Annotation as a Didactic Means

Heike Zinsmeister, Melanie Andresen, Fabian Barteld, Jiajin Feng, Johanna Flick, Renata Szczepaniak

Motivation

The term annotation is a polysemous notion. In linguistic contexts, it denotes (i) the process, in which we enrich text by systematically adding interpretative information to it, (ii) the added information in terms of conceptual classes, hierarchies and other relations that model some linguistic phenomenon, and (iii) the digital representation of the added information (cf. Leech 1997, Kübler & Zinsmeister 2015).

In this poster, we explore annotation in all of its three dimensions from a didactic perspective.

The process of annotation is detailed as an iterative procedure, which enables students to conceptualize linguistic phenomena in a systematic way. The development of conceptual classes in terms of an annotation tagset is done either inductively by inducing generalizations and classifications from the data, or deductively, based on a given annotation scheme or theory. Even if annotation can be done with paper and pencil, annotation tools allow for an efficient realization with a sustainable outcome in class.

Study: Inductive Annotation of Causality I

The goal of this task was for students to explore different types of causality.

(i) Die Rohre sind geplattet, weil Frost herrscht.

"The pipes burst because it is freezing." (Es herrst Frost, die Rohre sind geplattet."

Scheme: RESULT, well 'because' REASON

Task: Inductive development of an annotation scheme considering the syntax and semantics: the choice of connector, the source of coherence etc.

Causality describes a relation between two events: reason (a causing event) and result (a caused event). It is often indicated by causal markers, such as connectives (wel, desshalb…), prepositions (wegen, aufgrund…).

According to the source of coherence, we distinguish between a semantic and a pragmatic relation.

Semantic relation: two propositions are causally influenced; Pragmatic relation: the reason justifies the claim of the speaker or explains the speech act itself. (cf. Breindl & Walter 2009)

Study: Inductive Annotation of Causality II

Setting

• 25 linguistics students (B.A. and M.A., including teacher training)
• 90 minutes: group discussion (4-6 students) and individual work
• Paper and pencil
• Prior knowledge from previous classes: different layers of annotation; discourse annotation with Rhetorical Structure Theory (RST)

Table suggested for the initial scheme:

<table>
<thead>
<tr>
<th>Label</th>
<th>Definition</th>
<th>Cues</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Familiar or causality</td>
<td>$\exists$</td>
<td>$\exists$</td>
</tr>
<tr>
<td>S2</td>
<td>Task important is undifferentiated</td>
<td>$\exists$</td>
<td>$\exists$</td>
</tr>
</tbody>
</table>

Discussion

• Mixed but overall positive evaluation results.
• Introduction was too short; goal of the task was not clear enough.
• Partly confusion about actual annotation step.
• One group got completely "lost".

Steps

1. Group: Initial discussion of five examples that exemplified different readings, agreement on an initial annotation scheme.
2. Individual: Annotation of eight more sentences according to the initial annotation scheme.
4. Group: Discussion of problematic cases. Is it possible to achieve agreement?

Conclusions

→ Introducing "expert advisors".
→ Comparison between groups for more variety in annotations and discussions.
→ Relaxing time frame by reducing individual annotation load.

15 of 25 will be likely using annotation as a didactic means on their own.

The second LEA package is still work in progress. Using Sympathy as annotation tool it will cover phrases and syntactic functions.

Every LEA package contains:

• The exercise
• Information how to obtain the annotation tool
• Annotation guidelines
• A manual for students and teachers
• The sample solution
• A tool for automatic correction and evaluation

The first LEA package deals with parts of speech. The exercise comes in a tab-separated-value file that can be edited with common spreadsheet applications like Libre Office Calc. The correction and evaluation tool is created in Java and outputs helpful statistics (e.g. mean number of errors, common mistakes, a confusion matrix) for the teacher (see figure on the right).

Creating their own digitally annotated data can be a difficult task for linguistic beginners. This is especially true within the philologies where most students lack the computational expertise needed for corpus linguistics. Moreover, there is often not much time to teach the related methods.

With LEA we present a set of e-learning packages that combine linguistic exercises with easy to use annotation tools, simplified annotation guidelines, and tools for correction and evaluation. Thus, while doing their homework students learn to create sustainably annotated data.

LEA: Linguistic Exercises with Annotation Tools

Advantages for students

• Helps to understand linguistic concepts via new ways of visualization
• Introduces students to annotation tools and guidelines
• Tools to automatically analyze and evaluate students’ answers
• Ready to use in introductory courses
• Own exercises can be easily integrated

Advantages for teachers

• The development of the first LEA packages is funded by the Innovationsfonds für didaktische Forschung der Universität Hamburg (Innovationsfonds für die Lehre und das Lernen, IfD-Hamburg) 2015.

Study: Deductive Annotation of PoS with LEA

For further information please contact fabian.barteld@uni-hamburg.de

References

Breindl, Johanna Flick, Renata Szczepaniak, Melanie Andresen, Fabian Barteld. 2018. "Annotation as a didactic means on their own."


Corpora in Teaching and Research: Inductive Annotation of Causality I

For further information please contact fabian.barteld@uni-hamburg.de


Breindl, Johanna Flick, Renata Szczepaniak, Melanie Andresen, Fabian Barteld. 2018. "Annotation as a didactic means on their own."


Corpora in Teaching and Research: Inductive Annotation of Causality I

For further information please contact fabian.barteld@uni-hamburg.de


Corpora in Teaching and Research: Inductive Annotation of Causality I

For further information please contact fabian.barteld@uni-hamburg.de


Corpora in Teaching and Research: Inductive Annotation of Causality I

For further information please contact fabian.barteld@uni-hamburg.de