Interactive HamNoSys Notation Editor for Signed Speech Annotation

Jakub Kanis, Zdeněk Krňoul

We have created the domain-specific lexicon for our synthesis system from railway station domain. The signs which need to be notated were collected by the inspection of the Czech to Signed Czech (CSC) parallel corpus and translations of train announcements. The CSC corpus contains 1100 dialogues from the telephone communication between a customer and an operator in a train timetable information center.

We have begun the trial annotation process with six annotators to test the conveniences of the SignEditor and the feedback animation. We have divided the annotation process into two steps. In the first step, four annotators who are not familiar with CSC were employed to insert the signs in the direct editor. The annotators used the video dictionaries of CSC as a source for the sign notation. In the second step, two remaining annotators (instructors, familiar with CSC) were employed to correct the entered signs. The inspector used the SignEditor to reply signs and put comments about the correctness of the rendered animations.

The lexicon currently contains approximately 330 signs. By the inspection of the lexicon, we can observe that all signs include some variant of the starting point configuration. It is interesting that the most frequent symbols in the starting point configuration are the hand shape symbol described open hand and the location symbol for thorax. The most frequent sequences of symbols in the movement block are those for the relative change of hand shape and the finger orientation or absolute displacement of the vital position. More than 85 % of all signs in the lexicon contain these sequences. The majority of this sequences is annotated in combination with another simultaneously performed movements. The 51 % of all signs are notated with the symbols for straight movements which form the path of write. The symbols of circular movements were chosen for 7 % of all signs only. The symbols for the movement repetition are approximately used for 20 % of all signs.

The annotators have had problems with notation of very small movements. The modification symbols for small, normal and large movements seem not to be sufficient. Furthermore, they have had also problems with location symbol for thorax. It seems that the current annotation variant is not enough in this case. There could be more possibilities how to annotate more precisely this location. These experiences partially agree with comments by inspectors who checked signs in the lexicon.