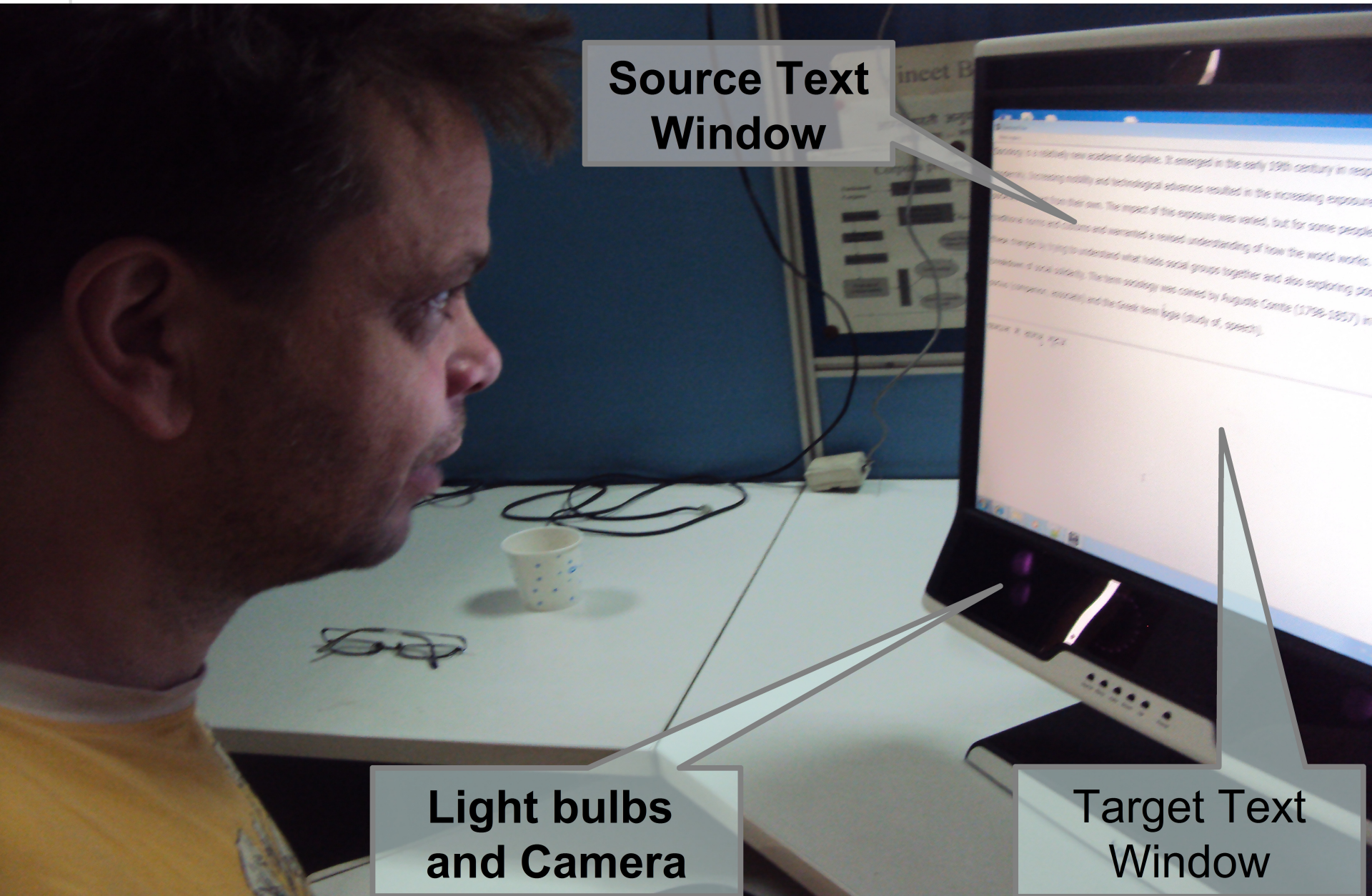


Human Translation Process Research and the CRITT TPR-Database

Michael Carl

Center for Research and Innovation in Translation and Translation Technology
Institute for Business Communication, Copenhagen Business School
Denmark

- Translation Process Research at CRITT
 - Data Acquisition with Translog-II
 - Translation styles professional/novice translators
- Vertical and horizontal Translation
 - Alternating vs. parallel translation activity
 - Translation time and writing segments
- Post-editing and the TPR Database

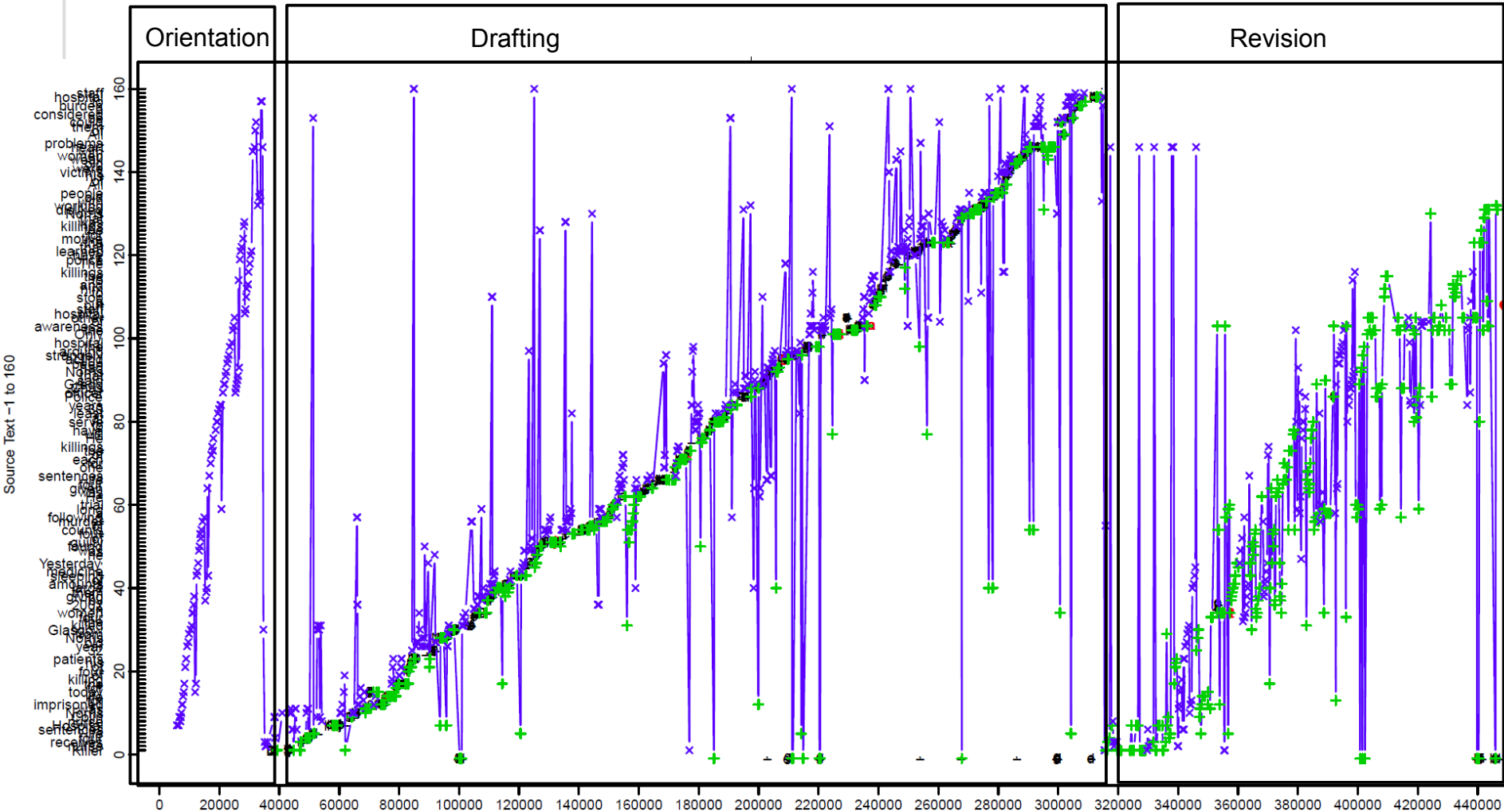


Source Text Window

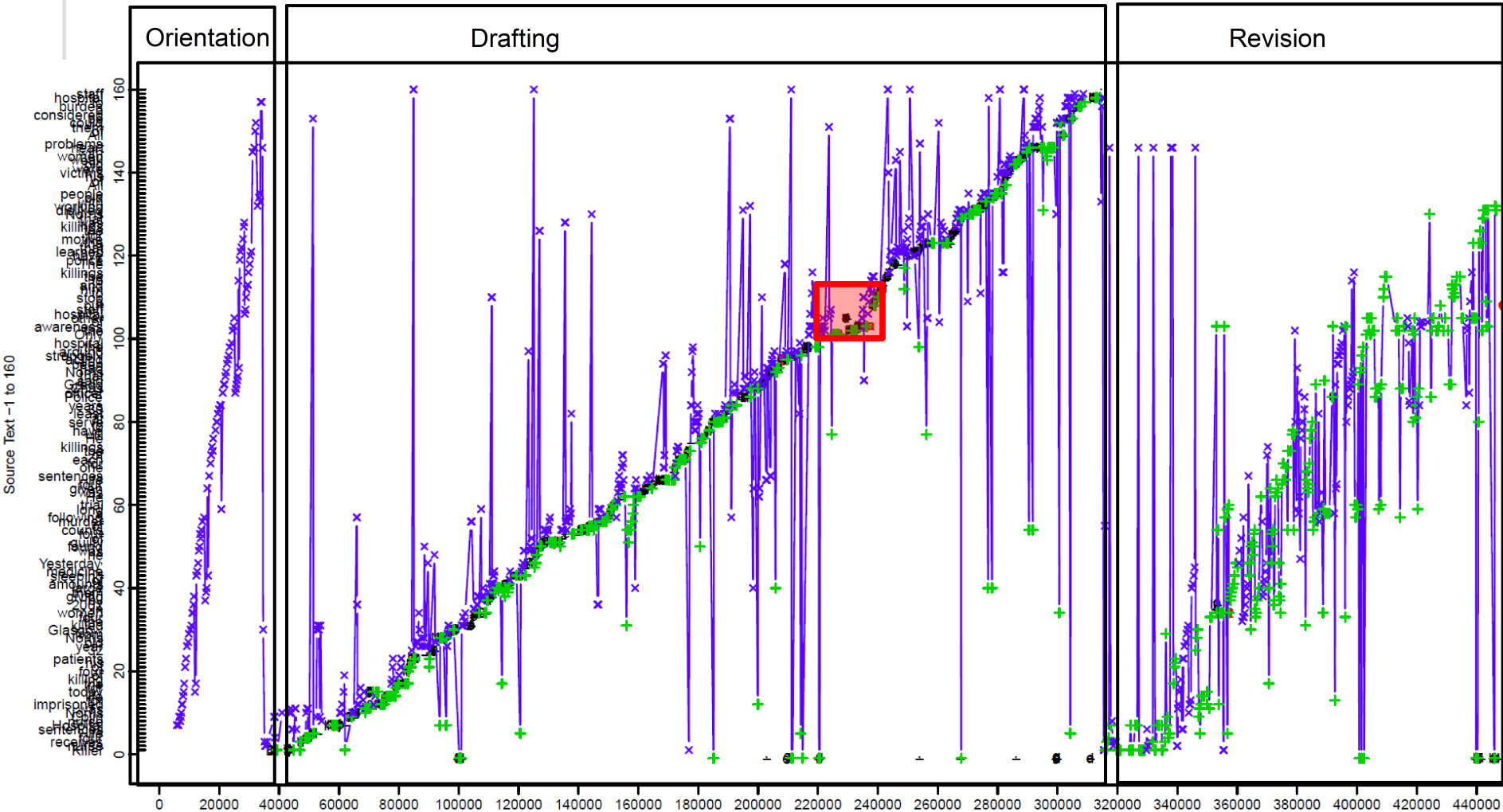
**Light bulbs
and Camera**

**Target Text
Window**

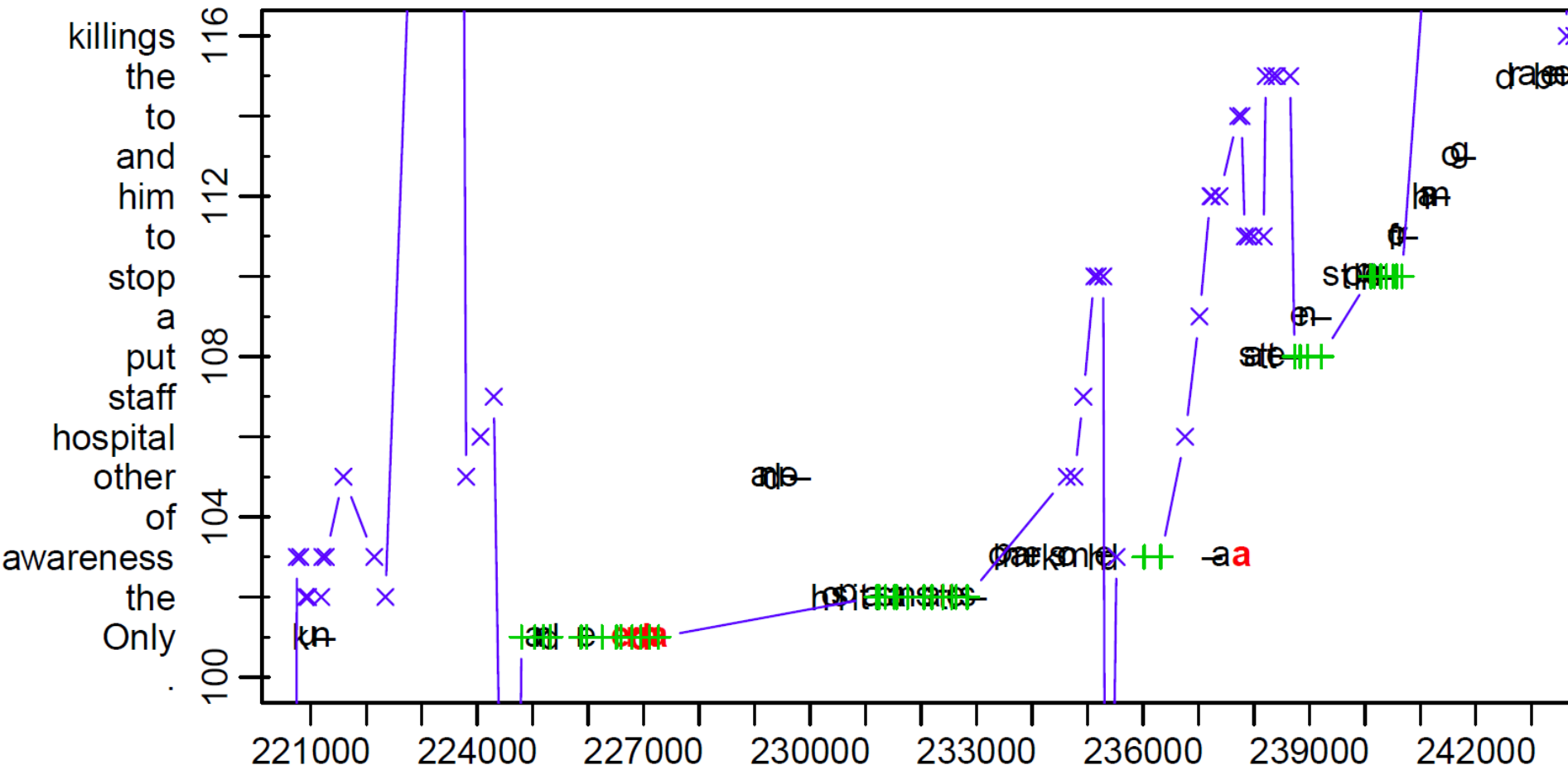
Translation Progression Graphs



Translation Progression Graphs



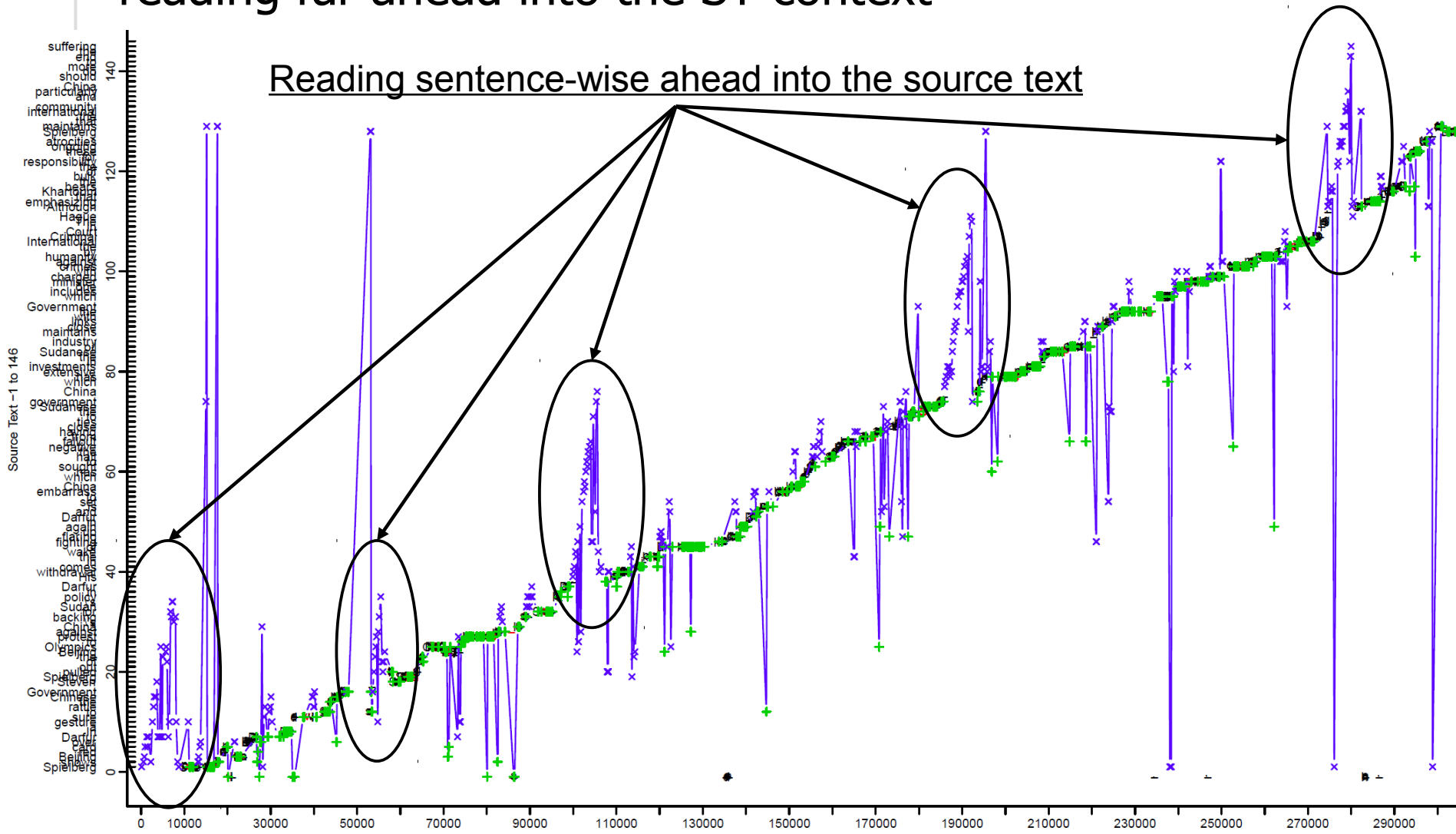
Translation Progression Graphs



Large Context Planner:

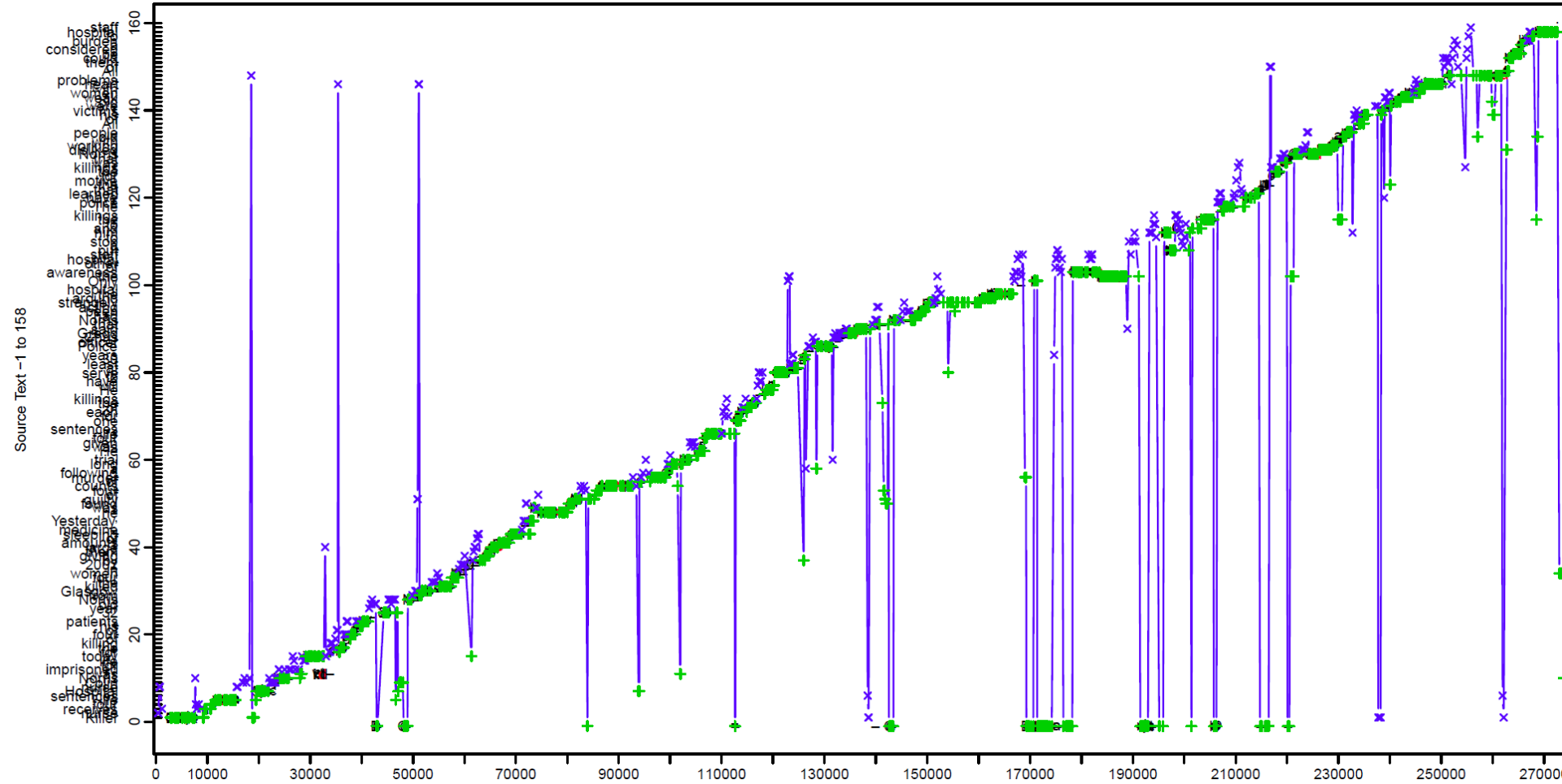
reading far ahead into the ST context

Reading sentence-wise ahead into the source text

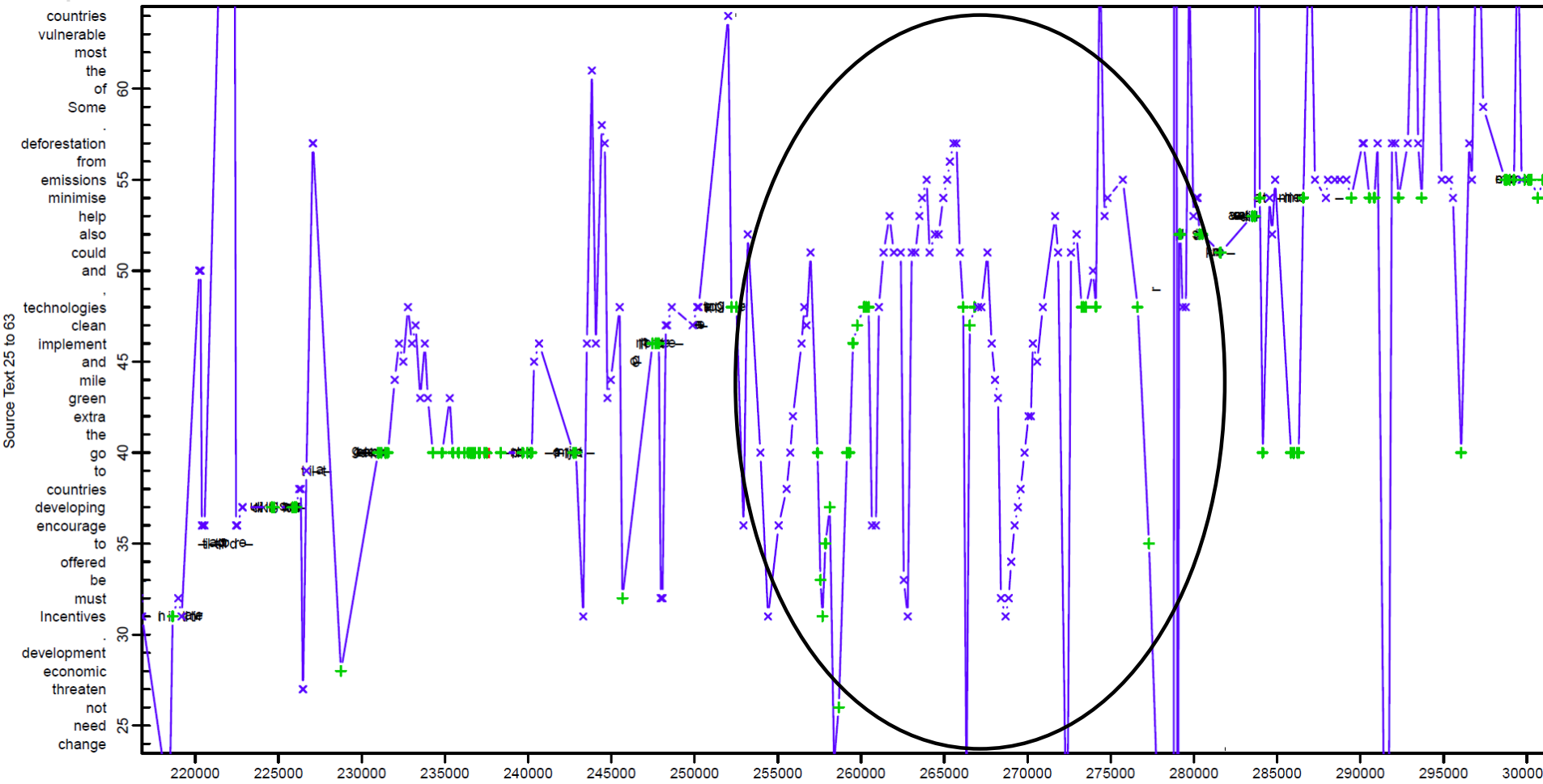


Small Context Planner:

looking only few words ahead in the ST



Back-tracker: looking into translated context



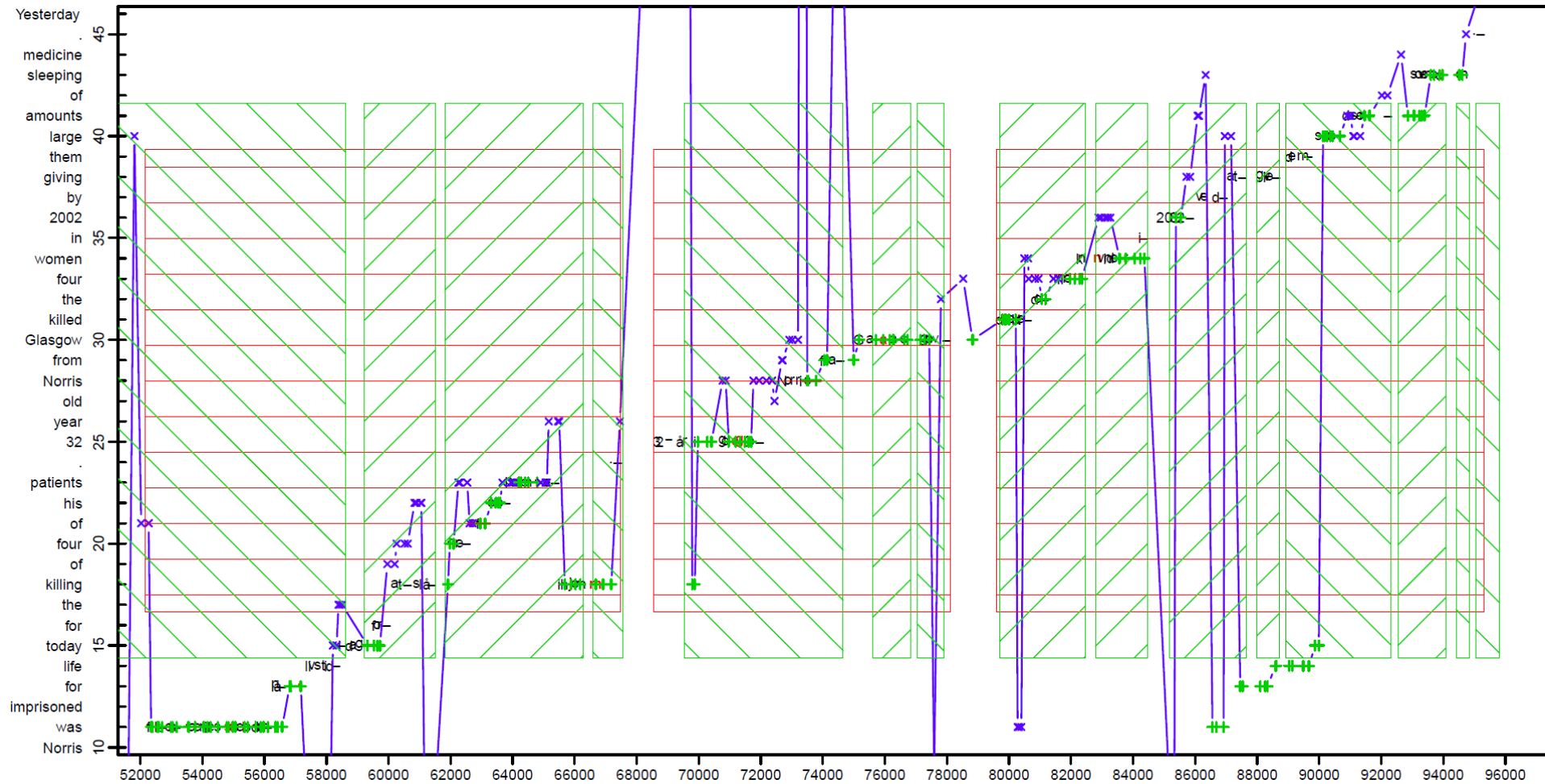
Translation Styles

- Initial orientation:
 - systematic orientation vs. skimming
 - quick planning vs. head start
- Translation drafting:
 - large-context vs. small-context planning
 - backtracking vs. non-backtracking
- Translation Revision:
 - online revision vs. end revision
 - constant revision

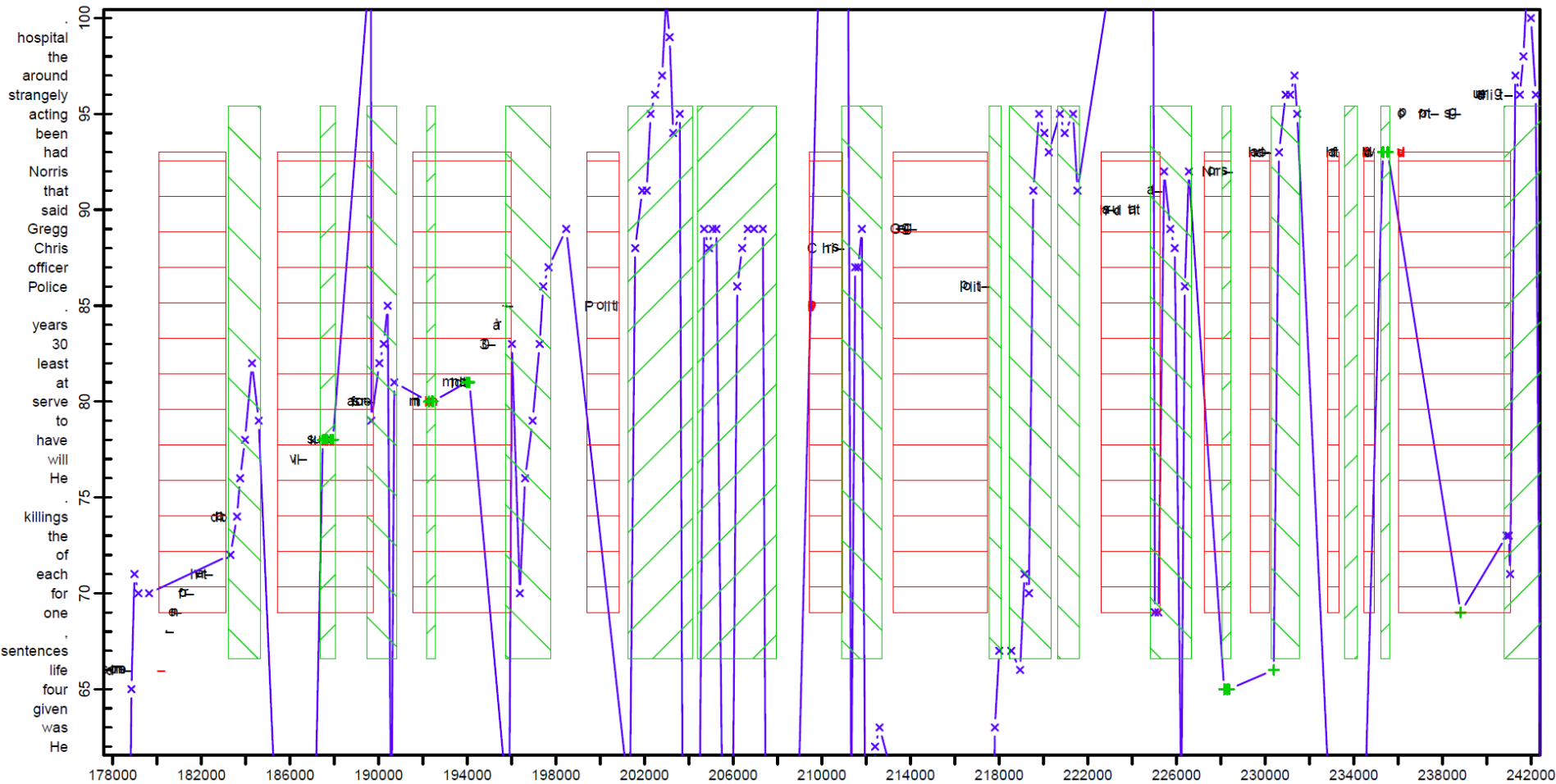
Professional vs. Novice Translators

- Professional translators:
 - head-starting
 - small-context planning
 - end-revising
- Student translators:
 - systematic initial orientation
 - large-context planning
 - online revising

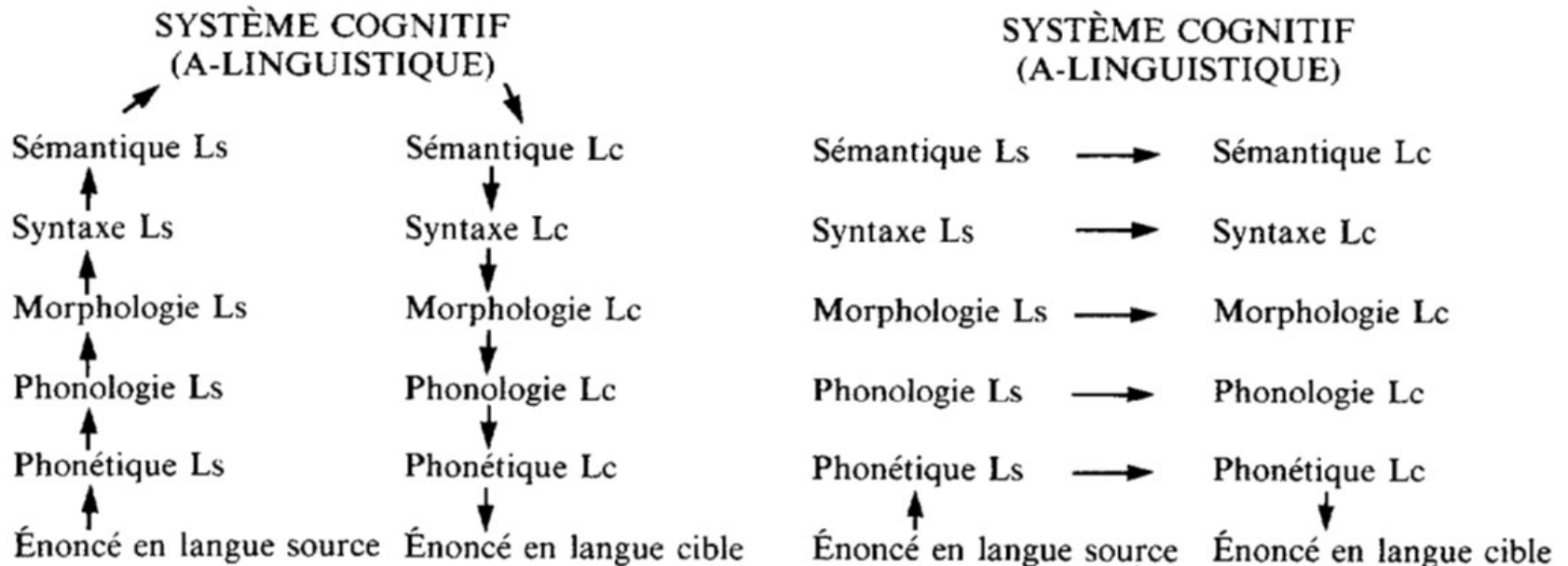
Parallel ST Reading during TT Production



Alternating Translation Production



Vertical and Horizontal Models of Translation



Translation = ST Reading + TT Writing + X

X = problem solving

X = activation of shared memory representations

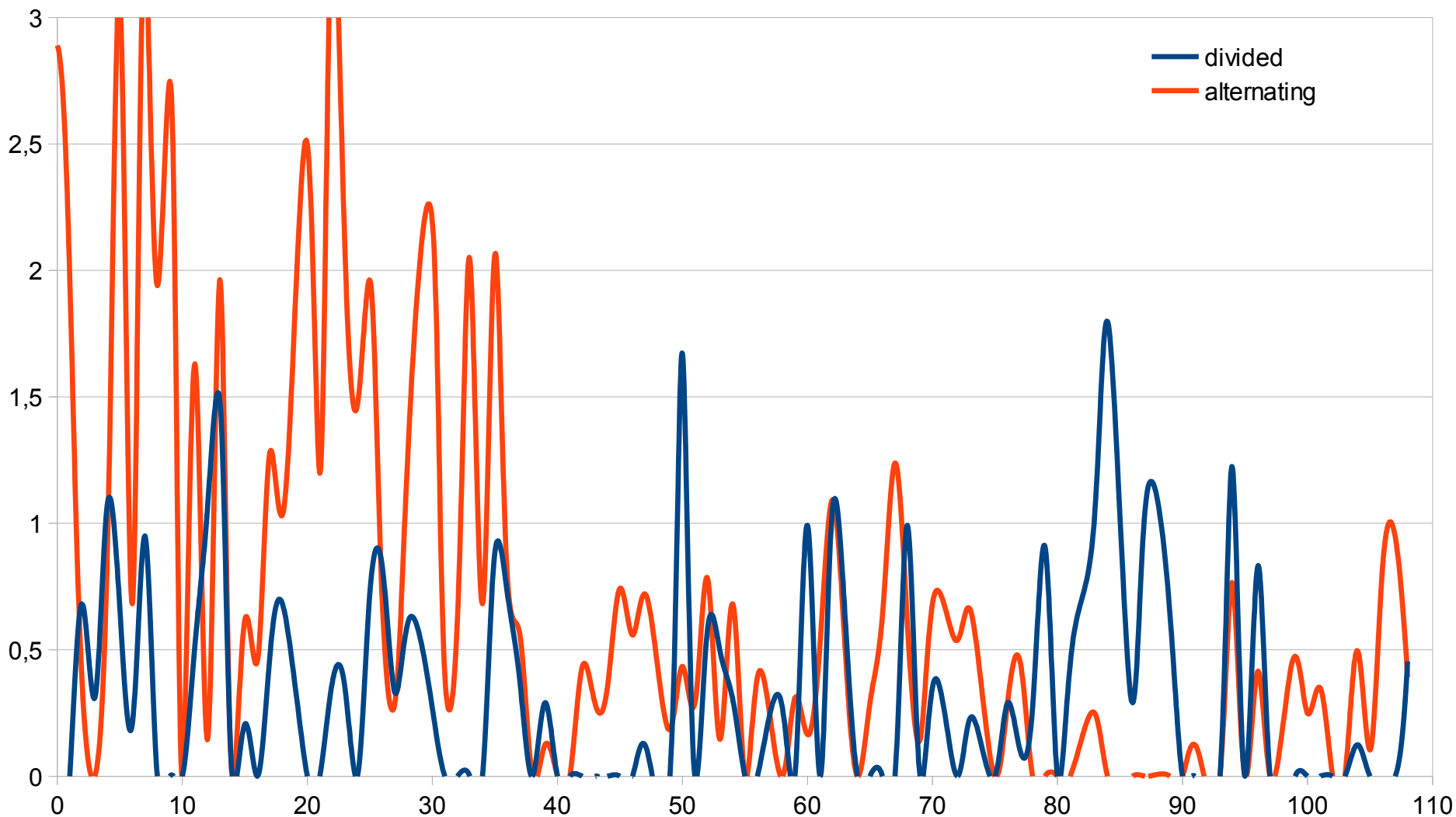
The Monitor Model

- Translators start in a literal default rendering mode
- A monitor interrupts the default procedure when a problem occurs

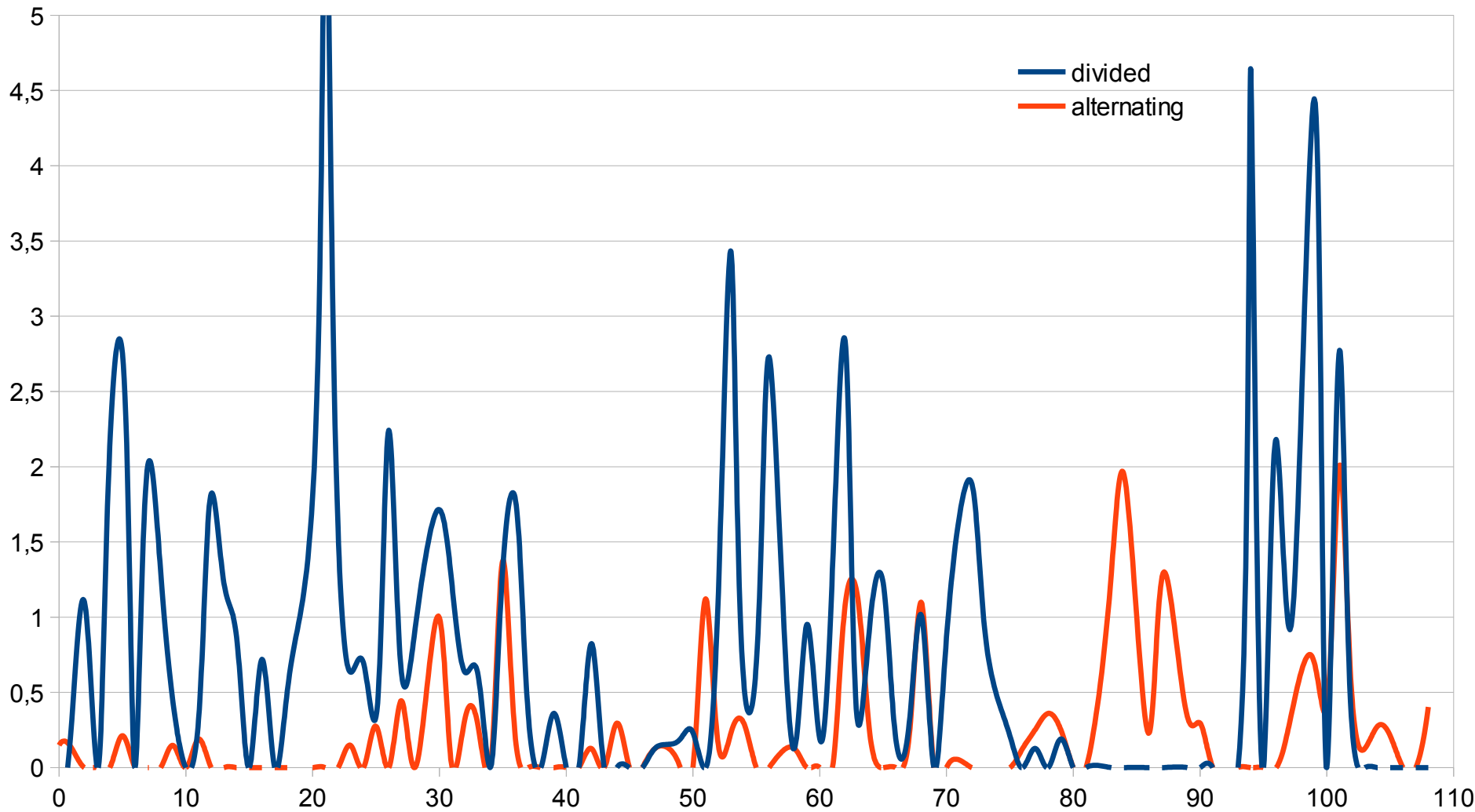
During the literal translation mode, ST primes the translation:

if particular features (phonetic/syntactic/semantic) exist in both languages, then they are likely to prime across languages

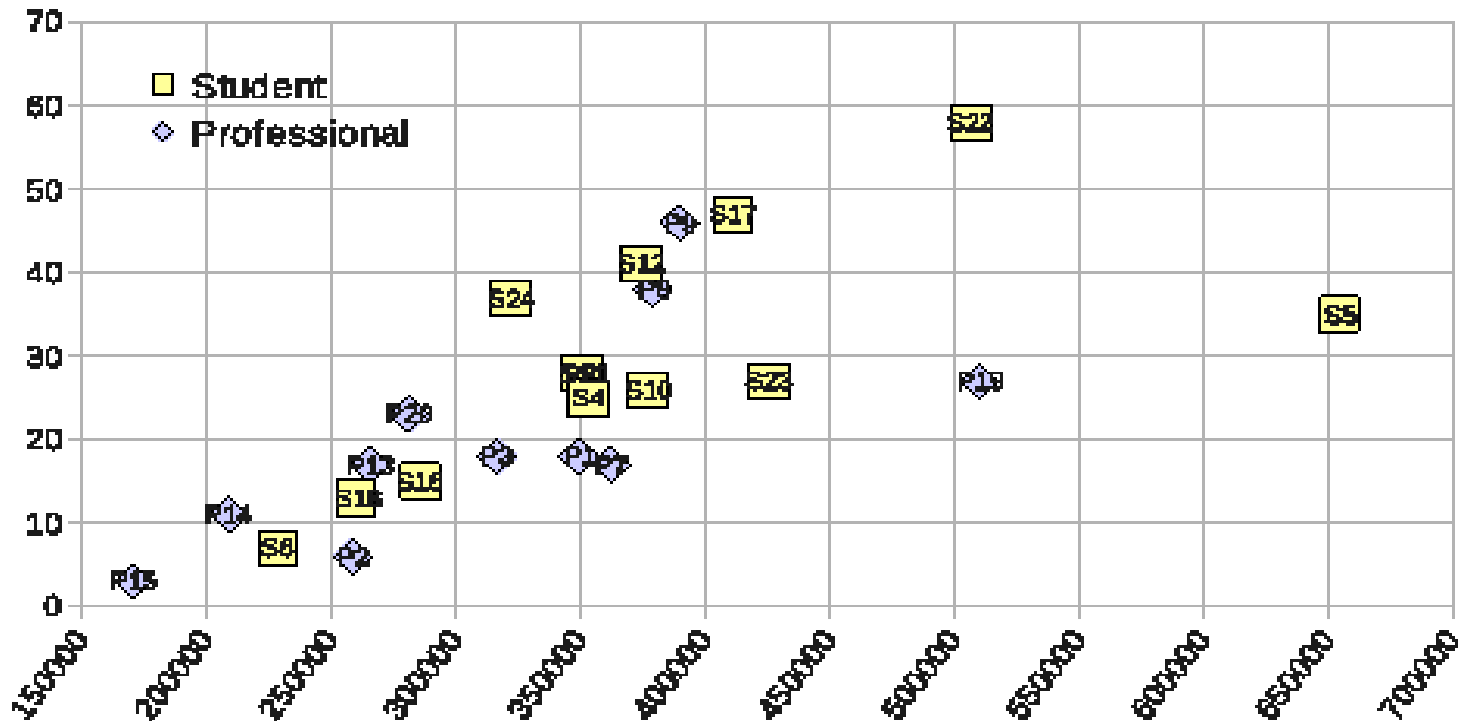
Parallel/Alternating Fixation Distribution



Parallel/Alternating Fixation Distribution



Translation Time and Number of alternating FUs for Professional and Student Translators

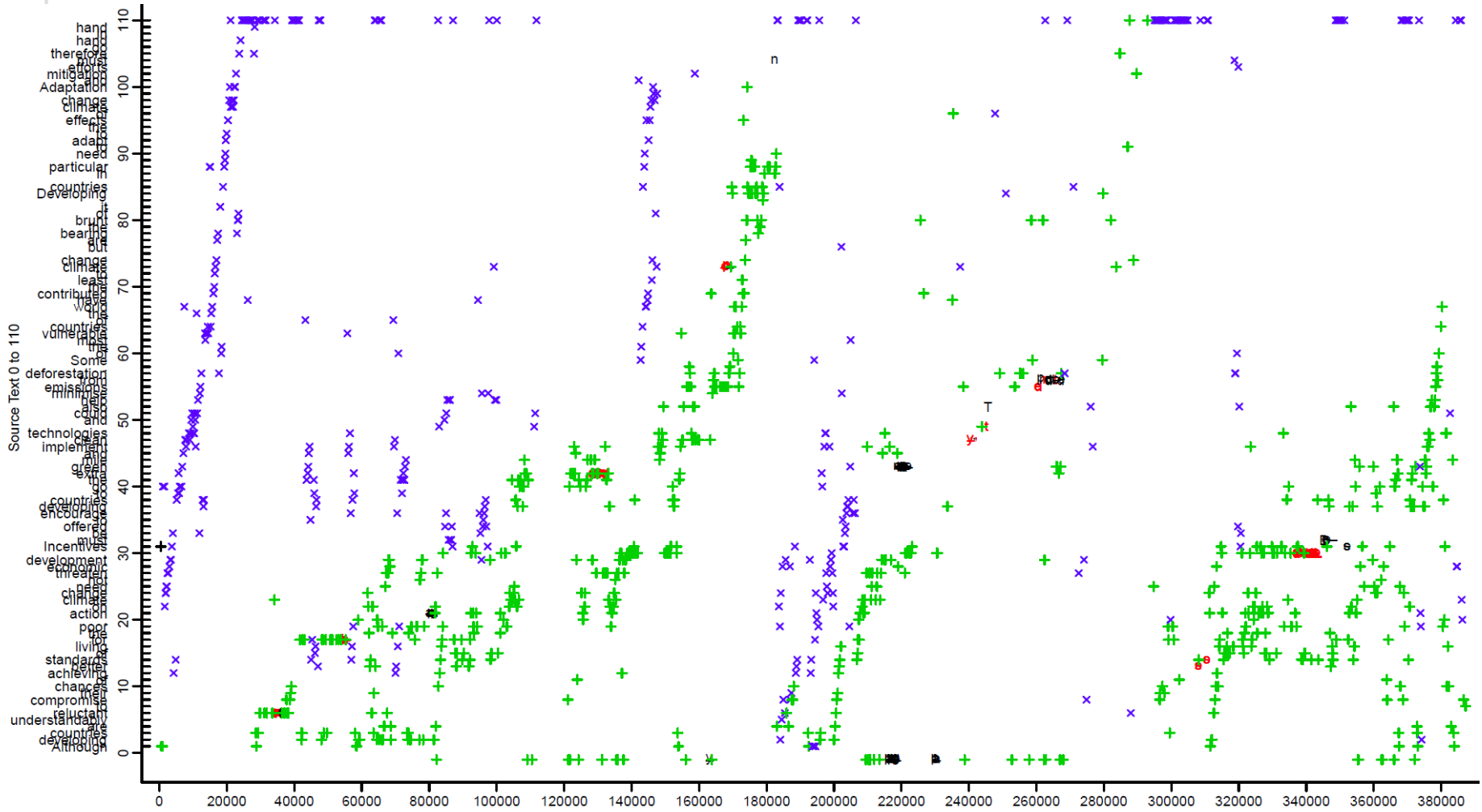


Average PU duration, length and typing speed of Professional and Student Translators

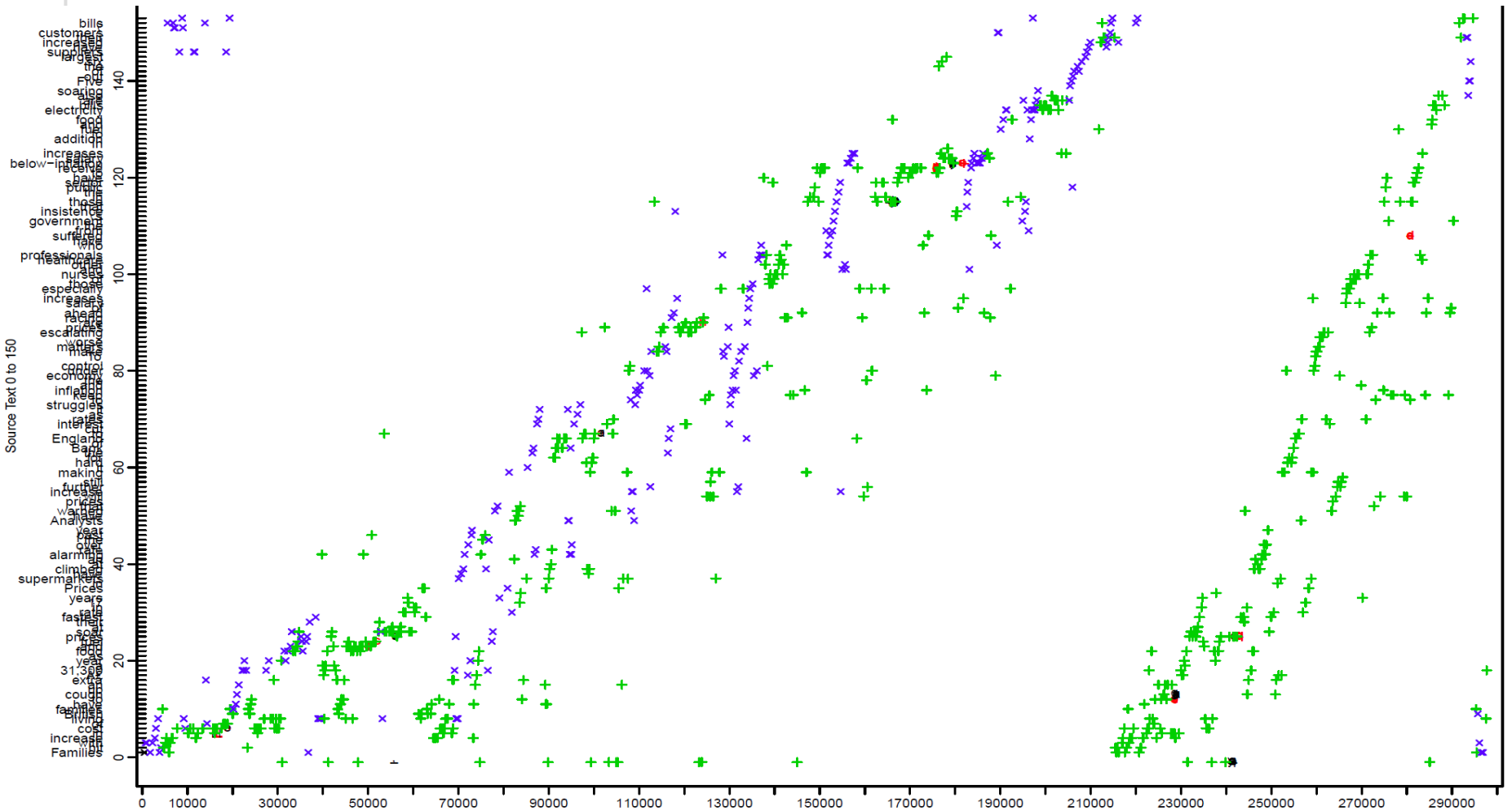
	Professionals	Students
av. PU duration in ms	3113	2216
av. PU length in chars	17.61	12.54
av. (PU chars/PU dur.)	5.44	5.70
median (PU chars/PU dur.)	5.61	5.53

What about Post-editing?

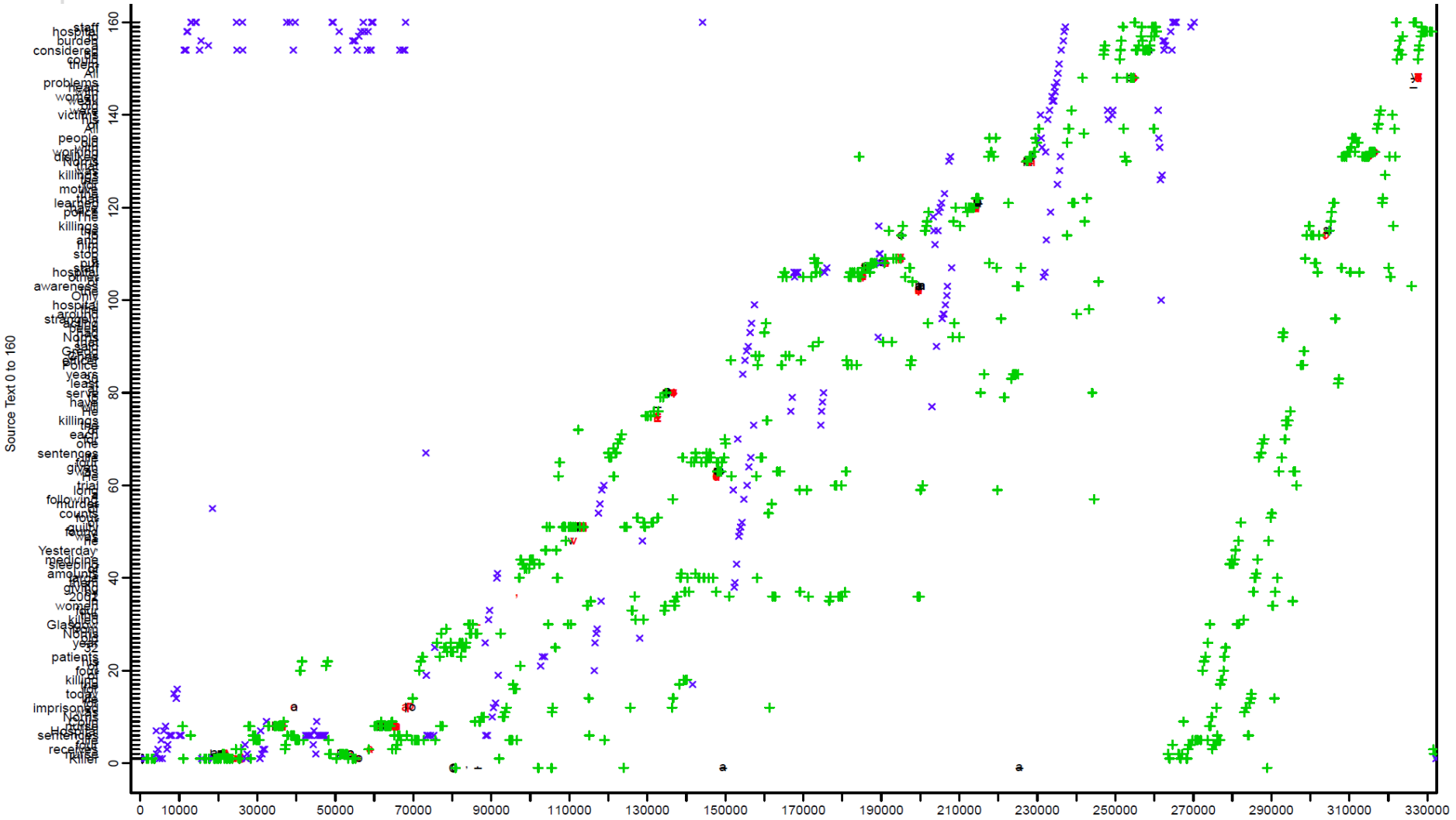
Unexperienced post-editing



Post-editing: reading ST; checking TT



Post-editing: reading TT; checking ST



What are Post-editing styles?

- Novice vs. experienced post-editors
- Impact of post-editing briefing
- Attention in post-editing (ST vs. TT)
→ TPR Database

What are appropriate Post-editing tools/

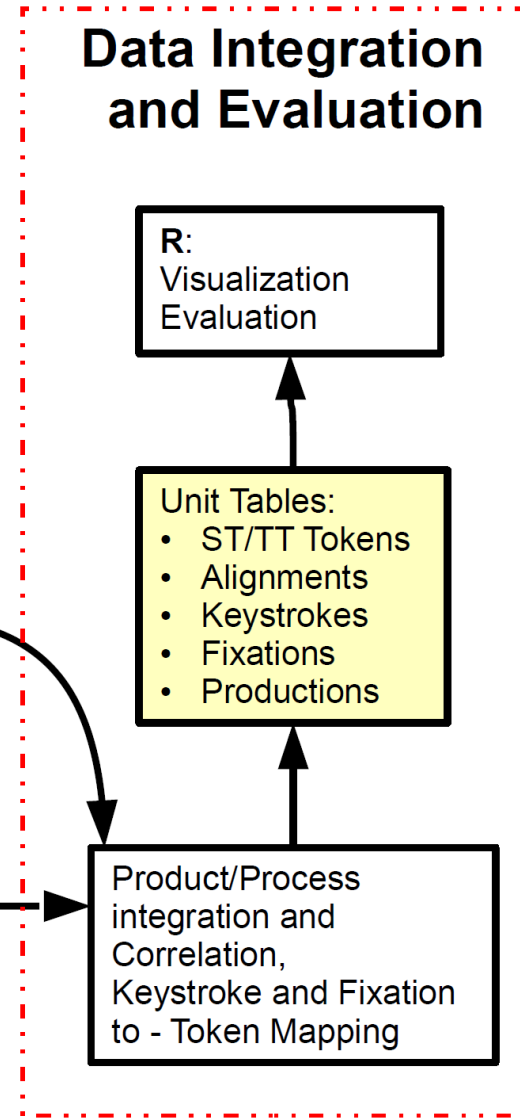
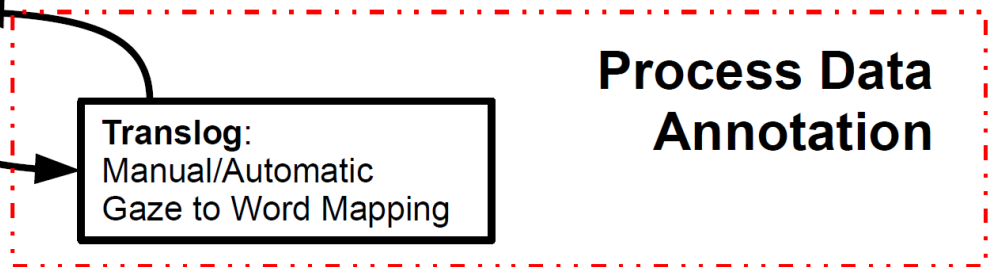
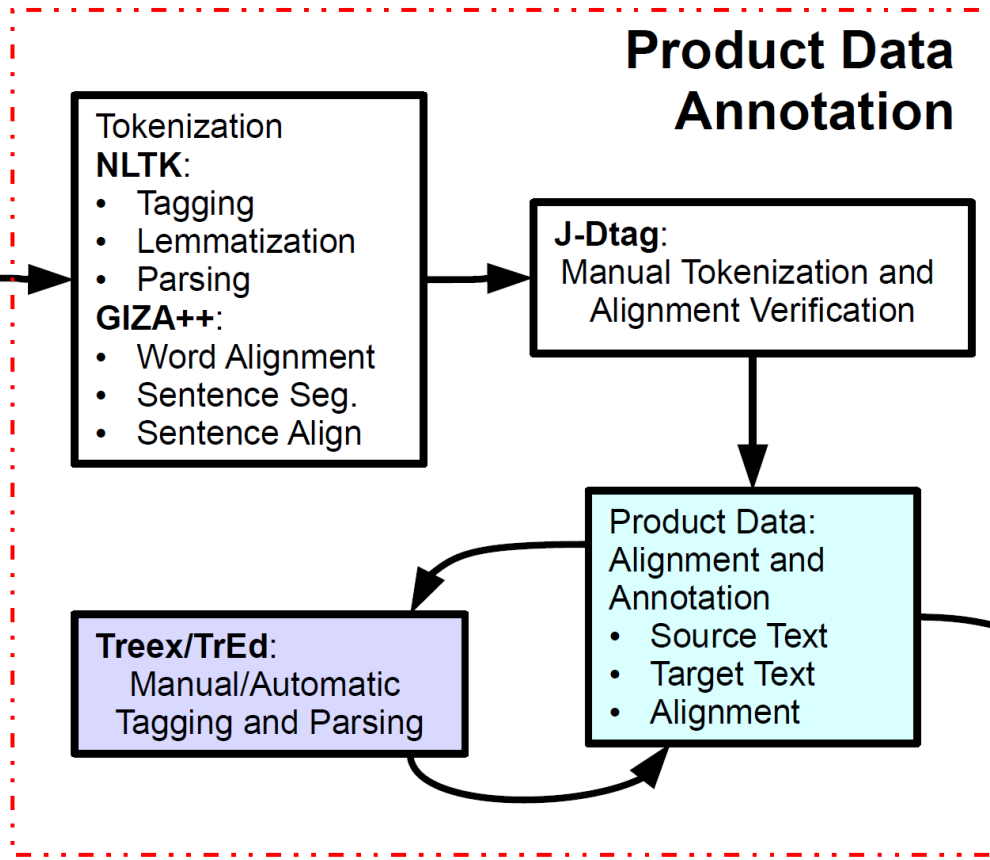
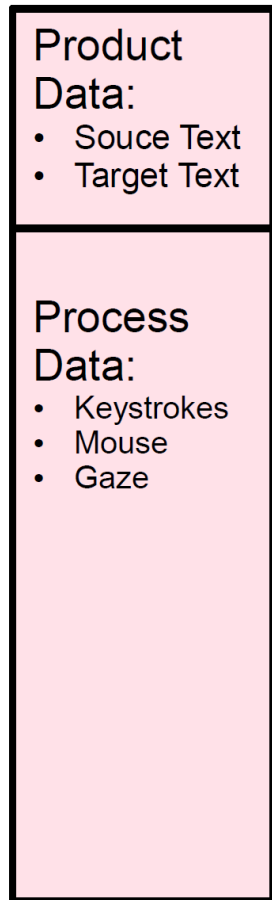
- Different tools for different post-editors
→ Casmacat

How can post-editing be taught?

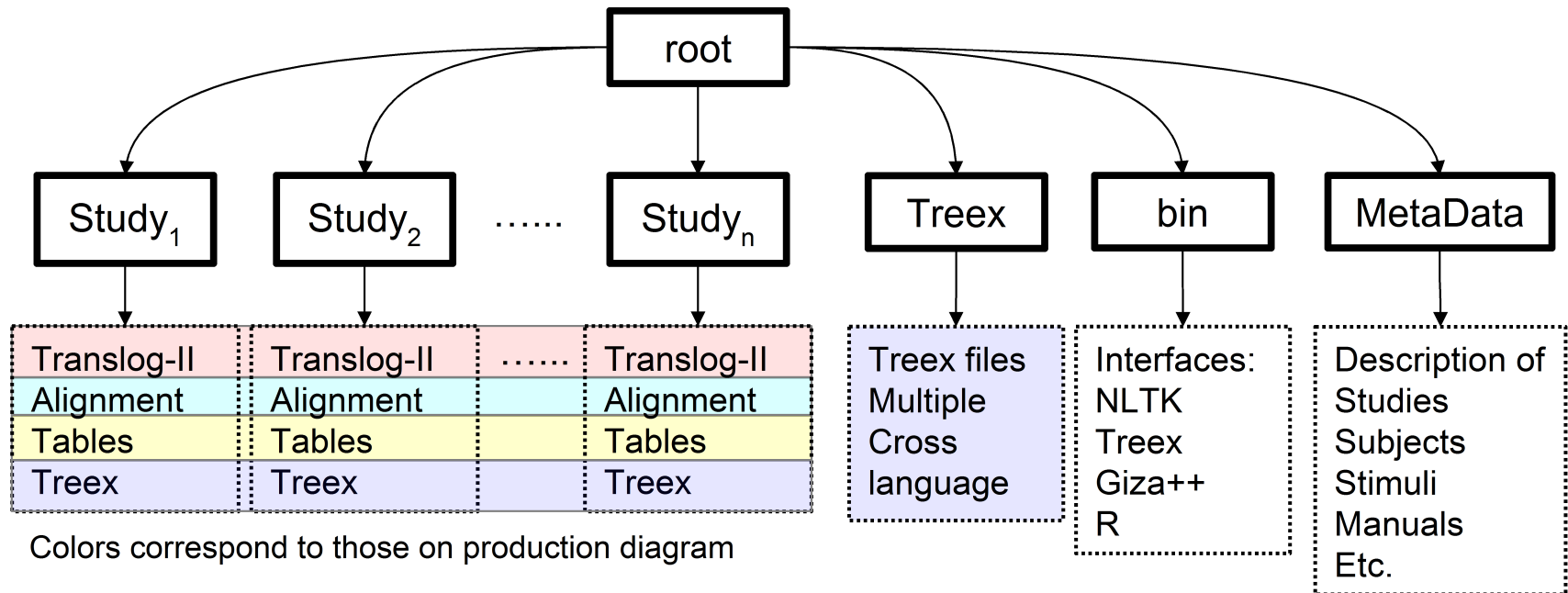
Source	Target	Transl.	PE	Edit	Copy	Total
en	da	111				111
en	hi	39	61			100
en	es	20	20	20		60
en	zh	15	19	10		44
en	de	12	19	10		41
da	en	40				40
en	en				30	30
en	pt	10	10			20
pt	en	10				10

TPR-DB Annotation and Compilation

Logged UAD:



TPR-DB Architecture



Study folder correspond to an experiment, containing sub-folders:

- Translog-II: recorded logging data of translation/reading/writing session
- Alignment: annotated and aligned product data for each recording
- Tables: extracted features for tokens, fixation, productions and alignment units
- Treex: sentence-aligned ST-TT correspondencies for each recording

Treex folder contains multi-lingual aligned representations of the product data

MetaData folder contains description of translation studies, stimuli and translators

CASMACAT: A tool for PE research

Thanks

CRITT TPD-DB V1.0 downloadable from: <https://dl.dropbox.com/u/7757461/TPR-DB.zip>

WPTP workshop paper: <https://dl.dropbox.com/u/7757461/TPR/CRITT-TPR-DB1.0.pdf>

Casmacat Demo: <http://130.226.34.13/prototype1/login.php>

User: test Password test