Chinese learning
- ICT for Chinese
- What’s missing
- Planned implementation
- Optimize gramrels (rules) for Chinese
- Implement multi-level segmentation

Mandarin Chinese has most native speakers of all languages.
- GCSE entries up 18% in 2015, despite overall decline language exam entries (Guardian 2015).
- Popularity at primary school (Chen 2014)
- Mandarin mastery important for careers (Lo 2016)
- Trade with China
- Confucius Institutes

Existing ICT provision
- Acclaimed corpus analysis platform
- English learner dictionaries
- Resources for many languages
- Suitable for expert users
- Can be used by teachers/learners

Sketch Engine
- Sketch Engine for Language Learning
- Free on web
- English, Czech & Russian currently available
- German & Italian soon
- Suitable for teachers/learners
- Extensible to other languages

Chinese segmentation privileges longer words
- Longest match procedure (Tseng & Chen 2002)
- OOV words often long (Ma & Chen 2005)
- Sun et al (1998) use MI segmentation without lexicon
- MI threshold is parametrizable
- Different word lengths could be favoured

References