Analysis of verbal data

Understanding the processes of collaborative learning
Overview

- Theoretical background of CSCL process analyses
- Steps in analysing CSCL processes based on verbal data
  - Analysing individuals in small groups
  - Transcription
  - Unit of analysis / Segmentation of verbal data
  - Categorisation
  - Determining reliability
  - Automatic analysis of verbal data
- Examples
  - Analysis of cognitive processes based on think-aloud data
  - High level analyses on the base of process analyses
General research paradigm

Triangle of hypotheses:

- Specific (learning) activities are positively related with a desired outcome. (b)

- An instructional support facilitates the specific (learning) activities. (a)

- The intervention fosters the desired outcome mediated by the specific (learning) activities. (c)
Framework on cooperative learning
(O‘Donnell & Dansereau, 1992)

- Individual acquisition of domain-specific and domain-general knowledge
- Type of task
- Individual Differences
- Incentive structure
- Scripts
  - external
  - internal
- Small group interactions
- Individual acquisition of domain-specific and domain-general knowledge
Framework on cooperative learning (O’Donnell & Dansereau, 1992)

<table>
<thead>
<tr>
<th>Type of task</th>
<th>Scripts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>external</td>
</tr>
</tbody>
</table>

Blind spot without process analyses
- n→∞ Interactions of conditions of cooperative learning
- Analysis of process-based phenomena (e.g., knowledge as co-construct, internal scripts)
- Examination of process-oriented theories

Individual acquisition of domain-specific and domain-general knowledge
Text-based communication

Self-transcription of dialogues
Joint, argumentative knowledge construction: Talking, Thinking, Learning

Example coding scheme:
Weinberger & Fischer, 2006
# Granularity of segmentation

<table>
<thead>
<tr>
<th>Fine granularity</th>
<th>Theoretical relation to learning?</th>
</tr>
</thead>
<tbody>
<tr>
<td>- signs</td>
<td>How many letters do the learners use?</td>
</tr>
<tr>
<td>- words</td>
<td>How many technical terms are being used?</td>
</tr>
<tr>
<td>- speech acts</td>
<td>How do learners coordinate discourse?</td>
</tr>
<tr>
<td>- sentences</td>
<td>How do learners structure their utterances?</td>
</tr>
<tr>
<td>- propositions</td>
<td>Which concept do learners discuss?</td>
</tr>
<tr>
<td></td>
<td>What claims are being made?</td>
</tr>
<tr>
<td>- arguments</td>
<td>How do learners link concepts to construct arguments?</td>
</tr>
<tr>
<td>- argumentations</td>
<td>What standpoints are being defended?</td>
</tr>
</tbody>
</table>

| Coarse granularity | The granularity of the segmentation represents (different) types of knowledge in discourse (Chi, 1997) |
Example of Different Degrees of Fine-grainedness for Segmentation

Original messages

Jim:
The teacher attributes Michael’s failure in an internal variable manner.
She argues that Michael is just plain lazy.

Carolyn:
I don’t think so! The teacher is just making Michael feel bad.

Segmented messages

Jim:
[The teacher attributes Michael’s failure in an internal manner]
[She argues that Michael is just plain lazy. ]

Carolyn:
I don’t think so! The teacher is just making Michael feel bad.
Categorisation

- Qualitative steps
  - (Development of) categories is related to state of the art of research
  - Generating hypotheses: Paraphrasing (Mayring), Coarse analyses (Forming clusters)

- Development of a coding scheme
  - Exhaustive: Every segment is being coded
  - Exclusive: Only one category applies per segment per dimension
  - Documentation of rules, e.g., in the form of a decision tree
Decision Tree for Epistemic Activities

Does the segment contain information from problem space?

Does the theoretical concept relate to problem space?

Subject mentioned

Does the segment contain a theoretical concept?

elaboration of problem space

Parents

Concept mentioned

Concept mentioned

elaboration of theory

off-topic

Teacher

Concept mentioned

Concept mentioned

Concepts mentioned (and part of the theory space):

A: locus
B: stability
C: controllability
D: attributing failure
E: attributing success
F: re-attribution training

Legend:
Top-down decision nodes: white rectangles
Bottom-up decision nodes: grey rectangles
End nodes (categories): triangles
Example for coding rules in form of a decision tree (Wecker, 2006)

1. Is there any talk in the segment at all (incl. mumbling)? yes: 2, no: 4

2. Is there any talk longer than 1 sec.? yes: 6, no: 3

3. Do the learners ask about the (i) reading progress (e.g., „Are you done?“), (ii) protest against scrolling down (e.g., „Stop!“), (iii) comment about any text (e.g., „Haha: ‘chacked’!“; „What means ‘focused’?“) or (iv) describe the current activity (e.g., „We are reading.“)?
   
   1. yes: Coding „Information intake“ for the current segment and all prior segments up to that segment that has been coded as „no activity (silence)“
   
   2. no: 4
Example for a framework for analysing verbal data in CSCL environments (Weinberger & Fischer, 2006)

- **Multiple dimensions:**
  - Participation dimension
  - Epistemic dimension
  - Formal-argumentative dimension
  - Dimension of social modi of co-construction (incl. transactivity)
# Multiple Dimensions of Argumentative Knowledge Construction

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participation</strong> <em>(Words and messages; Cohen, 1994)</em></td>
<td>Do learners participate (at all) in Online-Discourse?</td>
</tr>
<tr>
<td>• Quantity</td>
<td></td>
</tr>
<tr>
<td>• Homogenity</td>
<td></td>
</tr>
<tr>
<td><strong>Epistemic Activities</strong> <em>(κ = .90; Fischer, Bruhn, Gräsel, &amp; Mandl, 2002)</em></td>
<td>Do learners argue on task? Do learners construct arguments based on the relevant concepts?</td>
</tr>
<tr>
<td>• construction of problem space</td>
<td></td>
</tr>
<tr>
<td>• construction of conceptual space</td>
<td></td>
</tr>
<tr>
<td>• construction of relations between conceptual and problem space</td>
<td></td>
</tr>
<tr>
<td><strong>Argumentation</strong> <em>(κ = .78; Leitão, 2000)</em></td>
<td>Do learners construct formally complete arguments and argument sequences?</td>
</tr>
<tr>
<td>• construction of single arguments</td>
<td></td>
</tr>
<tr>
<td>• construction of argumentation sequences</td>
<td></td>
</tr>
<tr>
<td><strong>Social Modes of co-construction</strong> <em>(κ = .81; Teasley, 1997)</em></td>
<td>Do learners operate on the reasoning of their learning partners? How do learners build consensus?</td>
</tr>
<tr>
<td>• Externalization</td>
<td></td>
</tr>
<tr>
<td>• Elicitation</td>
<td></td>
</tr>
<tr>
<td>• Quick consensus-building</td>
<td></td>
</tr>
<tr>
<td>• Integration-oriented consensus-building</td>
<td></td>
</tr>
<tr>
<td>• Conflict-oriented consensus-building</td>
<td></td>
</tr>
</tbody>
</table>
Testing and documenting reliability

- Objectivity of coding -> interrater reliability
  - Two or more coders code the same segments
  - Similarity between codes is compared
    (-> Cohen‘s Kappa, Krippendorff‘s alpha, ICC)
- Objectivity requires training
Standard training process

- Explanation phase
  - Definition of dimensions and codes

- Modelling phase
  - Joint coding of example data

- Practice
  - Individual coding of example data
    - if objectivity sufficient -> training successful
    - if objectivity not sufficient -> modelling phase + feedback
Training material

- Rule of thumb: 10% of the raw data per testing/practice
- Randomly selected data
  - All experimental conditions have to be represented
  - All codes need to be coded at least several times to test objectivity
## Feedback: Crosstables

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th></th>
<th></th>
<th></th>
<th>Gesamt</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>88</td>
<td>99</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>13</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>88</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>99</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Gesamt</td>
<td>16</td>
<td>18</td>
<td>1</td>
<td>8</td>
<td>10</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maß der Übereinstimmung</th>
<th>Wert</th>
<th>Asymptotischer Standardfehler</th>
<th>Näherungswert $T^a$</th>
<th>Näherungswert Signifikanz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kappa</td>
<td>.456</td>
<td>.078</td>
<td>7.440</td>
<td>.000</td>
</tr>
<tr>
<td>Anzahl der gültigen Fälle</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Typical consequences of low objectivity

- Refinement of coding scheme, i.e. clarification of rules, additional examples
- Adaption of coding scheme
  - combination of codes
  - additional codes
- Beware of skewed data:
  - High objectivity due to code „other“
## Micro-Coding

<table>
<thead>
<tr>
<th>Lombard et al. - Criteria</th>
<th>1st wave of studies 00/01</th>
<th>2nd wave of studies 02/03</th>
<th>3rd wave of studies 03/04</th>
</tr>
</thead>
<tbody>
<tr>
<td>size of reliability sample</td>
<td>ca. 500 Seg.</td>
<td>199 Seg.</td>
<td>176 Seg.</td>
</tr>
<tr>
<td>relationship of the reliability sample to the full sample</td>
<td>105 participants 2821 segments</td>
<td>289 participants 6296 segments</td>
<td>243 participants 9825 segments</td>
</tr>
<tr>
<td>N of coders</td>
<td>2 students</td>
<td>6 students</td>
<td>5 students</td>
</tr>
<tr>
<td>amount of coding</td>
<td>50% each</td>
<td>ca. 17% each</td>
<td>ca. 17% each</td>
</tr>
</tbody>
</table>
| Reliability indices | Seg.: 87%  
Epi.: $\kappa = .90$  
Arg.: $\kappa = .78$  
Soz.: $\kappa = .81$ | Seg.: 83%  
Epi.: $\kappa = .72$  
Arg.: $\kappa = .61$  
Soz.: $\kappa = .70$ | Seg.: 85%  
Epi.: $\kappa = .89$  
Arg.: $\kappa = .91 \bar{\Omega}$  
Soz.: $\kappa = .87$ |
| Reliability of each variable | --- | --- | --- |
| amount of training | ca. 500 h in each wave trained with 1000 to 1500 discourse segments |
| references | Weinberger, Fischer, & Mandl, 2001; Weinberger & Fischer, 2006 |
Automatisation of coding

- Machine learning algorithms learn from already coded data

- Features of written text need to be extracted (e.g. word count, unigrams, bigrams, punctuation)
  - LightSIDE or TagHelper extract features and prepare them for the training of machine learning algorithms
Automatisation: Step 1

- Get the software „LightSIDE“ (it’s free):

  http://ankara.lti.cs.cmu.edu/side/download.html
Automatisation: Step 2

- Prepare your data
  - First column: text
  - Second column: code
- Save as csv-file
- Load file csv-file into LightSIDE
Automatisation: Step 3

- Extract features
Automatisation: Step 4

- Train model
Automatisation: Step 5

- Improving models
  - exclude rare features
  - exclude misleading features
  - add semantic rules
Automatisation: final step

- Apply model to new material
  - Must not be different from training material -> change of context, topic, sample may cause problems

- Automatically coded data require careful supervision
# Automatisation: Does it work?

**Table 2** Comparison without and with the layer of extracting attributes to automate the content analysis

<table>
<thead>
<tr>
<th>Segmentation layer II</th>
<th>Without extracting attributes</th>
<th>With extracting attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kappa SIDE-Training Material</strong></td>
<td>Cohen’s Kappa: 0.84</td>
<td>Percent Agreement: 96.7 %</td>
</tr>
<tr>
<td><strong>Kappa SIDE-Testing Material</strong></td>
<td>Cohen’s Kappa: 0.86</td>
<td>Percent Agreement: 97.0 %</td>
</tr>
<tr>
<td><strong>Major choice</strong></td>
<td>Cohen’s Kappa: 0.80</td>
<td>Percent Agreement: 96.7 %</td>
</tr>
<tr>
<td><strong>Math</strong></td>
<td>Cohen’s Kappa: 0.86</td>
<td>Percent Agreement: 96.6 %</td>
</tr>
<tr>
<td><strong>Class reunion</strong></td>
<td>Cohen’s Kappa: 0.87</td>
<td>Percent Agreement: 97.0 %</td>
</tr>
<tr>
<td><strong>Between-culture variance</strong></td>
<td>Cohen’s Kappa: 0.90</td>
<td>Percent Agreement: 97.7 %</td>
</tr>
<tr>
<td><strong>Text-analysis</strong></td>
<td>Cohen’s Kappa: 0.83</td>
<td>Percent Agreement: 96.9 %</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coding layer III</th>
<th>Without extracting attributes</th>
<th>With extracting attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kappa SIDE-Training Material</strong></td>
<td>Cohen’s Kappa: 0.70</td>
<td>Percent Agreement: 75.6 %</td>
</tr>
<tr>
<td><strong>Kappa SIDE-Testing Material</strong></td>
<td>Cohen’s Kappa: 0.61</td>
<td>Percent Agreement: 67.8 %</td>
</tr>
<tr>
<td><strong>Major choice</strong></td>
<td>Cohen’s Kappa: 0.63</td>
<td>Percent Agreement: 71.2 %</td>
</tr>
<tr>
<td><strong>Math</strong></td>
<td>Cohen’s Kappa: 0.67</td>
<td>Percent Agreement: 72.3 %</td>
</tr>
<tr>
<td><strong>Class reunion</strong></td>
<td>Cohen’s Kappa: 0.47</td>
<td>Percent Agreement: 58.5 %</td>
</tr>
<tr>
<td><strong>Between-culture variance</strong></td>
<td>Cohen’s Kappa: 0.53</td>
<td>Percent Agreement: 63.1 %</td>
</tr>
<tr>
<td><strong>Text-analysis</strong></td>
<td>Cohen’s Kappa: 0.68</td>
<td>Percent Agreement: 75.0 %</td>
</tr>
</tbody>
</table>
Checklist for argumentation analyses

• Theoretical framework

• Research questions and methods that can address those questions in a valid manner

• Explicit and theory-based set of rules for segmentation and categorization

• Testing and documenting reliability (see Lombard et al., 2002)

• Replication
Testing and documenting reliability:  
Part 1  
(Lombard, Snyder-Duch, & Braaken, 2002)

- the size of and the method used to create the reliability sample, along with a justification of that method;
- the relationship of the reliability sample to the full sample;
- the number of reliability coders and whether or not they include the researchers;
- the amount of coding conducted by each reliability and non-reliability coder;
Testing and documenting reliability: Part 2
(Lombard, Snyder-Duch, & Braaken, 2002)

- the index or indices selected to calculate reliability and a justification of these selections;

- the inter-coder reliability level for each variable, for each index selected;

- the approximate amount of training (in hours) required to reach the reliability levels reported;

- where and how the reader can obtain detailed information regarding the coding instrument,

- procedures and instructions (for example, from the authors).
Conclusions

- CSCL is an ideal context to investigate collaborative and individual knowledge construction processes, which can be reliably assessed with a multi-granular and multi-dimensional framework (Weinberger & Fischer, 2006).

but

- which requires major training efforts
- which captures most, but maybe not all cognitive processes of knowledge construction
Example 1

- Analyses of cognitive processes of learning through think-aloud protocols in CSCL
Analysis of cognitive processes

- Think-aloud protocols
- 10-Sec segments
- coding (κ = .78):
  - Elaboration depth
  - Elaboration focus
    - Elaboration of content
    - Elaboration of peer contributions
Good learner, no script
Learner with script, role of analytic
Learner with script, role of critic

Transactivity of discourse
- High
- Medium
- Low

Elaboration of contributions of the learning partner

Epistemic quality of discourse
- High
- Medium
- Low

Elaboration of the learning material

Seconds

0 200 400 600 800 1000
Example 2

- CSCL-assumption learners are influencing each other
- Requirement for analysis is independence of observations
- Analyzing individuals, groups, or both with **multi-level modeling**
Example 3

Collaboration (Knowledge sharing in collaboration)

- Use public transportation
- Use solar energy
- Use biodegradable bottles
- Save water
- Use wind energy
- Turn TV off
- Recycle more
- Plant trees
- Use solar energy
- Use wind energy

Shared prior knowledge

Learning from fellow learner

Shared knowledge

Text

Pre-test

Student A: Text
- Use public transportation
- Save water

Post-test

Student A: Text
- Use biodegradable bottles
- Recycle more
- Use solar energy

Pre-test

Student B: Text
- Use solar energy

Post-test

Student B: Text
- Use biodegradable bottles
- Save water
- Use wind energy

Learning from fellow learner
Example 4

- Inadequate inference (-1) (based on irrelevant prior knowledge)
- Writing aloud (0) *
  Adequate inference between problem and theoretical concept (0)

Learning partner (-2)*
- Adequate inference between problem and theoretical concept (-2)
- Grounded claim (-2)
- Grounded claim with qualification (-2)

Learning partner (-2)*
- Conflict-oriented consensus building (-2)
- Integration-oriented consensus building (-2)
- Counter argumentation (-2)

Task description (-2)
- Writing/Thinking aloud (-1) *
  Grounded claim (-1)
  Grounded claim with qualification (-1)

Problem information (-2)
- Theory paper (-2)

Positive relation
Negative relation
Covariance
Literature


