Annotation and Maintenance of the Greek Sign Language Corpus (GSLC)
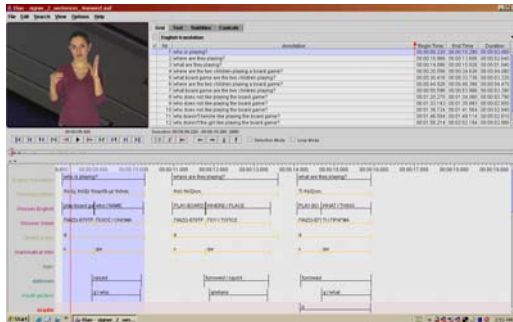
Eleni Efthimiou, Stavroula - Evita Fotinea

Institute for Language and Speech Processing (ILSP) / R.C. Athena
Artemidos 6 & Epidavrou, GR 151 25 Athens, Greece
e-mail: eleni_e@ilsp.gr , evita@ilsp.gr

Abstract : Here we present the design and development of a representative language corpus for the Greek Sign Language (GSL). Focus is put on the annotation methodology adopted to provide for linguistic information and annotated corpus maintenance and exploitation for the extraction of a linguistic model intended to support both sign language recognition and creation of language based tools and educational content.

GSLC

Corpus creation

Native signer digital video recordings by use of 3 synchronised cameras (stereoscopy)

Corpus development mechanisms

Native dialogues

Video & picture based elicitation

Lemmas

Selected sets of morpheme groups to support sign recognition

Elicited phenomena

controlled sentence groups to capture the main GSL structural mechanisms

Free narrations

Spontaneous story telling and dialogue driven utterances

Annotation procedure

Morpheme & sentence level manual annotation –

27 tiers of linguistic & meta-linguistic information

Quality control

Corpus creation control and annotation evaluation on peer review basis and external expert assessment

Exploitation of the annotated parts of GSLC

Extraction of measurements for sign recognition

Extraction of linguistic model for GSL

Language based tools

Educational content