

Dutch Public Broadcaster NPO's Signing Avatar Makes a Gesture

JOOST NEGENMAN (NPO), **INGE ZWITSERLOOD** (RU NIJMEGEN), **JOHN GLAUERT** (SCHOOL OF COMPUTING SCIENCES UEA) AND **MARTIN PRINS** (TNO)

In 2012, the Research and Development (R&D) team at NPO was asked: “What is the state of the art on automated sign production?” We decided to take a closer look and lead a market consultation on the topic. Soon after, we discovered that providing automated sign language was far from possible at any reasonable level of quality. However, we did find one solution that was being used in several European projects which caught our eye.

These projects used a technique generating animated signing based on a structured mark-up language called ‘SiGML’ (Signing Gesture Mark-up Language). We also took note of other international projects using animated sign language i.e. some were starting to use motion capture. The graphical results were good and probably better than the SiGML approach, but creating signs this way was more content creation than a sustainable broadcast chain solution. Therefore, it was time to investigate SiGML further.

We found that featuring items like NPO news or talk shows was not feasible due to the complexity of the programmes (i.e. vocabulary, rate of speech). So the search for a programme where animated signing does offer support for the hearing impaired began. While scrutinizing our accessibility programmes we became aware that there was one target audience that was not supported – children. Because young deaf and hearing impaired children in the youngest age group are incapable of reading subtitles, we decided that we should focus on this age group. Very soon the choice was made to use the NTR title ‘Sand Castle’, an educational series with limited speech and few characters (puppets) as a pilot.

NPO consulted several interested parties that, at first, were reluctant, fearing that this might replace human interpreters. After emphasizing that we sought various ways to create added value for deaf children, a common goal was recognized.

The project

The main aim of the TNO-led project was to develop and evaluate a proof of concept for a signing avatar service using SiGML. Our goals was to identify to what extent a signing avatar service could be developed with existing technologies, and to evaluate to what extent a signing avatar service could be of added value for hearing impaired children. Finally, we aimed to establish guidance on the possibilities and limitations of existing technology, with respect to usefulness, technical readiness, user-friendliness, and suitability for a broadcasting environment.

It was clear from the start of the project that automatic translation of speech or text (i.e. subtitles or a transcript) to a sign language was an unsolved issue. Sign languages are visual languages with their own grammar constructs, in which the (3D)-signing space plays an essential role. Speech and text often do not contain all the necessary information (metadata and context) that would allow a sign translation algorithm to establish how subjects and objects are interrelated (as these relations may need to be visually shown). If you don't have the context, it is difficult to address something that occurs in the TV-programme.

It was decided that we would not focus our proof of concept on the translation problem, but instead, on aspects that could realistically be provided. The trade-

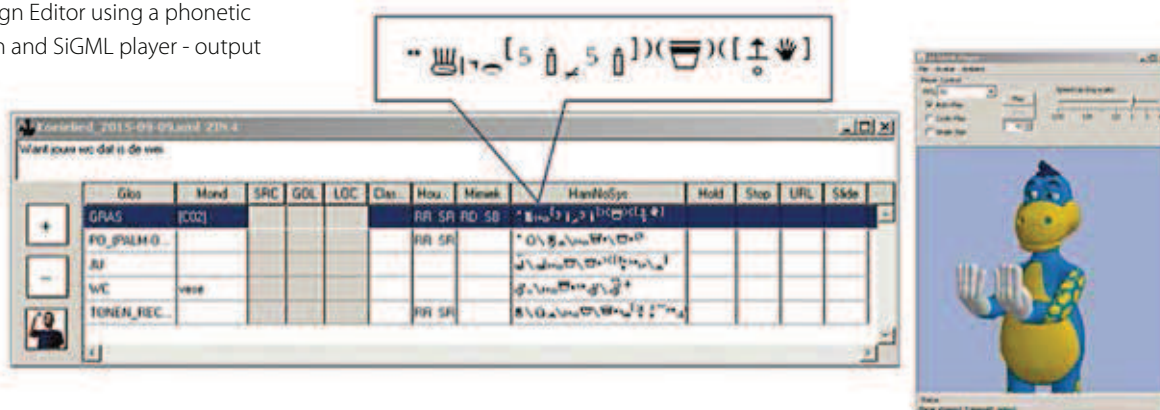
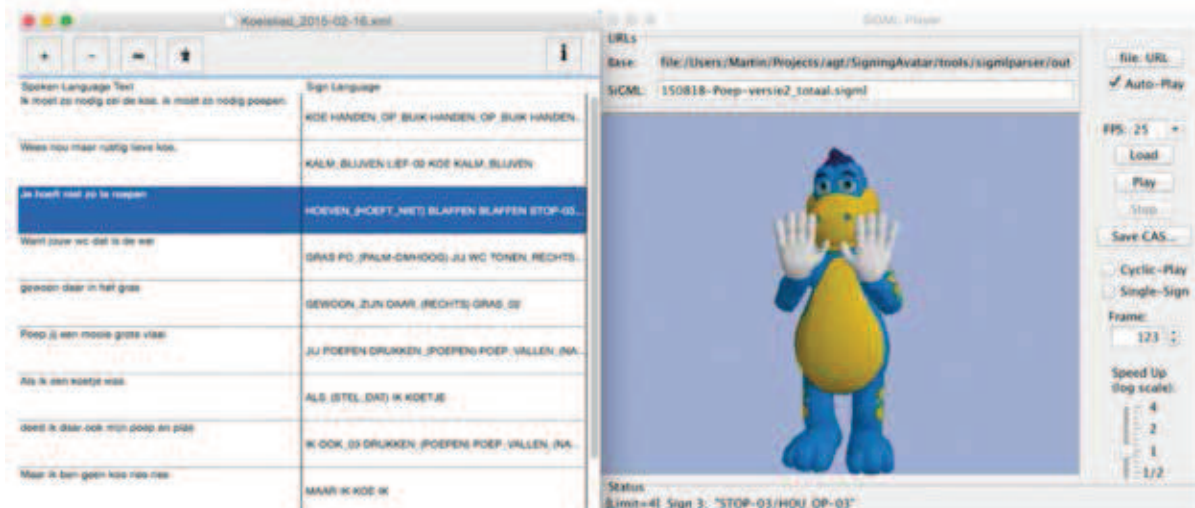


Figure 1: eSign Editor using a phonetic transcription and SiGML player - output

Figure 2: eSign Editor constructing sentences and SiGML player - output



off was to use Sign supported Dutch (Nederlands met Gebaren – NmG) instead of the Sign Language of the Netherlands (Nederlandse Gebarentaal – NGT). NmG uses signs from NGT for content words and Dutch grammar. This allowed us to look at a transcript of a ‘Sand Castle’ episode and use text analysis techniques to filter out words and grammatical aspects that were not expressed in NmG. This reduced the transcript that was used as input for the signing. Additional reductions were achieved by asking a deaf person to compare the episode with the transcript and to highlight the sentences that the child would probably comprehend merely by looking at the episode. These results were used to construct the signs and sentences that would be rendered by the signing avatar.

Creating the Avatar

Signing is constructed in several steps. First, the form of each sign is coded in an editing programme (eSign Editor), using a phonetic transcription system (Hamburg Notation System: HamNoSys) with symbols

for the shape, orientation, location, and movement of the hand(s). The hand in a sign may be a fist, flat, rounded or have any number of extended and curved or bent fingers. The sign can have a variety of movements, for example: straight, circular, or opening of the hand. The hand may move towards or away from a body part such as the shoulder or the side of the head or in the space in front of the body.

Second, sentences are created by selecting the desired signs in the database and placing them in the correct order. They are then provided with additional codes for prosodic and affection facial expressions and postures or movements of the head and body. Third, the finished text is translated into an XML application language such as SiGML (developed at the University of East Anglia), that is based on HamNoSys. Specialised software generates animation data for an avatar from the SiGML for a sign. Signs can be played in sequence and the software inserts smooth transition movements from the end of one sign to the start of the next.

Part 2 to be continued in our next issue...

Figure 3: Avatar designs – from human-like to cartoon-like

