

# Pragmatics and Discourse SS08

Ivana Kruijff-Korbayova

## cohesion, coherence, RST

Magdalena Wolska

# What is a unit of communication?

Theories of discourse meaning depend in part on a specification of the basic units of a discourse and the relations that can hold among them. Discourse processing requires an ability to determine to which portions of a discourse an individual utterance relates. Thus the role of discourse structure in discourse processing derives both from its role in delimiting units of discourse meaning and...

M: Queen beats the laufer at e1. Check.

...

## What is a discourse?

“Consider, for **you** have collected an arbitrary set of well-formed (and independent) utterances, for instance, arbitrary by randomly selecting one **sentence** from each of **the pages**, or **chapters**, of **this book**. Do you have a discourse? A list of arbitrary words? The answer is that **these utterances**, which may or may not have **widely exhibited coherence**. Consider, for example, the difference between **passages (1871) and (1872).**”

(Jurafsky and Martin:695)

VS....

# Sentences cannot be processed in isolation

Textual links

Segmentation and Ordering

Please use the toilet, not the pool.

ple

The pool for members only.

Discourse is *internally linked*; it „hangs together”

Discourse exhibits patterns of lexical connectivity

→ ***cohesion***

linguistic *text-forming* devices:

**lexical repetition, synonymy/antonymy, ellipsis/pro-forms,  
enumeration, parallelism, co-reference (anaphora)**

- Time flies.
- You can't; ~~they fly~~ **they fly** too quickly.  
(Halliday and Hasan 1982)

## Anaphora: pronominal

My neighbor has a monster Harley 1200. *They* are huge but gas-efficient bikes.

One should mind *their* own business.

We had two hurricanes hit us in the teeth, one of which you can read about in the current New Yorker if you find the place. It's called "Our Windswept Correspondents" and I am *they*.

## Anaphora: nominal (definite NP)

Al bought a car the other day. [...] He took it out of the garage last night with the help of George Cottrell, and *the thing* gave forth such immense clouds of smoke that one man came running up and asked me where the fire was.

[...] I wanted a Trumpeter Swan who could play like Louis Armstrong, and I simply created him and named him Louis. The cutting of the webs between his toes is also fantastical, just as *the bird* itself is; [...].

## Anaphora: surface-count and demonstrative

Sarah could leave but she was also given an option to stay; she chose *the latter*.

Have just driven to town, carrying our cook<sub>1</sub> and our cook's dog<sub>2</sub>. Gave *the one*<sub>1</sub> \$300 in currency and placed *the other*<sub>2</sub> in the infirmary, with eczema.

## Strained anaphora

John became a guitarist because he thought that *it* was a beautiful instrument.

# Temporal anaphora

If I must declare *today* that I am not a Communist, *tomorrow* I shall have to testify that I am not a Unitarian. And *the day after*, that I never belonged to a dahlia club.

# Spatial anaphora

The awful hot spell broke last night and today is clear and beautiful, [...]  
*Across the street*, the entire janitorial family has blossomed out in pink carnations,  
[...]

# Abstract entity anaphora

Each Fall, penguins migrate to Fiji.

*That's* where they wait out the winter.

*That's* when it's cold even for them.

*That's* why I'm going there next month.

*It* happens just before the eggs hatch.

(Webber 1988)

Send engine to Elmira.

*That's* six hours.

(Byron 2002)

# Ellipsis

The well water had chemicals in it and nothing in the house worked as it should [*work*].

[I] Have been uncommunicative lately, and [*I have been*] lagging in life's race.

I'm afraid my poem isn't as nicely written as "Paradise Lost," but anyway, it's shorter [*than "Paradise Lost"*].

Ultimately, even after Garcia was gone, Ruelas was able to cope and move on with his career. And indeed, he has [*coped and moved on with his career*].

## Discourse has *structure*

Discourse segments can be connected in a limited number of ways

→ ***coherence***

Discourse is linguistic choices that make organizing the explicit

**identity** (sameness): *that is, that is to say, in other words, ...*

**opposition** (contrast): *but, yet, however, nevertheless, whereas, in contrast...*

**addition** (continuation): *and, too, also, furthermore, moreover, in addition,...*

**cause and effect**: *therefore, so, consequently, thus, it follows that, ...*

**concession** (willingness to consider the other side): *admittedly, true, I grant,...*

**exemplification** (shift from general/abstract to specific/concrete idea): *for example, for instance, after all, an illustration of, indeed, in fact, specifically,...*

## Coherence vs. Cohesion

coherence: structural, functional relations between sentences

cohesion: non-structural, text-forming relations that “tie”  
parts of discourse together

John went to his favourite music  
store to buy a piano.

John went to his favourite music  
store to buy a piano.

He was excited that he could  
finally buy a piano.

He was excited that he could  
finally buy a piano.

## Coherence vs. Cohesion

coherence: structural, functional relations between sentences

cohesion: non-structural, text-forming relations that “tie” parts of discourse together

**John** went to his favourite music store to buy a piano.

He was excited that he could finally buy a piano.

**John** went to his favourite music store to buy a piano.

He was excited that he could finally buy a piano.

## Coherence vs. Cohesion

coherence: structural, functional relations between sentences

cohesion: non-structural, text-forming relations that “tie” parts of discourse together

**John** went to his favourite music store to buy a piano.

He had frequented the store for many years.

He was excited that he could finally buy a piano.

**John** went to his favourite music store to buy a piano.

It was a store John had frequented for many years.

He was excited that he could finally buy a piano.

## Coherence vs. Cohesion

coherence: structural, functional relations between sentences

cohesion: non-structural, text-forming relations that “tie” parts of discourse together

**John** went to his favourite music store to buy a piano.

**He** had frequented the store for many years.

**He** was excited that he could finally buy a piano.

**John** went to his favourite music store to buy a piano.

**It** was a store John had frequented for many years.

**He** was excited that he could finally buy a piano.

## Coherence vs. Cohesion

coherence: structural, functional relations between sentences

cohesion: non-structural, text-forming relations that “tie” parts of discourse together

**John** went to his favourite music store to buy a piano.

**He** had frequented the store for many years.

**He** was excited that he could finally buy a piano.

He arrived just as the store was closing for the day.

**John** went to his favourite music store to buy a piano.

**It** was a store John had frequented for many years.

**He** was excited that he could finally buy a piano.

It was closing just as John arrived.

## Coherence vs. Cohesion

coherence: structural, functional relations between sentences

cohesion: non-structural, text-forming relations that “tie” parts of discourse together

**John** went to his favourite music store to buy a piano.

**He** had frequented the store for many years.

**He** was excited that he could finally buy a piano.

**He** arrived just as the store was closing for the day.

**John** went to his favourite music store to buy a piano.

**It** was a store John had frequented for many years.

**He** was excited that he could finally buy a piano.

**It** was closing just as John arrived.

## Coherence vs. Cohesion

coherence: structural, functional relations between sentences

cohesion: non-structural, text-forming relations that “tie” parts of discourse together

When Forks by Kennedy paid a courtesy call on Ronald Reagan the night before he was elected, he suggested four equal Western surveillance targets to be set up in the Soviet Union. It was a significant achievement, and it was a piece of work. It grows through a process of evolution or nocturnalism, it is sleepy, you rediscover it and I suppose forks could have windows through evaporation.

# Intentional Approach

utterances : actions

speaker's plan wrt. communicating intentions ties the discourse together

discourse understanding : recognizing speaker's intentions

# Intentional Approach (Grosz and Sidner 86)

three dimensions of discourse

linguistic structure : the utterances

intentional structure : hierarchy of intentions

attentional structure : model of objects, properties and relations  
that are *salient* at each point in discourse  
(dynamically changing)

# Intentional Approach (Grosz and Sidner 86)

Linguistic structure

discourse segments + relations that hold between them

# Intentional Approach (Grosz and Sidner 86)

## Linguistic structure

discourse segments + relations that hold between them

(para-) linguistic expressions reflect discourse structure

cue phrases, aspect, tense, intonation, gesture

discourse structure constraints discourse interpretation

anaphora resolution

# Intentional Approach (Grosz and Sidner 86)

## Attentional structure

participants' focus of attention (what is „attend to”)

modeled by *focus spaces*: objects and relations in focus

changes: insertion and deletion rules

# Intentional Approach (Grosz and Sidner 86)

## Intentional structure

### Discourse Purpose (DP)

purpose/intention held by discourse initiator

e.g. make hearer:

intend to perform a task,

believe a fact,

believe that one fact supports another fact,

identify an object,

identify a property of an object

assumption: one per discourse

# Intentional Approach (Grosz and Sidner 86)

Intentional structure

Discourse Segment Purpose (DSP)

how given segment contributes to DP

# Intentional Approach (Grosz and Sidner 86)

## Intentional structure

### Hierarchy of intentions

#### dominance

DSP1 *dominates* DSP2 if satisfying DSP2 is intended to provide part of satisfaction of DSP1

#### precedence

DSP1 *precedes* DSP2 if DSP1 must be satisfied before DSP2

(more next time)

## Rhetorical structure theory (Mann and Thompson 87)

identifies (binary) relations between discourse units

relations may be made explicit by linguistic cues

nucleus(N) / satellite(S) segments

model: domain-independent rhetorical structure  
compositionally built discourse tree

# Rhetorical structure theory (Mann and Thompson 87)

example definition

**Evidence:** S provides evidence for what N claims

constraints on N:	Hearer might not believe N as much as Speaker think s/he should
constraints on S:	H believes S or will find it credible
constraints on N&S:	H's comprehending S increases H's belief of N
intention of S:	R's belief of N is increased
locus of effect:	N

[ This tax calculation software really works. ]N

[ I entered all the figures from my tax return and got a result which agreed with my hand calculations to the penny. ]S

# Rhetorical structure theory (Mann and Thompson 87)

## other relations

**Elaboration:** set/member, class/instance/whole/part...

**Contrast:** multinuclear

**Condition:** S presents precondition for N

**Purpose:** S presents goal of action in N

**Sequence:** multinuclear

**Result:** N results from something presented in S

## Problems with RST (cf. Moore and Pollack 92)

how many Rhetorical Relations are there?

how can we use RST in dialogue as well as monologue?

how to incorporate speaker's intentions into RST?

RST does not allow for multiple relations holding between parts of a discourse

RST does not model overall structure of the discourse

# Automatic identification of rhetorical structure

(Marcu 99 and later work)

parser trained on a discourse treebank

- 90 hand-annotated rhetorical structure trees
- Elementary Discourse Units (EDU) linked by Rhetorical Relations (RR)
- parser learns to identify N and S and their RR
- mainly shallow features: lexical, structural, Wordnet-based similarity

discourse segmenter (to identify EDUs)

- trained to segment on hand-labeled corpus (C4.5)
- mainly shallow features: 5-word POS window, presence of discourse markers, punctuation, presence/absence of particular syntactic items
- 96-8% accuracy

# Automatic identification of rhetorical structure

## evaluation of Marcu's parser

EDU identification:	recall 75%, precision 97%
hierarchical structure (related EDUs):	recall 71%, precision 84%
nucleus/satellite labels:	recall 58%, precision 69%
rhetorical relation:	recall 38%, precision 45%

⇒ hierarchical structure easier to identify than rhetorical structure

# Applications

anaphora resolution

text segmentation

text summarization

essay scoring

dialog processing

...

# References

- D. Byron. Resolving Pronominal Reference to Abstract Entities. Proceedings of ACL-02, pp.80–87, 2002
- B. J. Grosz, K. Sparck-Jones, B. L.. Webber. *Readings in Natural Language Processing*, Morgan Kaufmann Publishers, 1986
- B. J. Grosz and C. L.. Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics* 12(3):175–204. 1986
- M. Halliday, R. Hasan. *Cohesion in English*. Harlow: Longman, 1976
- J. Hobbs. Towards an Understanding of Coherence in Discourse, in W. Lehnert & M. Ringle (eds.), *Strategies for Natural-Language Processing*, Hillsdale NJ: Erlbaum, 1982
- W. C. Mann and S. A. Thompson. Rhetorical structure theory: A theory of text organization. Technical Report ISI/RS-87-190, USC/ISI, 1987
- D. Marcu. A decision-based approach to rhetorical parsing. *Proceedings of ACL-99*, pp. 365–372, 1999
- B. L.. Webber. Discourse deixis: Reference to discourse segments. *Proceedings of ACL-88*, pp. 113–123, 1988
- E.B. White. Letters of E.B. White, ed. D.L. Guth, Harper & Row, New York, 1972

slides in German with German examples: <http://rst.80686-net.de/downloads/rstfolien.pdf>

Daniel Marcu's website: <http://www.isi.edu/~marcu/>

RSTTool (graphical interface for marking up of the RST structure of text; perhaps something to consider for homework...):  
<http://www.wagsoft.com/RSTTool/index.html>

# Rhetorical structure theory (Mann and Thompson 87)

**Evidence:** S provides evidence for what N claims

constraints on N:	Reader might not believe N as much as Writer think s/he should
constraints on S:	R believes S or will find it credible
constraints on N&S:	R's comprehending S increases R's belief of N
intention of S:	R's belief of N is increased
locus of effect:	N

**[ This tax calculation software really works. ]N**  
**[ I entered all the figures from my tax return and got a result which agreed with my hand calculations to the penny. ]S**

# Rhetorical structure theory (Mann and Thompson 87)

## Antithesis

constraints on N: on N:

W has positive regard for N

constraints on N+S:

N and S are in contrast

effect:

R's positive regard for N is increased

**[ I'm going to the theater today, ]N [ not to the movies ]S .**

## Contrast (multi-nuklear)

constraints:

No more than two nuclei; the situations in these two nuclei are (a) comprehended as the same in many respects (b) comprehended as differing in a few respects and (c) compared with respect to one or more of these differences

intention:

R recognizes the comparability and the difference(s) yielded by the comparison is being made

**[ Peter likes vegetables, ]N [ and Mary chocolate. ]N**

# Rhetorical structure theory (Mann and Thompson 87)

## Background

on N: R won't comprehend N sufficiently before reading text of S

N&S: S increases the ability of R to comprehend an element in N

intention: R's ability to comprehend N increases

**[ The famous German football player is in hospital, ]N**  
**[ he was injured in the last game. ]S**

# Rhetorical structure theory (Mann and Thompson 87)

## Concession

on N: W has positive regard for N

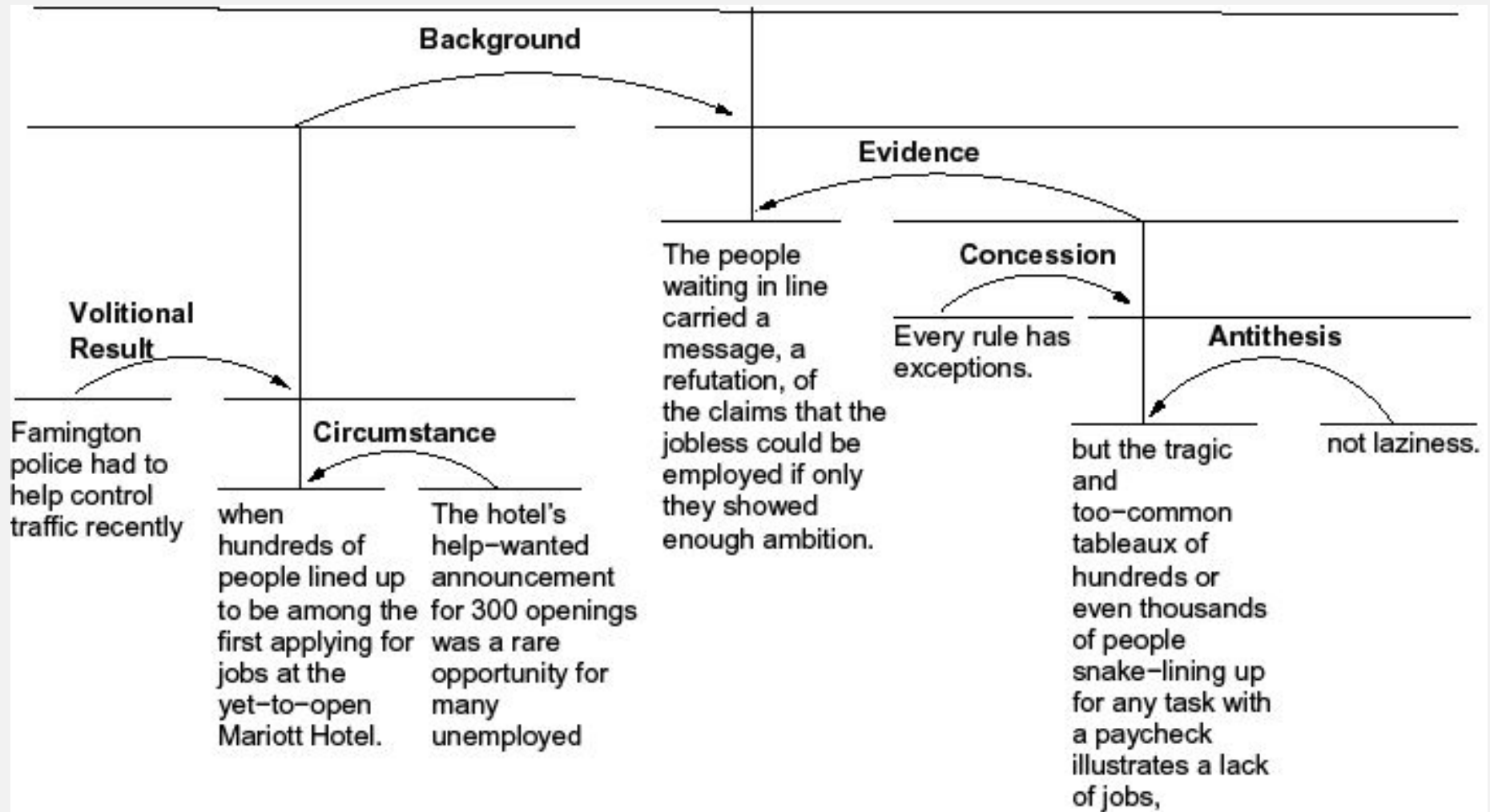
on S: W is not claiming that S does not hold;

constraints on N+S: W acknowledges a potential or apparent incompatibility between N and S; recognizing the compatibility between N and S increases R's positive regard for N

effect: R's positive regard for N is increased

[ Tempting as it may be, ]S

[ **we shouldn't embrace every popular issue that comes along.** ]N



# Rhetorical structure theory (Mann and Thompson 87)

## **Purpose**

on N: N is an activity;

on S: S is a situation that is unrealized

constraints on N+S: S is to be realized through the activity in N

effect: R recognizes that the activity in N is initiated in order to realize S

[ To see which Syncom diskette will replace the one you're using now ]S  
[ send for our free 'Flexi-Finder' selection guide. ]N

# Intentional Approach

utterances : actions

speaker's plan wrt. communicating intentions ties the discourse together

discourse understanding : recognizing speaker's intentions

# Intentional Approach (Grosz and Sidner 86)

three dimensions of discourse

linguistic structure : the utterances

intentional structure : hierarchy of intentions

attentional structure : model of objects, properties and relations  
that are *salient* at each point in discourse  
(dynamically changing)

# Intentional Approach (Grosz and Sidner 86)

Linguistic structure

discourse segments + relations that hold between them

# Intentional Approach (Grosz and Sidner 86)

## Linguistic structure

discourse segments + relations that hold between them

(para-) linguistic expressions reflect discourse structure

cue phrases, aspect, tense, intonation, gesture

discourse structure constraints discourse interpretation

anaphora resolution

# Intentional Approach (Grosz and Sidner 86)

## Intentional structure

### Discourse Purpose (DP)

purpose/intention held by discourse initiator

e.g. make hearer:

intend to perform a task,

believe a fact,

believe that one fact supports another fact,

identify an object,

identify a property of an object

assumption: one per discourse

# Intentional Approach (Grosz and Sidner 86)

Intentional structure

Discourse Segment Purpose (DSP)

how given segment contributes to DP

# Intentional Approach (Grosz and Sidner 86)

## Intentional structure

### Hierarchy of intentions

#### dominance

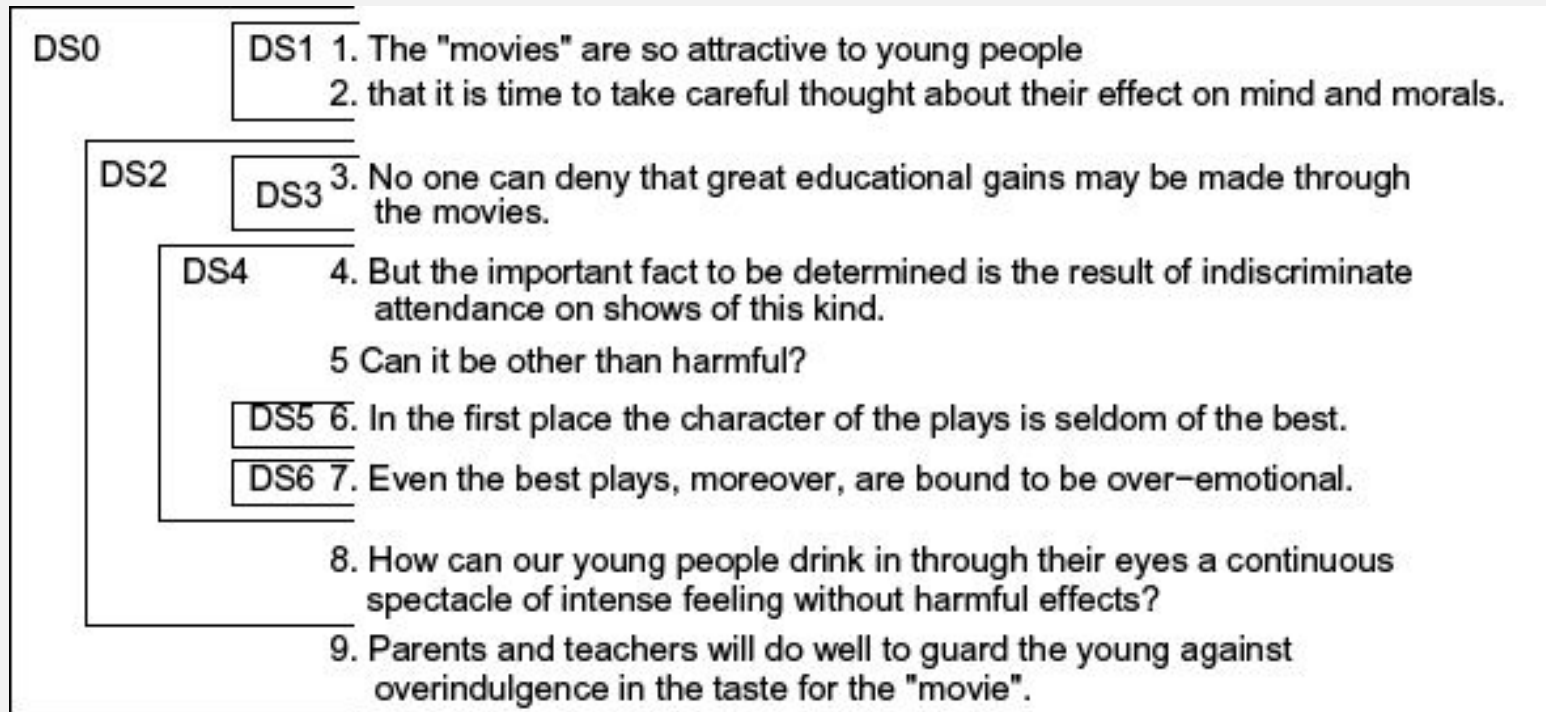
DSP1 *dominates* DSP2 if satisfying DSP2 is intended to provide part of satisfaction of DSP1

#### precedence

DSP1 *precedes* DSP2 if DSP1 must be satisfied before DSP2

1. The "movies" are so attractive to young people
2. that it is time to take careful thought about their effect on mind and morals.
  
3. No one can deny that great educational gains may be made through the movies.
4. But the important fact to be determined is the result of indiscriminate attendance on shows of this kind.
- 5 Can it be other than harmful?
6. In the first place the character of the plays is seldom of the best.
7. Even the best plays, moreover, are bound to be over-emotional.
  
8. How can our young people drink in through their eyes a continuous spectacle of intense feeling without harmful effects?
9. Parents and teachers will do well to guard the young against overindulgence in the taste for the "movie".

(thanks to Caroline Sporleder for the images)



**DSP 0: make reader believe that young people should be guarded from overindulgence in the movies**

**DSP 1: make reader believe that it's time to consider the harmful effects of movies**

**DSP 2: make reader believe that young people can't be exposed to the movies without harmful effects**

etc. . . .

**DSP 0 dominates DSP 1, DSP 0 dominates DSP 1, . . .**

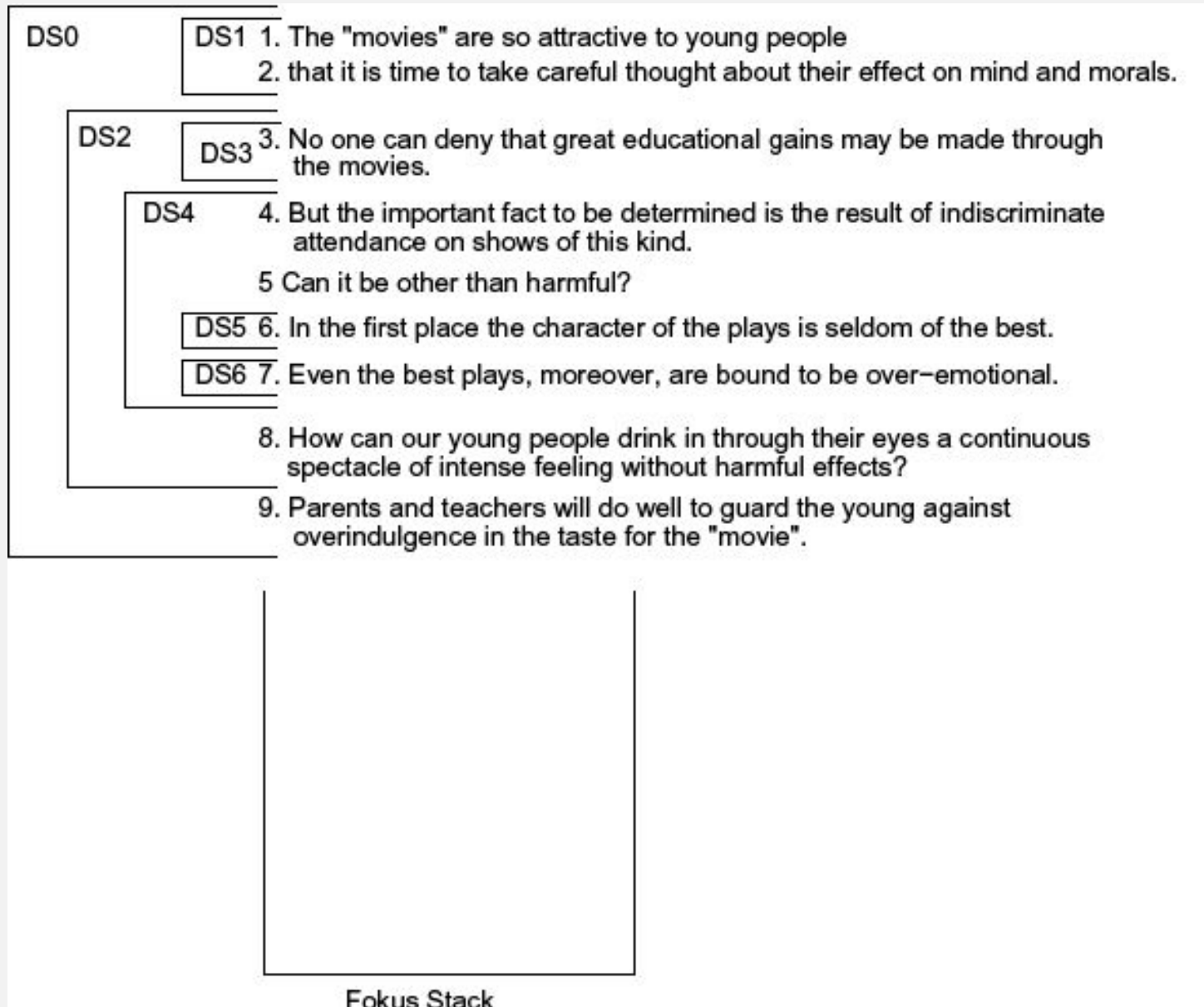
# Intentional Approach (Grosz and Sidner 86)

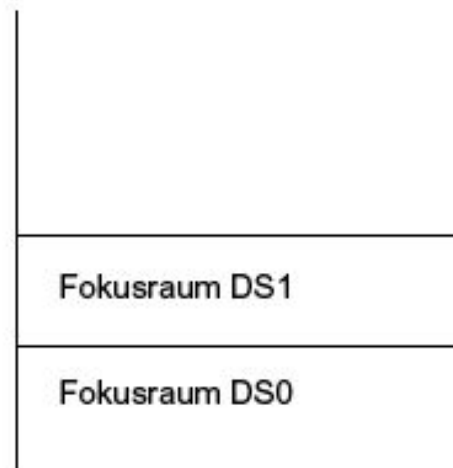
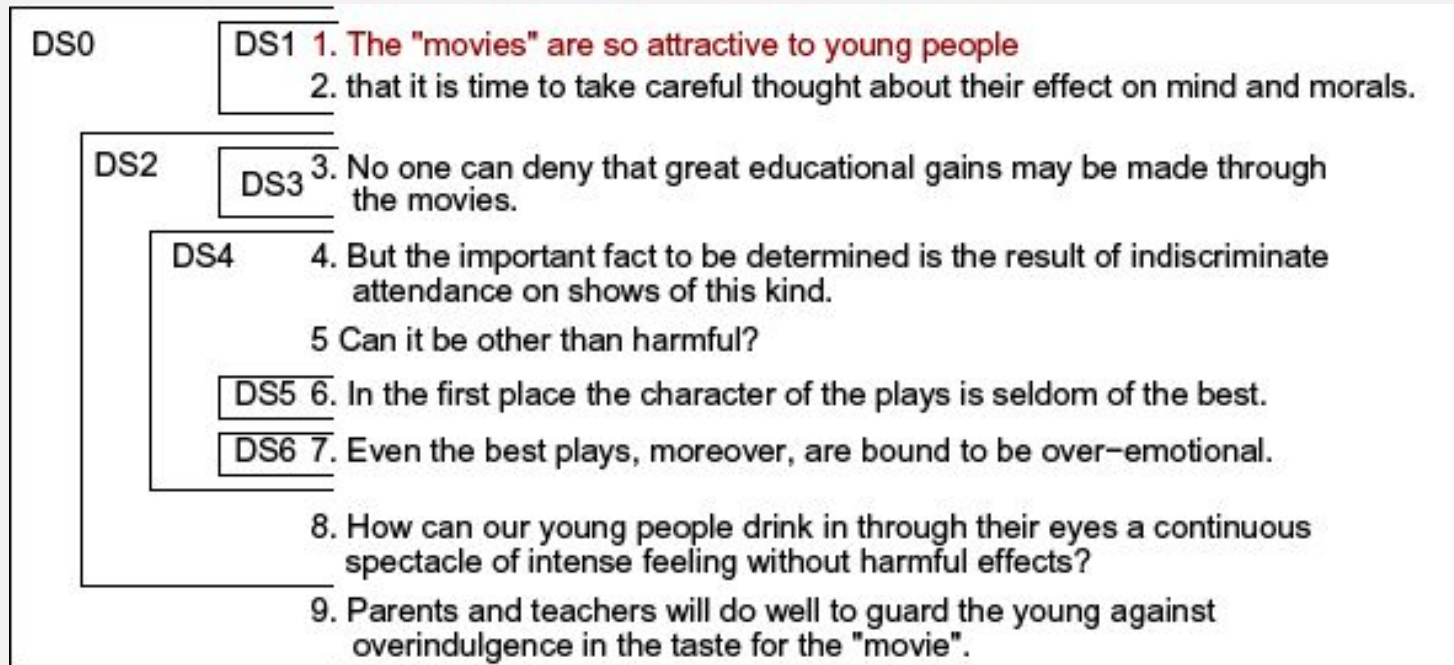
## Attentional structure

participants' focus of attention (what is „attend to”)

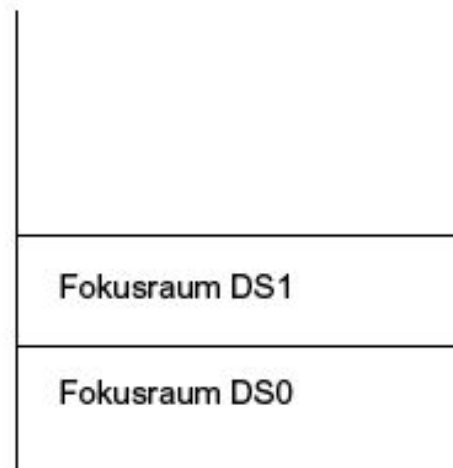
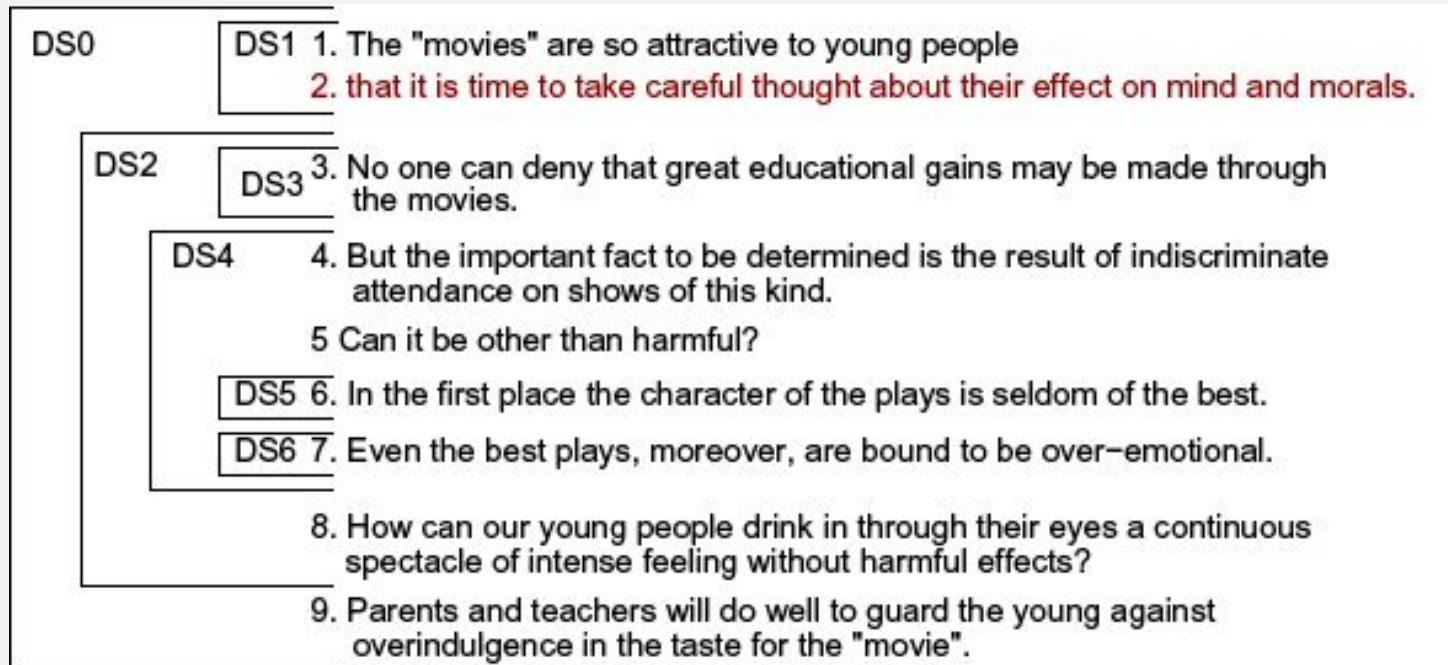
modeled by *focus spaces*: objects and relations in focus

changes: insertion and deletion rules

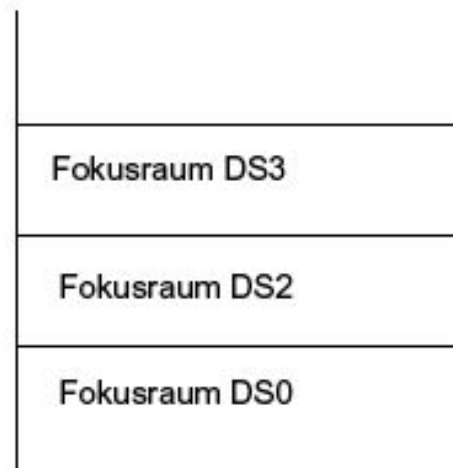
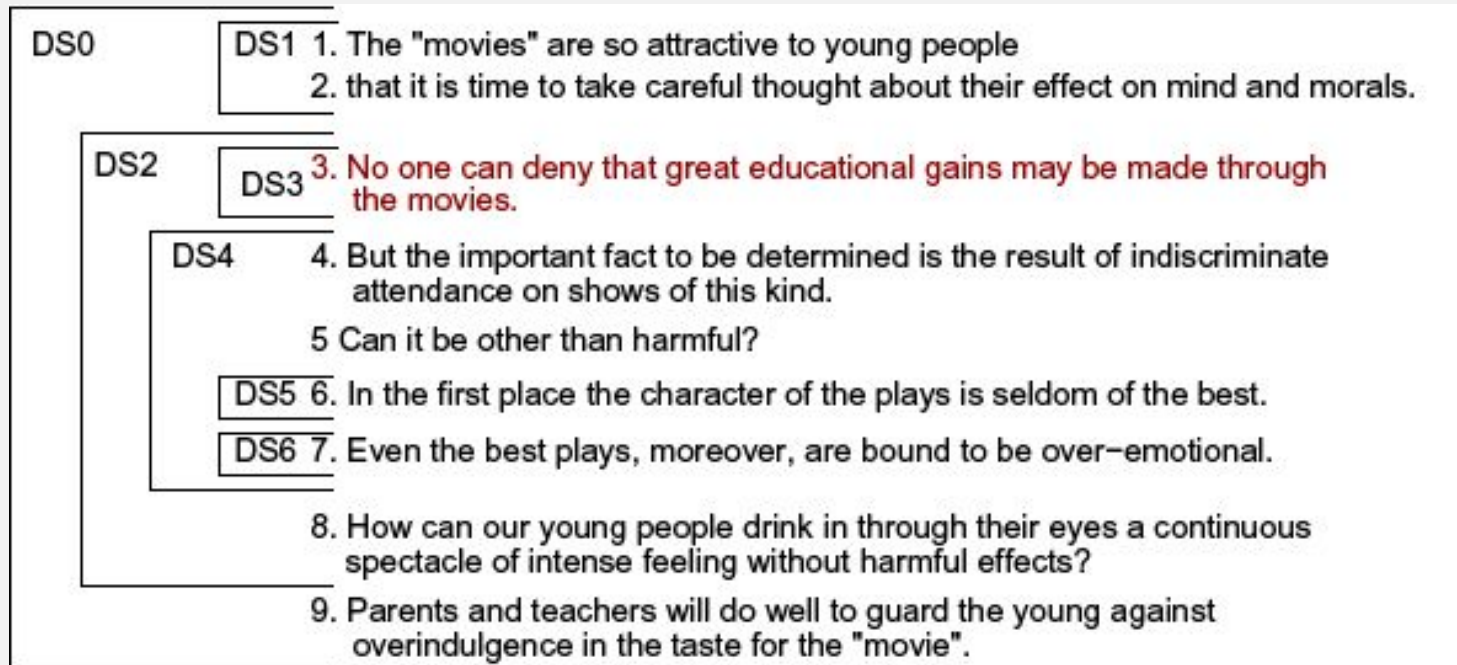




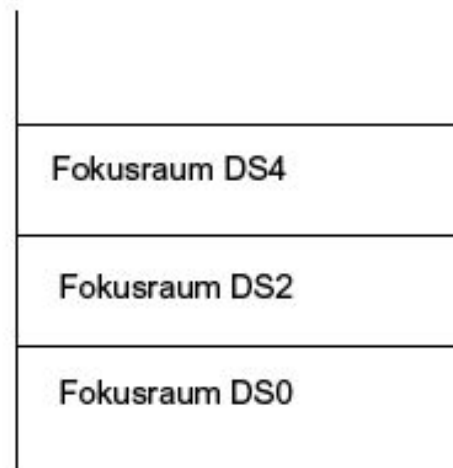
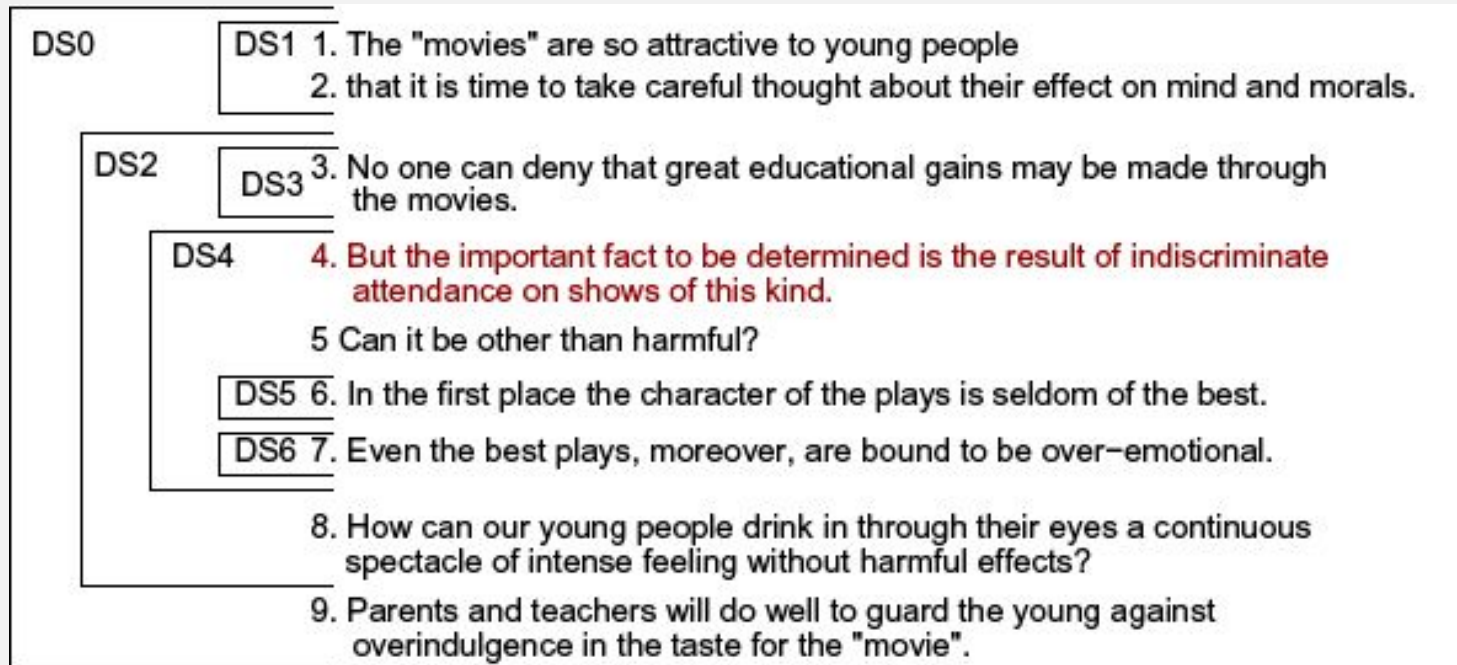
Fokus Stack



Fokus Stack



Fokus Stack



Fokus Stack

1. Es hatte etwas Versöhnliches, wie die Deutsche Telekom und die Gewerkschaft Verdi gestern den Schulterschluss demonstrierten.
2. Jetzt ging es darum, das gemeinsam erreichte Ergebnis in ein positives Licht zu rücken.
3. Einen Verlierer soll es nicht geben.
4. Es ist aber vor allem die Telekom, die die Einigung für sich als Durchbruch reklamieren kann.
5. Das beharrliche Management hat damit sein Ziel erreicht, die Kosten im Konzern künftig maßgeblich zu senken.
6. Außerdem werden die Beschäftigten mehr arbeiten,
7. was die Produktivität erhöht.
8. Der schwerfällige T-Dampfer nimmt langsam Fahrt auf.
9. Dass der Aktienkurs nach der Einigung deutlich kletterte, spricht Bände.