# Corpus-based translation error analysis and implications for pedagogy

*textbooks meet error-annotation data (S 161)*

Maria Kunilovskaya and Natalia Morgoun

*13th Teaching and Language Corpora Conference*

Cambridge, July 20, 2018

[1] *supported by the Russian Foundation for Basic Research within Project No. 17-06-00107*

# Outline

# How useful is error annotation as a source of knowledge?

General framework:

- research on linguistic properties of translations, translational variety with the focus on learner translations, translationese, quality and professionalism

This reseach aims:

- complement comparative approach to translation properties with parallel and negative data analysis
- to develop a corpus-informed approach to translator education, i.e. to find a way to use error-annotation to indicate

> *"areas of the learning curriculum where teaching is most needed"*

[Castagnoli et al., 2011]

# The context and scope of this research

## Learner translations sample

is limited to

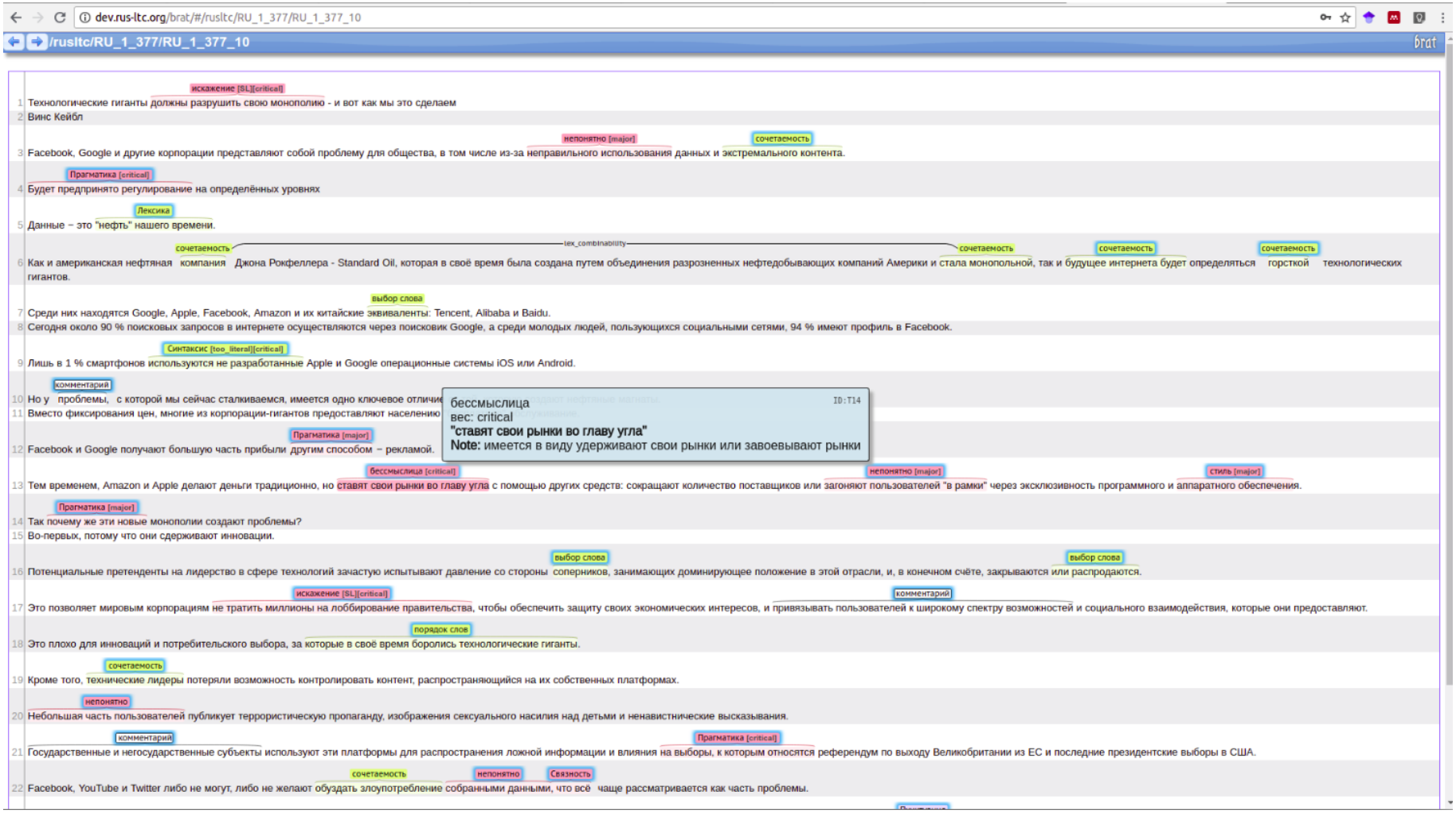
- ST with at least 6 translations
- produced independently, i.e. in the exam or translation contest settings
- English > Russian (Russian = L1)
- by final year students (higher L2 competence, prior translation practice, theory of translation in the recent edu background)

## Translation textbook sample

everything on *translation* from *Library Genesis* search engine is filtered to fit our educational context

- aimed at translators-to-be (BA,MA), not as part of ESP
- based on general domain informational texts, not fiction, technical or academic texts
- practice oriented and methodology, not TS theory
- English > Russian

# The annotation environment is used as a teaching tool




Students get the feedback and can compare solutions.
Multiple translations are ranked on the number and types of errors

# The context and scope of this research, cont.

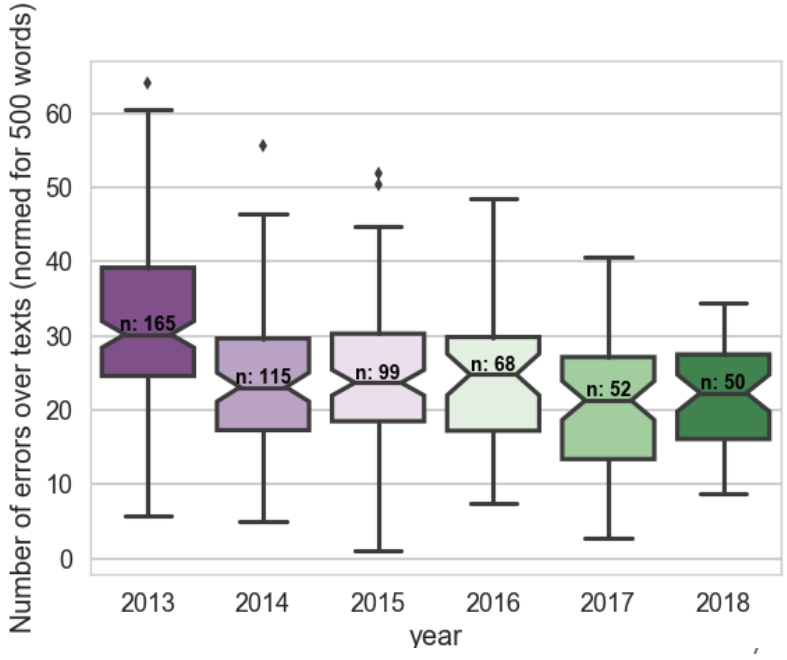
## Limitations of the research data

1. interpretative and subjective nature of translation error annotation

   *Annotation is the "process of adding [. . . ] interpretive, linguistic information to an electronic corpus of spoken and/or written language data"*

   [Leech, 1997]

2. one annotator, but
   - theory-neutral formal guidelines
   - consistency (see boxplot)
3. learners from one university



# DATA1: multiple independent student translations

| | English sources | Russian translatior |
|---|---|---|
| texts | 32 | 405 |
| sentences | 653 | 8098 |
| words | 14640 | 164805 |
| tags | - | 8330 |


## Data formats

- set of *.txt files
- RusLTC[1]customized TMX, sentence-level manually aligned, all translations in one translation unit
- set of stand-alone annotations from brat annotation tool[2]

[1] *www.rus-ltc.org/about.html*
[2] *brat.nlplab.org/*

# DATA2: textbooks for translation students

| | books[3] |
|---|---|
| Translation practice | 32 |
| Methodology | 6 |
| Total | 38 |

| published in Russia | |
|---|---|
| 1963-1993 | 6 |
| 1994-1999 | 4 |
| 2000-2005 | 9 |
| 2006-2010 | 10 |
| 2011-2016 | 11 |

## Major approaches to training

training focus on:

- shifts in translation
- difficult language items
- semantic categories
- text-based (speech aspects)

## Books: top level category

| | |
|---|---|
| shifts | 7 |
| difficulties | 20 |
| semantics | 2 |
| text-based | 3 |

[3] *see list of references at www.rus-ltc.org/about.html*

# Theoretical premises of error classification

Translation quality criteria:

1. fitness for communicative purpose [House, 2001, Nord, 2006]
2. fidelity, semantic accuracy
3. readability (linguistic acceptability)

Translation error is any defect in a translation in breach of the above requirements.

Methodological principles behind the classification used for annotation:

- target text-centered (TT = the object of evaluation)
- formative (not summative) assessment
- hierarchical structure; top-level distinction: content transfer or target language expression
- manageable number of classes

# RusLTC error classification

## content errors

1. reference
   - omission
   - distortion
   - nonsense
   - inexact
   - unclear
2. cohesion
   - theme-rheme
   - logic
3. pragmatics
   - tenor
   - field

NB! 70% of tags have Annotator's Notes

## language errors

1. lexis
   - choice-of-words
   - combinability
2. morphology
   - wordform
3. syntax
   - incomplete
   - ungrammatical
   - word order
   - preposition
4. hygiene
   - typo
   - capitals
   - punctuation
   - delete

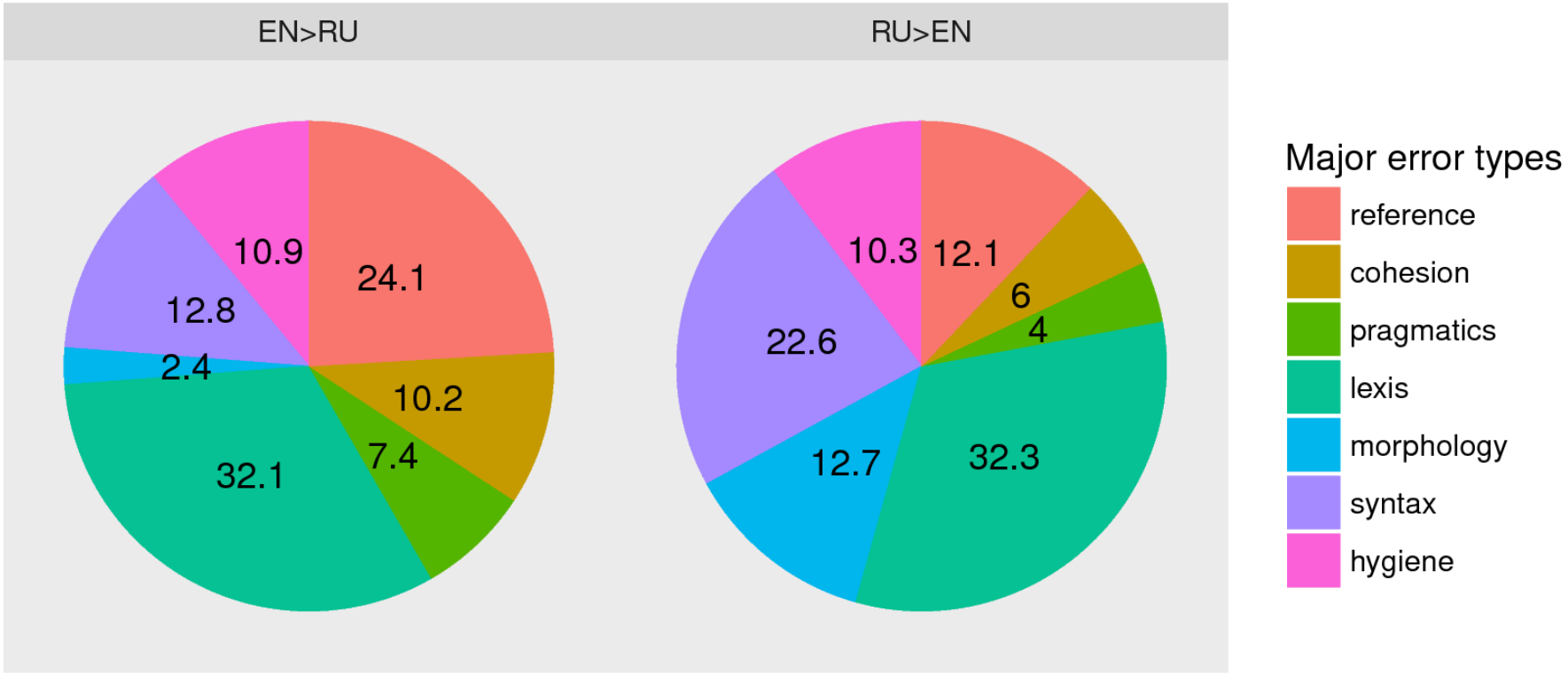
## attributes and positive feedback

- error weight (critical, major, minor)
- technology (background info, SL, TL, literal, free, proper name, inconsistency)
- Good solution!

# Direct error type statistic analysis

Error classification says what went wrong in the TT quality-wise

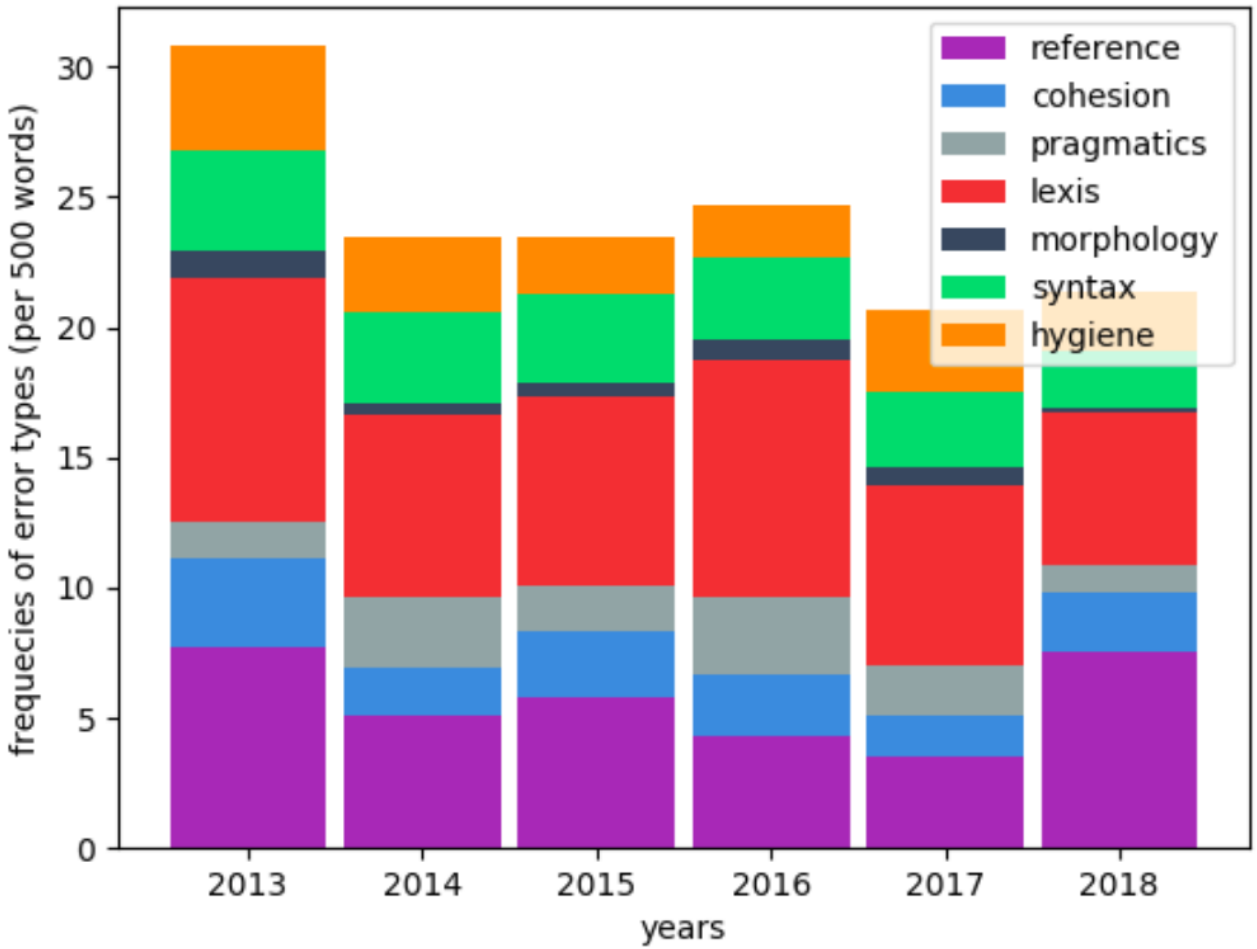
*Distibution of error types (%) in EN>RU and RU>EN*




Good for increasing awareness and indicating stages of translation that require more effort

# Track changes over years

*Ratio of normalized error type frequencies over years*




2016 third year students started taking the course *Corpora in translation practice*

# Analysis of technology attributes

*Frequency of (arbitrary) technology labels over years*

## Step-by-step

**How textbook content correlates with real translation problems?**

1. build a representative corpus of translation textbooks
2. draw up a list of the most addressed translational issues (=categories for error analysis)
3. learner data: define a set of translationally most challenging source sentences (and extract translations and error stats for them)
4. manual analysis to identify
   - error-prone ST items (that triggered inadequacies in several independent translations) and
5. compare issues featured in the textbooks and error analysis results

# Featured didactic items: approach

A didactic unit is a named shift (=translation solution) or difficulty which has dedicated exercises in the textbook
DATA: 26 textbooks with identifiable discrete didactic units
RESULT:

- an inventory of ($\approx 65$) topics discussed
- their frequency in the textbooks (how often the authors choose to include them among the preferred average of 13.5)

**Had to interpret**

the textbooks differ in

- granularity of description and names of items
- categorisation of the same items (*attributive phrases: lexis?grammar?*)
- emphasis on a given issue (*some textbooks are more universal than others*)

# The usual suspects (in the order of importance)?

**grammar**
(32 issues in 26 bks)

- complex NP (esp. N+N)
- modal verbs
- polysemantic functionals (*while, as, it*)
- passive voice
- non-finite clauses
- infinitives

**lexis and stylistics**
(19 types in 20 bks)

- false friends
- neologisms, new coinages
- phraseology
- polysemy, contextual
- proper names
- SL-specific items (*glimpse, inferiority*), inc. lexical lacunae (*cream tea, tundra*)

**shifts/solutions**
(13 types in 8 books)

- splitting
- compensation
- antonymic shift
- modulation
- word order change
- concretisation

# Filtering data on error-prone sentences

Translation unit = source sentence + all its translations

**All error annotated data used**

- $\approx 9K$ sentences
- $\approx 180K$ words
- a lot of idiosyncratic variation

**Data reduced to the top 5 most challenging source sentences**

- 160 source sentences, $\approx 2K$ target sentences
- **assumption: this sample reflects (negative) commonalities in translator's linguistic behavior**

4 times smaller data for analysis and a subcorpus of translation challenges and least adequate translations (!)

# Identifying most difficult source sentences

Sentence translation difficulty index is normalized and weighed score of errors per source sentence (ssent)

**index = (minor*1 + major*2 + critical*3)/number of targets for this ssent**

Intermediary task solved: link monolingual annotation output to sentence-aligned bi-text (TMX)

- disambiguate short annotated spans by increasing them to at least 20 char (50% improvement on accuracy of match)
- find the spans in the bi-text, extract the ssent and error statistics for its translations
- rank sentences of a ST and get the top of the list (top-5 in this research)

# Manual analysis procedure

- have in mind the textbook inventory of difficulties
- identify ST or SL items (constructions, categories) that trigger errors in several (usually more than half) independent translations in each of 160 sets
- exclude sets where translators happen to fail on different "rich points" [Beeby et al., 2005], though it can be a sign of cognitive overload [Ovtchinnikova and Pavlova, 2016]

# Manual analysis results: top 20 most frequent triggers

| trigger | cases |
|---|---|
| complex noun phrases | 25 |
| non-human S as agents | 22 |
| theme-rheme (FSP[4]) | 21 |
| cliché | 20 |
| nominalisations | 19 |
| terms | 15 |
| contrastive combinability | 15 |
| compression | 14 |
| complex sentences | 13 |
| SL-specific lexis | 13 |

| trigger | cases |
|---|---|
| infinitives | 10 |
| detailed descriptions | 9 |
| word order | 9 |
| polysemy/contextual | 9 |
| proper names | 8 |
| figurative speech | 8 |
| discourse markers | 8 |
| modal verbs | 7 |
| passive voice | 5 |
| plural of nouns | 4 |

[d][Firbas and Jan, 1992]

# Examples

## (1) multi-component attributive noun phrases

- **harsh, increasingly insecure economy**
  (1.1) ... в этой суровой и всё более опасной экономике...; (1.2) ... жесткой и незащищенной экономики...; (1.3) ... в условиях слабой и все более ненадежной экономики ...
- **deep-thinking intellectuals**
  (1.4) глубоко-мыслящих людей; (1.5) глубоких мыслителей; (1.6) глубоко мыслящего интеллектуала; (1.7) умных людей
- towards **smart, sustainable and inclusive growth**
- Celinda Lake, **a Democratic pollster**, who ...
- his **mad picture-book mind** (about Tim Burton's cinematography)

# Examples, cont

## (2) Non-human subjects of action verbs

**Saying that volcanoes toppled the Egyptians is obviously untrue.**
(2.1) ... вулканы свергли правящую династию ...; (2.2) ... именно вулканы повлияли на падение ...; (2.3) ... что вулкан изжил...; (2.4) ... вулканы погубили...; (2.5) ... вулканы разрушили Египет ...; (2.6) ... Египет пал из-за вулканической активности...

**The death and destruction wrought by these weapons was unprecedented and might have, in another world with another race of beings, ended the nuclear threat right then and there**
(2.7)Смерть и разрушение, вызванные этим оружием, были беспрецедентны и должны были положить конец ядерной угрозе еще тогда.

# Examples, cont

## (3) cliché

- inability to recognize the cliché, hence nonsensical or whimsical renditions
  ... US youth were in favour of volunteering, but **their voting record was clear**
  (3.1) ... американская молодежь выступает в пользу взаимопомощи, но **по результатам голосования все предельно ясно**; (3.2) ... поддерживает волонтерское движение, но их **выбор был очевидным**; (3.3) ... молодежь за любой движ, кроме голодовки, но к выборам это не относится
- literal rendition is understandable, but lacks idiomaticity
  Then I found one which **became my bible** for the whole of 1982
  (3.4) Мне посчастливилось найти талмуд, который стал моей библией ... :-)

## (4) Functional sentence perspective (theme-rheme structure, inc. with marked rhemes)

- Ancient Egyptian society saw its fair share of revolts and conquests, but a new paper hints that a surprising force may have been meddling in the affairs of the time.
- Walkers partnership with Gillian Cooke showed promise last season, but following intensive testing in training over the summer months Rebekah Wilson has now joined as brakeman in the womens lead two-man crew.
- There are several hundred well-documented cultures in the world, most of them belonging to major nation-states
- It's now widely accepted that rising household debt helped set the stage for our economic crisis
- Evo Morales, the president of Bolivia, was also present - a testimony to his commitment and leadership to this critical agenda.

# Examples, cont

## (5) polysemy (including need for contextual interpretation)

- **flim-flam survey data**
  (5.1) ... часть ложных данных ... (5.2) ... по некоторым данным шуточного опроса ... (5.3) ... по некоторым данным любительского опроса... (5.4) ..., но они не внушают доверия
- **people** > frequency calque
- **failed to deliver**
- **ensure/make sure**
- **wrong**
- **assumption**
- **funny** as in *I hunted out the longest titles and the authors with the funniest names, I scoured the library for completely unread books.*

# Towards a new syllabus for a practical course

There is a theoretical component that overviews all observed and typical difficulties

**grammar**

- complex NP
- theme-rheme
- non-human S as agents
- syntactic complexity, inc. "wise" (not crude force) decompression
- word order
- nominalisations

Praise more!

**lexis and stylistics**

- contrastive combinability
- differences in representation as to detailisation
- phraseology esp. non-figurative, cliche, bundles
- polysemy, contextual
- proper names
- terms and SL-specific items (with emphasis on the

# Translation error annotation in use

- Error annotation describes translational product and can be used for relative assessment and to diagnose translation competence
- Error types statistics do not explicitly reflect translational difficulties
- Errors can be used to identify translationally challenging source sentences for class feedback and further analysis
- Frequency lists and concordances for annotator notes were not helpful in our system, because they are too contextualized

## Outlook: formalize analysis

- on ST side: mark-up known triggers. Do they correlate with the probability of a specific error type?
- on TT side: mark-up TT items that give away typical errors, given a ST form
- Is there a correlation between errors and known translationese features?

# Thanks for helping me

- learn to manipulate error data
- revise our approach to translation education
- get some knowledge from error-annotation

## Questions?

Maria Kunilovskaya
mkunilovskaya@gmail.com
Russian Learner Translator Corpus

# References I


- Beeby, A., Fernández Rodríguez, M., Fox, O., Kozlova, I., Neunzig, W., Presas, M., Rodríguez-Inés, P., and Romero Ramos, L. (2005). Investigating translation competence: Conceptual and methodological issues. *Meta: journal des traducteurs*, 50(2):0609–619.
- Castagnoli, S., Ciobanu, D., Kunz, K., Kübler, N., and Volanschi, A. (2011). Designing a Learner Translator Corpus for Training Purposes. *Corpora, Language, Teaching, and Resources: From Theory to Practice*, (12):221–248.
- Firbas, J. and Jan, F. (1992). *Functional sentence perspective in written and spoken communication*. Cambridge University Press.

# References II


- House, J. (2001). Translation Quality Assessment: Linguistic Description versus Social Evaluation. *Meta: Journal des traducteurs*, 46(2):243.
- Leech, G. N. (1997). Introducing corpus annotation. In *Corpus Annotation: Linguistic Information from Computer Text Corpora*, pages 1–18.
- Nord, C. (2006). Translating as a purposeful activity: a prospective approach. *TEFLIN Journal: A publication on the teaching and . . .*, 17(2):131–143.
- Ovtchinnikova, I. and Pavlova, A. (2016). *Translational bilinguism. Based on translation error analysis*. Moscow: FLINTA, Nauka.