



UNIVERSITY
OF TAMPERE

Kotus Experience: Crowdsourcing the Old Literary Finnish for the Research's Benefit

OCR Webinar

5.3.2014

Minna Vanhasalo

Finnish Language

School of Language, Translation and Literature studies

minna.vanhasalo@uta.fi



Institute for the Languages of Finland

Kotimaisten kielten keskus (Kotus)

- Some history and background (Kuutti 2014)
 - a need for even better corpuses for research's benefit
 - on 2010 dozen of old texts from the Old Literary Finnish era were digitized (ocr-format)
 - lots of mistakes on the material -> unreliability -> demand of extensive proofreading -> resource?
 - collaboration with The National Library of Finland was initiated (Toni Suutari)



Institute for the Languages of Finland

- Goals of collaboration (Kuutti 2014)
 - crowdsourcing as a resource for labor intensive proofreading
 - through crowdsourcing usable and reliable corpus
 - for research in general (linguists and historians)
 - for Dictionary of Old Finnish
 - for students of Finnish Language
 - for interested laymen



Collaboration with University of Tampere and Kotus

- Goals of collaboration from teacher's perspective
 - a need for developing teaching of Old Literary Finnish
 - a need for getting better learning results
 - a desire to change attitudes towards Old Literary Finnish



Before: Low motivation, low learning results (E&F 2013)

- Experiences from the previous courses
 - course of old literary Finnish was considered more of as something difficult, boring, useless, and less of as something interesting, useful, easy and worth learning
 - low motivation for the course
 - low learning results: superficial learning, lots of mistakes and misconceptions, low grades
 - grading system 1,2,3,4,5 (5 = excellent)



After: High motivation, great learning results

- Profound learning, in-depth analyses of phenomena of old Finnish
 - grades have shifted from the lower end to the higher end: 1,2,3,4,5
- Positive attitudes towards the Old Literary Finnish -course and its contents
 - I really enjoyed proofreading the material, because as I was doing it I learned a lot of old literary Finnish.
 - The contents of teacher's lectures started to make sense after seeing the same phenomena on the edited text.
 - Even if there were problems with the editor, I really enjoyed working with it. Now I feel confident I can read old Finnish.
 - After an exam one easily forgets everything learned. I don't think I will



But I got more than I was asking for... (SRA 2014)

- Students perceptions of themselves as learners and as members of the academia seem to be evolving to the right direction!
- According to student counsellors number of first year students share common concerns about themselves as university students (Eerola 2010)
 - Am I competent enough? Do I belong to the university?
 - Will someone notice that I am actually a fake student, and I got into university by chance and luck?



UNIVERSITY
OF TAMPERE

Students' experience and opinions on working Erici Ericus Sorolainen's Postilla (1621) with OCR-editor (SRA 2014)

- In the beginning of the course I was startled that they are actually going to let first year students work with this kind of material for true and authentic scientific purposes, then I was happy that I could do it.
- When I heard about the exercise, I was a bit worried if I can do it. Now I hope they will do this always on this course, because it wasn't that difficult and I learned so much.



Benefits for the research and for Kotus

- 200 pages of Postilla has been proofread and re-edited.
 - 5 pages / student
 - could be more
 - number of mistakes made by students in proofreading: 0 - 2 / page, (average being 1)
- Hands-on experience on pro's and con's of crowdsourcing.
- Ideas on what other features are needed from the editor.



UNIVERSITY
OF TAMPERE

First it was Crustus, now it is Christus

Students of the Old Literary Finnish 24.2.2014





References

- Eerola, Satu (2010) Yliopistopedagogiikan perusteet 1 -kurssi. Luento.
- Kuutti, Pirkko (2014) Personal communication. E-mail on 28.2.2014.
- Study Reports and Analyses (2014) On the course Old Literary Finnish. (SRA)
- Exams and Feedback by students (2013) On the course Old Literary Finnish. (E&F)