Motion Capture 3D Sign Language Resources

Pavel Jedlička   Zdeněk Krňoul   Miloš Železný   Luděk Müller

University of West Bohemia, Faculty of Applied Sciences, NTIS - New Technologies for the Information Society Univerzitní 8, 306 14 Pilsen, Czech Republic
{jedlicka, zdkrnoul, zelezny, muller}@ntis.zcu.cz

Project Aim

- **Create a Motion Capture (MoCap) data-set for Czech sign language (CSE)**
- Data-set primary purpose: sign language synthesis and machine translation
- Proof of concept for large scale motion capture 3D data-set recording

Motion Capture Setup

Hardware Specification
- MoCap device: 8x VICON T-20 camera (2MPx, @100fps)
- reference RGB video
- marker sizes: from **4 mm** hemispherical (face, fingers) to **14 mm** spherical (body)

Two-step Recording Approach

Macroscopic movement recording:
- Movement of body, head, arms, palms.
- Simplified finger setup: 3 markers on fingertips (Thumb, Index, Pinky)
- Simplified facial setup: 2 markers on eyebrows, 4 on mouth, 1 on nose
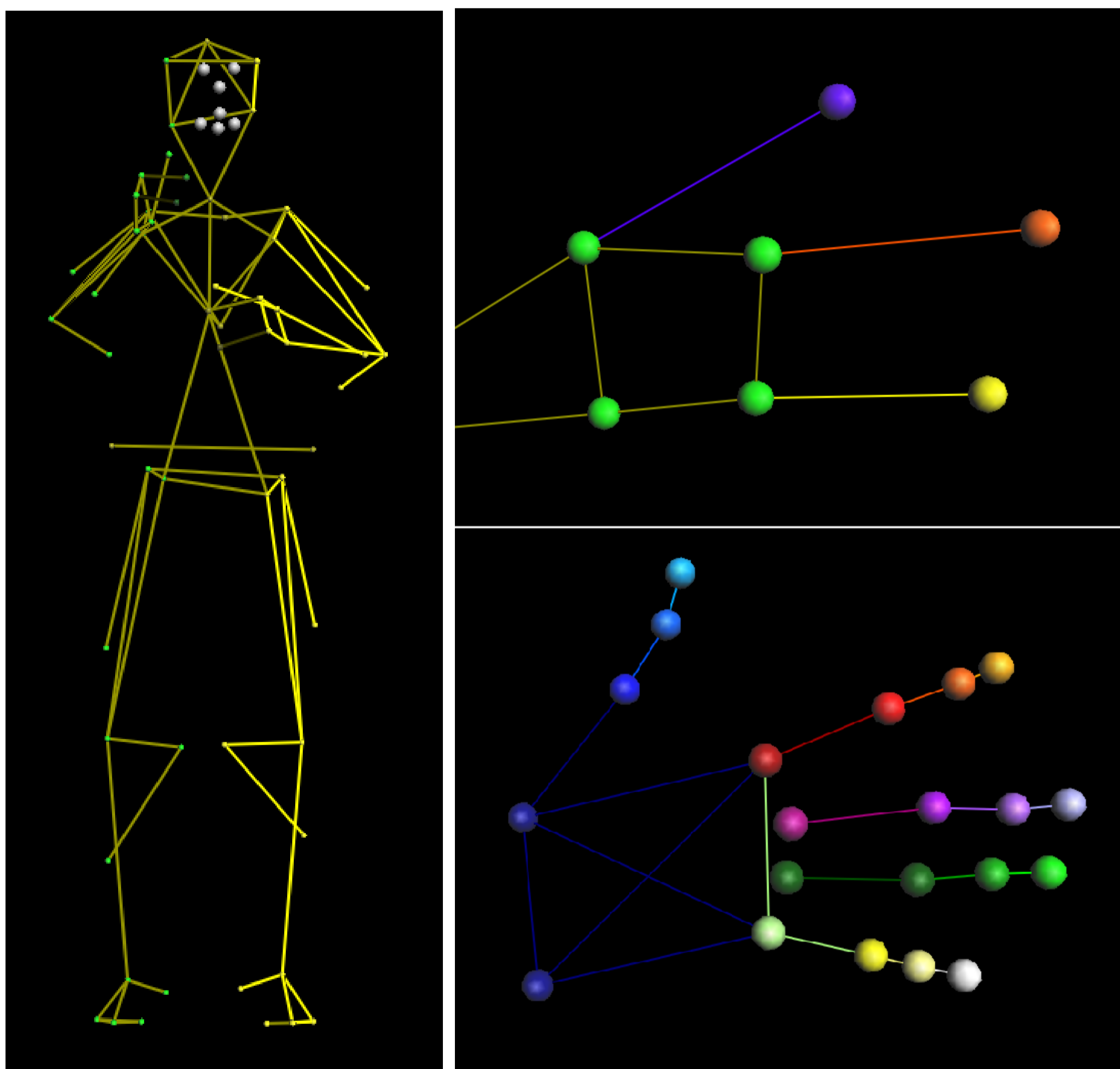
Detailed hand-shape setup:
- 3 markers per finger (fingertip, distal and proximal interphalangeal joints)
- Shares palm tracking with the macroscopic to provide connectability
- 52 selected hand-configuration recorded
- recording of hand-configuration dynamics (relaxed-configured-relaxed)

Data-set Properties

- Czech Sign Language (CSE)
- Cooperation with CSE linguists - Institute of Special Education Studies, Palacký University Olomouc
- Topics (limited vocabulary): **weather forecast**, **zoo tour**
- 6 different native CSE speakers (age 18 - 50, 4 males, 2 females)
- Each individual speech is recorded twice by same the SL speaker
- **Linguistic annotation**
  - annotated by CSE native speakers
  - annotations in ELAN [1] based on reference video
  - 3-tier annotation
    - identification of individual signs (glosses)
    - identification of both hand-shapes
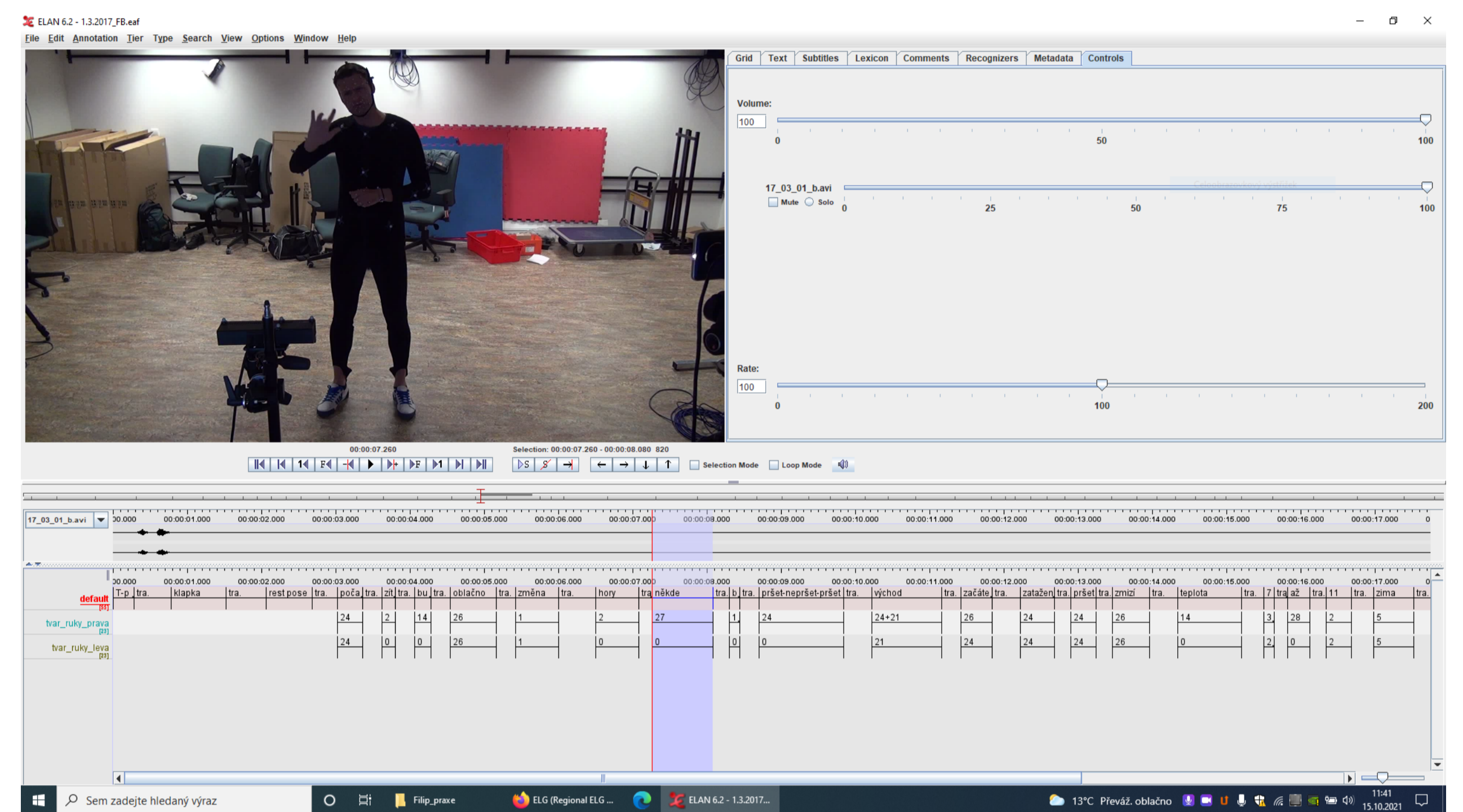
Data-sets Details

Topic 1: **Weather forecast**
- 36 individual forecasts (one forecast $\sim 30sec$ continuous speech)
- Vocabulary diversity (3 forecasts per month)
- Large sample of numbers (temperatures)
- Limited vocabulary - multiple instances of same sign in different context (frequent signs more than 20 repetitions)
- All signs recorded also as dictionary items (separated signs)

Topic 2: **Zoo tour**
- 20 different animals (one animal $\sim 2 - 3min$ continuous speech)
- 4 chapters:
  - Czech name, Latin name (finger-spelling)
  - Scientific classification, biotope, food (highly structured)
  - lifespan, social (telling / low structured)
  - interesting facts (telling / low structured)
- Large sample of letters (Latin names - finger-spelling)
- Overlapping vocabulary (in highly structured part)
- Contains advanced grammatical structures (in low structured part)

Overall data-set information

- Sign Language 3D motion data-set
- Over 18 hours of continuous speech
- Separate dictionary recorded over 300 signs
- Dynamic hand-configuration recording of 52 shapes
- 6 different speakers
- Data in form of marker trajectories and annotation
- Part of the European Language Grid (ELG)

Acknowledgements

[1] ELAN [Computer software]. (2021). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan