

CONCLUSIONS & DISCUSSION

---

**#CLS15**

#CLS15

thanks to

- Dong Ok Lim



- Lin Jiang



- Xiao Wang



#CLS15

**thanks to**

- Proposals
- Attended today & those who wanted to but couldn't

@TaLC2018



18-21 July, 2018 - F. Education, University of Cambridge

**13 Teaching and Language Corpora Conference 2018**

# Submission opens Nov 1

[www.educ.cam.ac.uk/events/conferences/talc2018](http://www.educ.cam.ac.uk/events/conferences/talc2018)

[CHARLOTTE.TAYLOR@SUSSEX.AC.UK](mailto:CHARLOTTE.TAYLOR@SUSSEX.AC.UK)

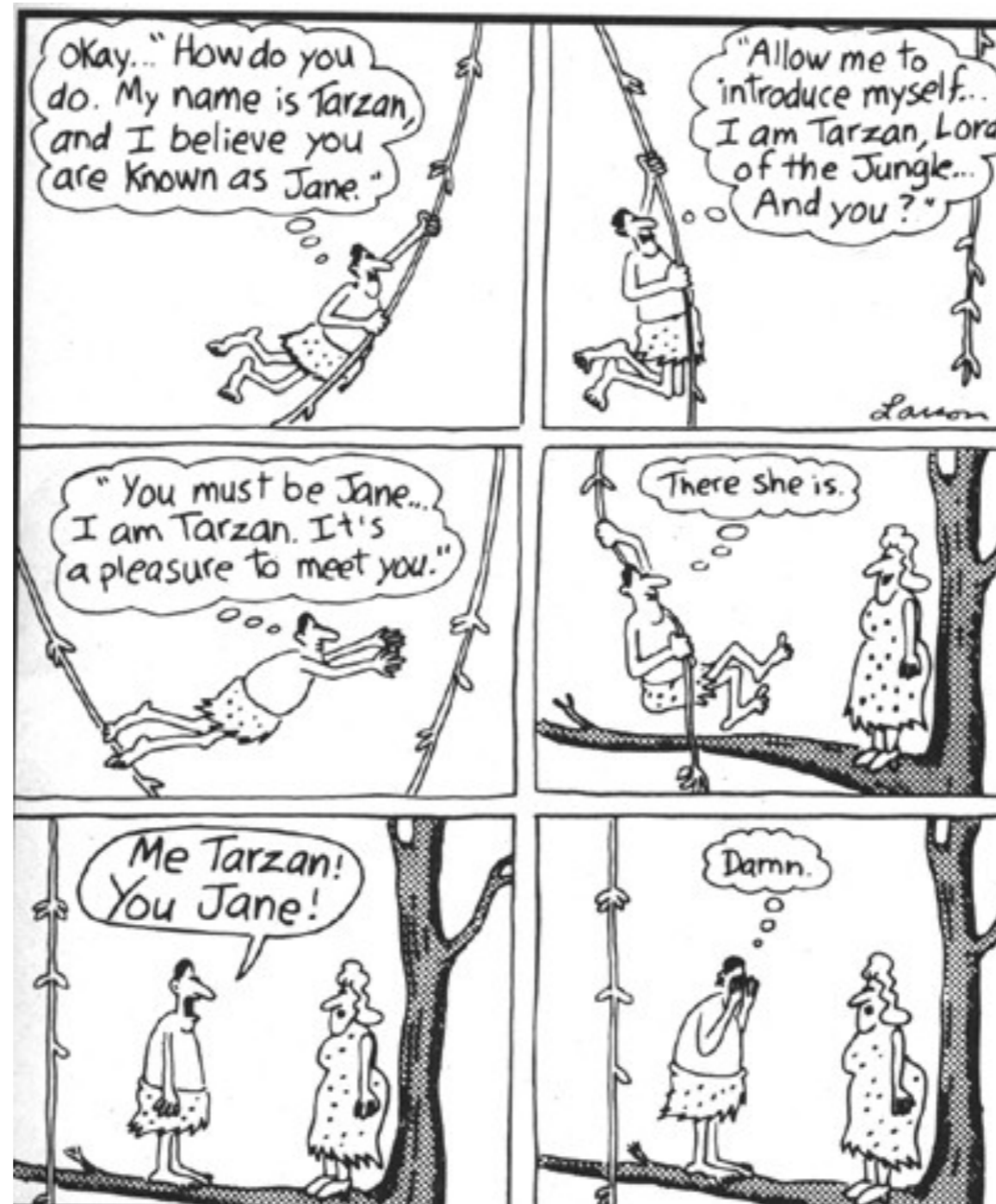
PFP23@CAM.AC.UK

---

**NEXT CLS?**

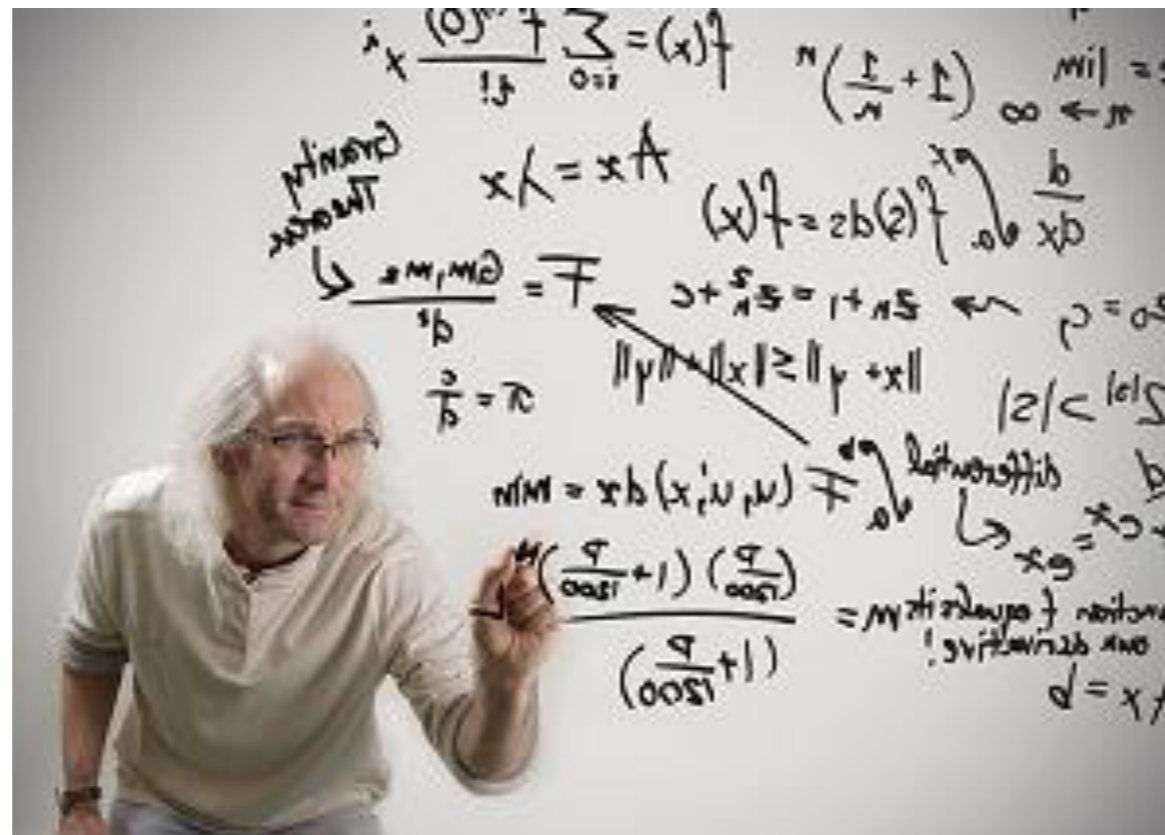
# THE STUFF OF CORPUS LINGUISTICS: ATTESTED USES AND ALL THAT

---



# WE ARE INTERESTING PEOPLE

- ▶ 2000 words in the Brown corpus



## ANALYSING DISCOURSE

- ▶ Corpus linguistics “constructs” (n-gram, collocation, keywords) as useful methods to inform principled annotation processes. How do they determine our findings? Combination of them?
- ▶ Annotation of corpora: coder reliability? What is a coder? Is reliability an issue outside CL?
- ▶ The sole annotator vs coder reliability: research paradigm: interpretative vs post-positivist tensions.



## PRAGMATICS, LEARNER LANGUAGE, TRANSLATION

- ▶ Software-aided annotation: not quite there?
- ▶ Learner language: frequency not enough, ways to improve our understanding of development.
- ▶ Learner language. Large “representative” corpus availability and the many small corpora out there.
- ▶ Limitations of POS Tags.
- ▶ More methodological awareness: within CL from outside CL

## GENERAL REMARKS

- ▶ CL: A scientific method? (Leech 1992). Assumptions about ontology and epistemology in our research area
- ▶ Construction of corpora & use of corpora
- ▶ Context-dependency & human interpretation: aren't "we" more similar to "QUAL people" that we may like to think?
- ▶ Don't we tend to play down aspects of QUAN analysis such as multi-word tagging, disfluencies, etc.?
- ▶ CL methods "in the making": Keyness.
- ▶ Corpus linguistics vs corpus research methods: representativeness, corpora, datasets, analysis.



# 15 Corpus Linguistics in the South

