraak aan elkaar te knopen, en dit gaat beter naarmate er meer spraak is om uit te kiezen. Niettemin is het toch steeds gelukt om een bruikbare computerstem te maken. We zullen dit illustreren met voorbeelden van de zes kinderstemmen die we tot dusverre hebben opgenomen.