Annotation Guidelines

Argumentation structures in scientific articles from the educational domain
August 2015

1 Introduction

The goal of any scientific publication is to present and to explain one's own research project to an expert audience. Accordingly, in a scientific publication as many statements as possible should be established and proved in terms of an argumentation. The goal of this annotation study is to identify the argumentation structures precisely described in the following paragraphs in order to later simplify reading comprehension and access to information.

In this annotation study we are concerned exclusively with empirical publications made by education researchers with a length of about ten pages and containing the following section: Abstract, introduction, theoretical background, methods, results, discussions. Each of these sections has a specific function within the publication. Further structural characteristics include subsections, paragraphs, sentences and phrases.

During the annotation we don't consider the Methods and Results sections because the arguments mentioned there are very special and technical and could not easily be considered out of context. In this study we are interested more in information which is also comprehensible for laymen (e.g.: "Mädchen sind besser in der Schule, weil Schule einen höheren Stellenwert in ihrem Leben einnimmt und sie dadurch mehr Zeit investieren"("Girls do better at school because school has a higher priority in their lives and they invest more time in it.")) than in information comprehensible exclusively in the context of the current study (e.g.: "In der 9. Klasse zeigten sich keine signifikanten Unterschiede: Mädchen berichteten eine Durchschnittsnote von M = 3,14, Jungen von M = 3,21"("In the 9th there were no significant differences noticed: the girls' average grade was M = 3,14; boys' – M = 3,21")).

1.1 Definition Argument

Arguments are justified (or proved) statements. They consist of argumentation units which are connected to each other by means of directed relations. With the help of these relations the argumentation units can support (support) or attack (attack). We have a support relation in the case where statement A1 is supported or proved by some additional statement A2 (A2 supports A1). We have an attack relation when some statement A1 is attacked, refuted, restricted, criticized or questioned by some other statement A2 (A2 attacks A1). An argumentation unit always corresponds to exactly one sentence, whereas the end of a sentence is indicated by the following punctuation: ., !, ?, ;, :.

Besides the relations support and attack we use the relations detail and sequence (see Paragraph 1.3 or 1.4) to mark further, non-argumentative relations which are nevertheless important for text comprehension.

Example 1 shows an argument which consists of three argumentation units. At the upper level (green) the author justifies the girls' educational success by the fact that they give a higher importance to school than boys. This thesis is supported by another argumentation unit (blue) (Study shows that girls make more effort (Studie zeigt, dass sich Mädchen mehr anstrengen)). To provide evidence for this claim, further sources (red) were quoted.

Example 1:

The current higher educational success of girls in comparison to boys is explained in the literature by (among other things) the fact that girls show much stronger patterns of behavior and attitude that prove the higher importance of school in their lives. Thus, different studies show that girls, according to their own statements, spend more time per week doing their homework than boys. (e.g. Wagner, Schober & Spiel, 2008) and thereby on average make a greater effort (Trautwein, Lüdtke, Kastens & Köller, 2006) (...).

Der derzeit im Vergleich zu Jungen relativ größere schulische Erfolg von Mädchen wird in der Literatur unter anderem dadurch erklärt, dass Mädchen in stärkerem Maße als Jungen Verhaltensweisen und Einstellungen zeigen, die für einen hohen Stellenwert der Schule im eigenen Leben sprechen. So zeigten verschiedene Studien, dass Mädchen nach eigenen Angaben mehr Zeit pro Woche mit der Bearbeitung von Hausaufgaben verbringen als Jungen (z. B. Wagner, Schober & Spiel, 2008) und sich dabei im Durchschnitt mehr anstrengen (Trautwein, Lüdtke, Kastens & Köller, 2006) (...).
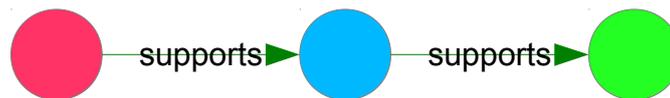
Example 2 shows an attack relation: the statement "Der Fragebogen erweist sich als geeignet"(The questionnaire proves to be suitable) is attacked by some further statement.

Example 2:

The questionnaire proves to be suitable for the classification of children with little or no German knowledge and of children with advanced German knowledge. However, it is also important to remember that children with little German knowledge, in comparison with children with a speech disorder, do not belong to one group with some discrete characteristic features.

Der Fragebogen erweist sich hinsichtlich der Klassifikation von Kindern mit keinen oder geringen Deutschkenntnissen und Kindern mit fortgeschrittenen Deutschkenntnissen als sehr geeignet. Es gilt jedoch zu bedenken, dass es sich bei Kindern mit geringen Deutschkenntnissen im Vergleich zu Kindern mit einer Sprachentwicklungsstörung nicht um eine Gruppe mit einer diskreten Merkmalsausprägung handelt.

If a statement is supported or attacked by some other element (say, by some reference), then an argument is present. It should be noted, however, that supports and attacks can be also implicit. A typical example of this is the summary at the end of the following discussion paragraph:

Example 3:

The DaZ-E questionnaire can be considered a time- and cost-effective method to assess the language of education of children with German as a second language at the age of 33 to 48 months.

Der Fragebogen DaZ-E kann als zeitökonomisches und kostengünstiges Verfahren zur Erfassung der

This statement is not justified here. But it is clear for the reader that it is based on the results of the recent study which provides support for this statement. The sentence must therefore be annotated as an argumentation unit (without relations, if necessary). It is even clearer if the sentence starts with words such as, "Die Befunde der Studie deuten darauf hin, dass ..."(The results of the study indicate that...). It is clear, without providing a support relation for each finding of the study, that the statement made here is supported by its own results.

When identifying arguments and argumentative relations it is helpful to reformulate the relevant units and to organize them into some structure as in the following examples (where "A" –supports–> "B" is applicable):

- B is valid because of / due to / as a result of A (Mädchen sind besser in der Schule, weil sie sich mehr anstrengen / Girls do better at school because they work harder)

- A, therefore / thus / hence B (Mädchen strengen sich mehr an. Deshalb sind sie besser in der Schule / Girls work harder. Therefore they do better at school)

- B. The study A demonstrates this. (Mädchen sind besser in der Schule. In einer Studie waren die Noten im Durchschnitt um 0.4 Einheiten besser / Girls do better at school. In a study the grades were on average 0.4 units better)

Furthermore, it can be helpful to pay attention to certain words which connect different sentences (discourse connectors). The following examples should give an impression. However, the occurrence of these words doesn't always mean that the corresponding relation exists!

- Support: ":", "so (so)", "because (weil)", "due to (wegen)", "as (da)", "therefore (deshalb)", "hence (daher)", "from this follows (daraus folgt)", "thus (somit)", "for this reason (damit)", "as a result (in folge dessen)", "consequently (demzufolge)", "accordingly (folglich)", "ergo (also)", ...

- Attack: "however (hingegen)", "in contrast (dagegen)", "though (jedoch)", "nevertheless (trotzdem)", "yet (dennoch, allerdings)", "but (aber)", "on the other hand (andererseits)", "not...but (zwar... aber)", ...

1.2 Non-argumentative text

Before you proceed with identifying and connecting argumentative units it is necessary to decide whether any given text paragraph is argumentative in the first place. Non-argumentative are those paragraphs which report on the current state of research or give you other background information or definitions which support comprehensibility and structure without requiring further reasoning. For each individual paragraph we decide whether it is argumentative or not, and then mark the entire paragraph as argumentative or non-argumentative. When it is unclear whether the paragraph is argumentative or not, it can be helpful to question whether the paragraph contains any information important for a summary of the paper. Also, we've observed that a large proportion of the paragraphs in a document can be viewed as argumentative, so that in a case of doubt a paragraph should be considered argumentative.

Here are some examples of non-argumentative and therefore unannotatable texts:

Example 4:
In the following paragraph, we first of all define the most important terms. Afterwards, we describe the

study we carried out. (...)

Im Folgenden Abschnitt werden zunächst die wichtigsten Begriffe definiert. Anschließend beschreiben wir die durchgeführte Studie (...).

In this example, only the structure of the publication is described.

Example 5:

In this study we examine N = 180 multilingual children and N = 180 monolingual German-speaking children at the age of 5.

In dieser Studie wurden N = 180 mehrsprachige Kinder und N = 180 ausschließlich deutschsprachige Kinder im Alter von 5 Jahren untersucht.

Description of the study.

Example 6:

By social competence we mean the totality of a person's knowledge, skills and abilities which leads to socially competent behavior.

Unter sozialer Kompetenz verstehen wir die Gesamtheit von Wissen, Fähigkeiten und Fertigkeiten einer Person, die zu sozial kompetentem Verhalten führt.

Definition of the term "social competence".

1.3 The "detail" Relation

A pattern that appears very often in publications is as follows: At first one makes a statement in a sentence which is explained more precisely (in "detail") in the following sentence(s). Very often it involves the reasoning / proof in the sense of a support relation (see Example 1).

However, this pattern is very often used to give some background information: in this case the information to be conveyed is summarized in the introductory sentence and subsequently explained in detail. In principle, we could also call this a support, but without any argumentation. We annotate this relation as a detail relation (see Example 7). This relation can be very useful for rapid comprehension of a text or for an automatic summarization, because one can focus more on summarized sentences, while skipping the details.

Example 7:

Würth (2001) presents a stringent theory-based approach to athletic career development in her analysis of the young athletes' carrier paths for the German sport system. Her explanations rest on the Salmela's (1994) phase model which describes the athletic career in three phases: initiation, development and mastery.

Einen stringent theoriegeleiteten Ansatz zur sportlichen Karriereentwicklung stellt Würth (2001) in ihrer Untersuchung zu Karriereverläufen jugendlicher Athleten für das bundesdeutsche Sportsystem vor. Ihre Ausführungen lehnen sich an das Phasenmodell von Salmela (1994) an, der die sportliche Karriere in drei Phasen beschreibt: die Phase des Beginns, der Entwicklung und der Meisterschaft.
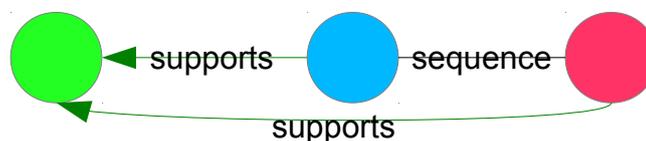

1.4 The "sequence" Relation

As we've already mentioned in the introduction, we observe each argumentative sentence as a separate argumentation unit. However, it is possible for several consecutive sentences to be closely linked in content or to make sense only as a unit. In this case these consecutive sentences are connected by the undirected sequence relation. In this case, support / attack / detail relations marked for individual sentences in the sequence normally apply to the whole sequence. However, it is also possible that other argumentation units correspond only to individual sentences of a sequence.

In Example 8, Sentences 2 and 3 build a sequence because only together do they demonstrate that the results are inconsistent and therefore only together do they support the first sentence.

Example 8:

The investigations provide inconsistent results. The students taught in immersion programs evaluate their education as exciting. Other investigations found no differences.

Die Untersuchungen liefern uneinheitliche Befunde. So beurteilen immersiv unterrichtete Schüler ihren Unterricht als spannender. Andere Untersuchungen finden keine Unterschiede.

The enumerations used in Example 9 are also typical for sequences. It should be noted here that the attack relation corresponds only to the second element of the sequence (In contrast, relations usually correspond to the whole sequence).

Example 9:

The parent questionnaires show similarly high or even higher correspondences between the parents estimation and the objective language measures than those of comparable and established parent questionnaires for monolingual children in German-speaking or Anglo-American countries. Furthermore, each of the ten language versions with correlations in the "strong effect" range is proved to be valid. However, when comparing the language versions one can also see differences.

Der Elternfragebogen weist ähnlich hohe oder gar höhere Übereinstimmungen zwischen der Elterneinschätzung und den objektiven Sprachmaßen auf, als vergleichbare und etablierte Elternfragebogen für monolinguale Kinder im deutschsprachigen und angloamerikanischen Raum. Zudem erweist sich jede der zehn Sprachversionen mit Korrelationen im hohen Effektstärkebereich als valide. Allerdings zeigen sich beim Vergleich der Sprachversionen auch Unterschiede.
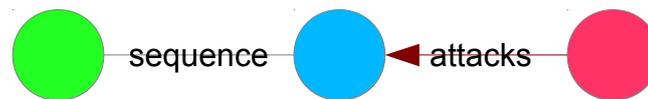
2 Annotation process

This paragraph describes the procedure that is to be followed when annotating a document. First, you should determine what the article is about by reading the abstract at the beginning of the document. Afterward read the document section by section (i.e. the text between each pair of headings) and divide the text into different concepts (see below). Afterward, within each concept, link the argumentation units via relations. After each fully annotated paragraph, the annotations must be read through and verified once more. The individual steps will be illustrated hereafter in detail.

Step 1: Get an overview

Read the abstract of the document and determine what the article is about: What is the problem? What has been done? What are the goals? What are the results?

Step 2: Identification of concepts

- Start to read the document section by section (i.e. the text between pairs of headings). Decide whether each paragraph is argumentative or not. Argumentative paragraphs must be subdivided into "Concepts". A concept is a text segment which deals with a particular aspect of the content and is distinct from every other concept.

- Argumentation units of different concepts can't be related to each other. A concept can contain one or more arguments (Reminder: An argument consists of several argumentation units combined via relations).

- Individual paragraphs often correspond to a single concept. In some instances, however, for example in Paragraphs 1 and 2 are discussed different aspects which in Paragraph 3 provide support for a conclusion. These three paragraphs should afterward be summarized in one concept so that the support relation can be annotated. It is also possible to draw conceptual borders within a paragraph. The sentences must not be divided.

- It can be helpful to pay attention to discourse connectors, words which connect two sentences (e.g. this, therefore, however, yet, because, furthermore). Such connectors are an indication that

the given sentence still belongs to the same concept. A concept, however, never goes beyond the paragraph's heading.

- It can be helpful to summarize in a few words the content of each selected concept and therefore separate it from other concepts (e.g. Concept 1: Discussion of related works, Concept 2: Discussion of the author's own approach).

- • For the sake of clarity it is generally recommended to delineate concepts such that they are as small as possible, but still big enough that all their relations can be represented (keeping in mind that relations can be identified only within a concept).

Step 3: Annotation and connection of argumentation units

The argumentation units are automatically marked within the concepts with the help of annotation tools. References (see Paragraph 3.2) as well as relations between them (supports / attacks / details / sequence) must then be manually marked.

Step 4: Verification of annotations

After you have fully annotated a paragraph you should once more verify the plausibility of all annotations. For this purpose you should go through the relation graph from the bottom up and verify whether the relation makes sense. It can be helpful to re-formulate each particular argument (which consists of two argumentation units connected via a relation). Any possible ambiguity should be made note of.

3 Further tips

3.1 Headings

Headings should not be annotated.

3.2 References

References are often used as support. The annotation tool therefore provides a separate argumentation unit called citation. All the units annotated with it automatically get a (unique) argument ID. However, it is essential to distinguish between references which are used to substantiate the statement and references which explain a term within a statement. In first case there is a support relation; in the second case, a details relation. If it is impossible or very difficult to decide which of the two cases is present (see Example 10), we always use the support relation.

Example 10:

In light of this, children with ADHD should receive a multimodal therapy (cf. Petermann & Hampel, 2009), in which the establishment of social skills should specifically follow from the concept of cognitive information processing.

Auf diesem Hintergrund sollten Kinder mit ADHS eine multimodale Therapie erhalten (vgl. Petermann & Hampel, 2009), wobei der Aufbau sozialer Kompetenzen besonders dem Konzept der sozial-kognitiven Informationsverarbeitung folgen sollte.

It is unclear in this example whether the reference is used to justify the multimodal therapy (in which case the reference could be a support) or whether the term "multimodale Therapie"(multimodal therapy) is explained in this cited source (detail). In most cases references are used as support and in case of uncertainty should be annotated as such.

3.3 Directionality of the relations when formulating a goal

When formulating a goal there are two ways of expressing it in the text flow:
1. The author describes the goal as well as what has been done to reach this goal
2. The author describes his approach and afterward explains which goal is pursued with it.

Text type 1 can be especially confusing:

Example 11:
A confounder should be ruled out.   For this purpose students had been randomized.

Eine Konfundierung sollte ausgeschlossen werden.   Dazu wurden Schüler randomisiert.

Example 12:
Students were randomized.   A confounder should thus be ruled out.

Schüler wurden randomisiert.   Eine Konfundierung sollte damit ausgeschlossen werden.

Basically both examples are identical in content. However, in Example 11 the students were randomized because a confounder must be ruled out. In Example 12, on the other hand, it follows from the randomization of the students that the goal to rule out a confounder can be reached. Basically, when annotating, one should always try to have the author's intention in mind. In most cases the relation in this example should look as follows: "Schüler werden randomisiert"(Students are randomized) –supports–> "Ziel Konfundierung auszuschließen kann erreicht werden"(Goal ruling out confounder can be reached.). At the word "sollte"(should) it becomes clear that in Example 11 a goal is defined so that the example corresponds to the text type written above. In such cases caution should be exercised because discourse markers (here: dazu / for this purpose) can be misleading, with regard to the actual direction of the relation.

Below is another example which must be annotated according to the graph. The argument in this example simply reads, "We investigate comprehensive schools with mono- and coeducational physics lessons, with the aid of an experimental design which rules out a confounder. Therefore we can bolster the validity (which is our goal)."

Example 13:
The goal of the current study is to bolster the validity of the findings already indicated in the Kiel-based pilot experiment concerning the effect of mono-educational physics elementary education.   For this, we scientifically conducted a Berlin-based pilot project, where the physics lessons in 8th Grade were carried out in either co- or mono-educational groups. Comprehensive schools were investigated in order to bolster the external validity of the effects observed in the Kiel-based pilot project for high school students. By choosing an experimental design that allows to rule out confounders we expect results with higher internal validity.

Ziel der vorliegenden Studie ist es, die Validität der bereits im Kieler Modellversuch erzielten Befunde zu den Auswirkungen monoedukativen Physik-Anfangsunterrichts zu verbreitern. Dazu haben wir einen Berliner Schulversuch wissenschaftlich begleitet, in dem der Physikunterricht in der achten Klasse entweder in koedukativen oder monoedukativen Gruppen durchgeführt wurde. Um die externe Validität der im Kieler Modellversuch für Gymnasiasten gefundenen Effekte zu verbreitern, wurden Gesamtschulen untersucht. Durch die Wahl eines Untersuchungsdesigns, das Konfundierungen auszuschließen erlaubt, sind Ergebnisse mit erhöhter interner Validität zu erwarten.

3.4 Explanation of results

It is particularly common in discussion sections for the results of the investigation to be (more broadly) repeated, and to present possible explanations for them. These explanations should be regarded as support for the results (see Example 12).

Example 14:

The gender makeup of the study groups had no impact on the score for physics in the 8th Grade certificate. One of the possible methodical explanations is that girls from the mono-educational groups which were enrolled in the elementary courses showed worse results in the 8th Grade certificate than the girls who came to this course from the coeducational groups.

Die Geschlechterkonstellation der Lerngruppen hatte keinen Einfluss auf die in der Physik im Abschlusszeugnis der achten Klasse erreichte Punktzahl. Eine mögliche methodische Erklärung ist, dass die den Grundkursen zugewiesenen Mädchen aus den monoedukativ zusammengesetzten Gruppen vergleichsweise schlechter im Schuljahresendzeugnis der achten Klasse abgeschnitten hatten als die Mädchen, die aus koedukativ unterrichteten Gruppen in diese Kurse kamen.

However, when the results are interpreted or when conclusions are drawn from them, then the results themselves serve as support (see Example 15).

Example 15:

The results show that effect of the youths' gender or of the gender makeup of the study groups didn't change for any of the variables between the middle and the end of the academic year. Because the female and male students had been randomly assigned to the study groups, the findings can be interpreted to mean that the mono-educational teaching had already been effective after half an academic year (...).

Die Ergebnisse zeigen, dass Effekte des Geschlechts der Jugendlichen oder der Geschlechtskonstellation der Lerngruppen sich für keine der abhängigen Variablen zwischen Schulhalbjahresmitte und Schuljahresende veränderten. Weil die Schülerinnen und Schüler den Lerngruppen randomisiert zugewiesen worden waren, können diese Befunde so interpretiert werden, dass der monoedukative Unterricht bereits nach einem halben Schuljahr Wirkung zeigte (...).


3.5 Further tips for the attack relation

Sometimes it is difficult to decide whether there is an attack relation or some other one. In Example 16 there is no attack relation despite the cue word "however"(allerdings) in the third sentence, because here it involves an enumeration. However, an attack relation usually appears when there is a cue word (but, yet, however,...). There is a typical procedure in scientific publications where at first a research gap is reported and in the next sentence it is restricted to some extent by the indication of the preexisting knowledge or the research gap (see Example 17). Such a procedure must be annotated with an attack relation due to its restrictive character.

Example 16:

As expected, children from childhood homes with two non-German-speaking parents have the greatest disadvantages when acquiring the German language. Furthermore the disadvantages are reflected in other competence areas (see above). However, in comparison with children from pure German families, a significantly weaker vocabulary is observed for children with only one natively German-speaking parent.

Erwartungsgemäß haben Kinder aus Elternhäusern mit zwei nicht deutschsprachigen Elternteilen die

größten Nachteile beim Erwerb der deutschen Sprache; Nachteile zeigen sich darüber hinaus auch in anderen Kompetenzbereichen (siehe oben). Kinder mit einem deutschsprachigen und einem nicht muttersprachlich deutschen Elternteil verfügen allerdings ebenfalls über einen signifikant schwächeren Wortschatz als die Kinder aus rein deutschsprachigen Familien.

sequence — sequence
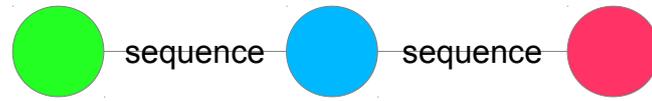
Example 17:

The relevant competences and associated education decisions result in turn from cumulative development and funding processes as well as from decision-making processes, whose correlations are still comparatively little known today. Important milestones of children's development of skills can be clearly located before school acceptance (educational context).

Die relevanten Kompetenzen und die damit verbundenen Bildungsentscheidungen resultieren ihrerseits aus kumulativen Entwicklungs- und Förderprozessen sowie Entscheidungsverläufen, über deren Wechselbeziehungen heute noch vergleichsweise wenig bekannt ist. Wichtige Meilensteine kindlicher Kompetenzentwicklung sind jedoch schon deutlich vor dem Eintritt in den Bildungskontext Schule zu verorten.

3.6 Further tips on relations

Experience shows that in some cases it is difficult to decide which sentence refers to an argumentative unit. In Example 18 the third sentence attacks the previous sentence and therefore restricts the statement that the study by Weber, Marx and Schneider supports its own results to some extent. And yet one can argue that it is not the second sentence that the third sentence attacks, but rather the first one, and it therefore challenges the authors' own study because here, in contrast to the study by Weber, Marx and Schneider, low effects occur. In case of doubt it is preferable to annotate short relations, i.e. relations between closely related sentences. Apart from that it is often reasonable to build argumentation chains (e.g. A –> B –> C).

Example 18:

By the time they enter kindergarten, a poorer vocabulary and worse knowledge of German grammar is exhibited by children with a migration background (with regard to German as a majority language, and above all for children from families with two non-German parents). This finding is in line with similar findings made by Weber, Marx and Schneider (2007). They compared native and non-native German-speaking preschoolers and also found strong effects on the vocabulary. However, the effects turned out to be even higher in the cited study than in the present one. This can be traced back to the cumulative effects on slightly older children.

Zum Zeitpunkt des Kindergarteneintritts verfügen die Kinder mit Migrationshintergrund (vor allem solche aus Familien mit zwei nicht deutschen Elternteilen) bezogen auf die deutsche Mehrheitssprache über einen geringeren Wortschatz und beherrschen die deutsche Grammatik vergleichsweise schlecht. Dieser Befund steht im Einklang mit Ergebnissen von Weber, Marx und Schneider (2007), die Vorschulkinder deutscher und anderer Muttersprache verglichen und ebenfalls starke Effekte im Wortschatz fanden. Allerdings fielen die Effekte in der genannten Studie noch höher aus als in der vorliegenden Arbeit, was möglicherweise auf kumulative Effekte bei den etwas älteren Kindern zurückgehen könnte.