

# From connectives to discourse relations

## An analysis of CONTRAST

Manfred Stede  
Applied Computational Linguistics  
Dept. of Linguistics & SFB 1087  
Universität Potsdam/Germany



## Overview

- Motivation: [connective-lex.info](http://connective-lex.info)
- Flashback: Some contrastive German connectives
- Towards a map
- Enhancing the connective lexicon
- Application: Argumentation mining
- Conclusions

Lexicon Selection <

<input type="checkbox"/> Arabic	Arabic	about
<input type="checkbox"/> DiMLex-Bangla	Bangla	about
<input type="checkbox"/> CzeDLex	Czech	about
<input type="checkbox"/> DiMLex	German	about
<input type="checkbox"/> DisCoDict	Dutch	about
<input checked="" type="checkbox"/> Eng-DiMLex	English	about
<input type="checkbox"/> LDM-PT	Portuguese	about
<input type="checkbox"/> LexConn	French	about
<input type="checkbox"/> LICO	Italian	about

Search Options

Filter Word

Syntactic Category

cco  csu  adv  prep  other

Discourse Relation (PDTB3)

COMPARISON  
 Contrast  
 Similarity  
 Concession (  Arg1-/  Arg2-as-denier )

CONTINGENCY  
 Cause (  Reason/  Result )  
 Condition (  Arg1-/  Arg2-as-cond )  
 Negative condition (  Arg1-/  Arg2-as-negcond )  
 Purpose (  Arg1-/  Arg2-as-goal )

EXPANSION  
 Conjunction  
 Disjunction  
 Equivalence  
 Instantiation  
 Level-of-detail (  Arg1-/  Arg2-as-detail )  
 Substitution (  Arg1-/  Arg2-as-subst )  
 Exception (  Arg1-/  Arg2-as-except )  
 Manner (  Arg1-/  Arg2-as-manner )

TEMPORAL  
 Synchronous  
 Asynchronous (  Precedence/  Succession )  
 +Belief  +SpeechAct  Other

COMPARISON:Concession:Arg2-as-denier  
*Ms. Levine had never been fired, although she had stopped working at the restaurant*

COMPARISON:Concession:Arg1-as-denier  
*Although started in 1965, Wedtech didn't really get rolling until 1975*

prep  
**COMPARISON**  
*Mr. Morishita's main business certainly appears to be thriving, although he won't disclose numbers*

EXPANSION:Exception  
*Although the Treasury will announce details of the November refunding tomorrow, it could be delayed if Congress and President Bush fail to increase the Treasury's borrowing capacity*

COMPARISON:Concession  
*Worth watching, although the music has lasted better than the plot or the humor*

COMPARISON:Contrast  
*And Beauregard was mentioned twice -- although very briefly and in passing  
 Although viewership soars when big news breaks, it ebbs during periods of calm  
 He doesn't fully succeed, although Mr. Strasheim lists an array of interest-rate scenarios*

COMPARISON:Concession:Arg2-as-denier  
*Ms. Levine had never been fired, although she had stopped working at the restaurant*

COMPARISON:Concession:Arg1-as-denier  
*Although started in 1965, Wedtech didn't really get rolling until 1975*

thoug Variants Synonyms English – Eng-DiMLex 3

other  
**COMPARISON**  
*But even though companies are paying more taxes, many are still paying less than the statutory rate*

COMPARISON:Concession  
*teachers for the most part teach in English, even though it is a foreign language for most Philippine children*

COMPARISON:Contrast  
*Giuliani has always been pro-choice, even though he has personal reservations  
 resulting in a two-year plan to reduce the deficit by more than \$76 billion -- even though the deficit actually rose by nearly \$12 billion during that period  
 and nuclear power (even though it would counteract the greenhouse effect*

COMPARISON:Concession:Arg2-as-denier  
*I'm more favorably impressed by "no-load life," even though it turns out to be low-load life*

COMPARISON:Concession:Arg1-as-denier  
*that hung over parts of the factory, even though exhaust fans ventilated the area*

# Connective: Definition

x is a connective if

1. x cannot be inflected,
2. the meaning of x is a two-place relation,
3. the arguments of this relation are abstract objects (propositions, events, states, or processes),
4. the arguments can be expressed as clauses,
5. x is a fixed expression and cannot be freely modified by inserting other expressions,
6. x is not semantically compositional w.r.t. to its components (~~particularly when~~, etc.).

## DiMLex-Ger

- First development in the late 1990s
- Inspiration and inventarization:  
Pasch et al.: *Handbuch der deutschen Konnektoren*. De Gruyter 2003 (Vol. 2: 2016)
- Major extension with addition of PDTB senses in 2015
- Current size: 275 entries

T. Scheffler, M. Stede: Adding semantic senses to a lexicon of German connectives. Proc. of LREC conference, 2016

## DiMLex-Eng

- Starting point: 100 connectives from the Penn Discourse Treebank corpus

D. Das, T. Scheffler, P. Bourgonje, M. Stede: Constructing a lexicon of English discourse connectives. Proc. of the SIGDIAL conference, Melbourne, 2018

# DiMLex-Eng

- Starting point: 100 connectives from the Penn Discourse Treebank corpus
- 2nd source: RST signalling corpus
  - Extends the RST Discourse Treebank with signal annotations
  - Decide which signal types correspond to our definition of *connective*
  - 71/100 PDTB connectives also appear in RST-SC
  - 46 RST-SC connectives do not appear in PDTB

D. Das, M. Taboada: RST Signalling Corpus: a corpus of signals of coherence relations.  
*Language Resources and Evaluation* 52(1), 2018

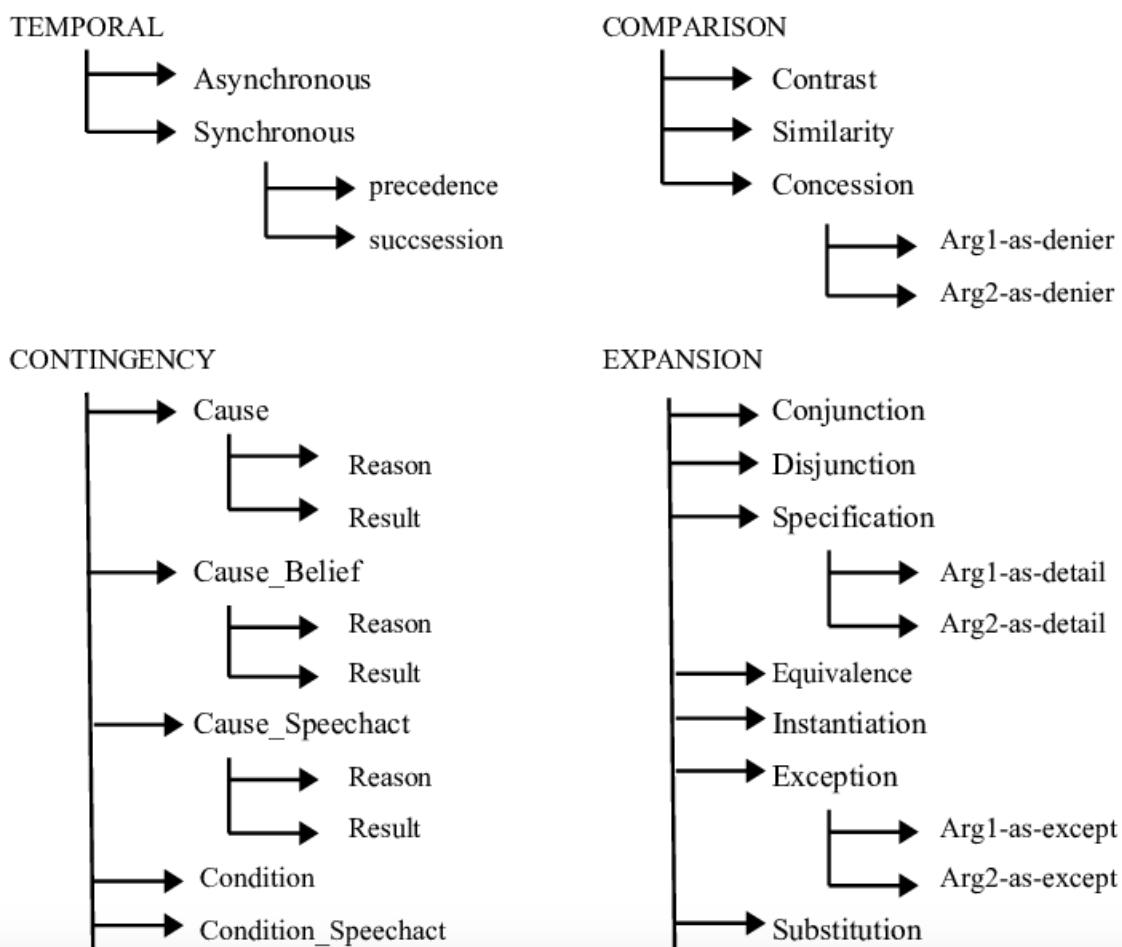
# DiMLex-Eng

- Starting point: 100 connectives from the Penn Discourse Treebank corpus
- 2nd source: RST signalling corpus
- 3rd source: List of relational indicators
  - Induce the most frequent word n-grams per coherence relation in the RST-DT
  - Decide which signal types correspond to our definition of *connective*
  - => 5 additional entries

Or Biran, Owen Rambow: Identifying justifications in written dialogs by classifying text as argumentative.  
*International Journal of Semantic Computing* 5(4), 2011

# connective-lex

- Features in the public version
  - syntactic category cco, csu, adv, prep, other
  - PDTB relation(s)
  - argument order
  - orthographic variants
  - non-connective readings
- Example: although
  - syntax: csu / prep
  - PDTB relations:
    - comparison
    - comparison:concession
    - comparison:contrast
    - expansion:exception
  - order: a1-a2 / a2-a1
  - orth. variants: though
  - non-conn readings: %



# Coherence relations: CONTRAST

- **PDTB:**
  - Expansion:Substitution
  - Expansion:Exception
  - Comparison:Contrast
  - Comparison:Concession
- **RST:**  
Contrast (multinuclear), Antithesis, Concession
- **SDRT/Lexconn:**  
Contrast (coord), Concession (subord), Opposition
- **SDRT/STAC:**  
Contrast, Correction

## Overview

- Motivation: [connective-lex.info](http://connective-lex.info)
- Flashback: [Some contrastive German connectives](#)
- Towards a map
- Enhancing the connective lexicon
- Application: Argumentation mining
- Conclusions

## A **but** B

- In Romania you don't pay with the Euro **but** with the Leu.
- My daughter attends university, **but** my son is still in high school.
- We tried hard to arrive on time, **but** the bus was terribly slow!
- His grades are OK, **but** we really shouldn't hire him.

## Whereas A, B

- \* **Whereas** in Romania you don't pay with the Euro, you pay with the Leu.
- **Whereas** my daughter attends university, my son is still in high school.
- \* **Whereas** we tried hard to arrive on time, the bus was terribly slow!
- \* **Whereas** his grades are OK, we really shouldn't hire him.

## Although A, B

- \* **Although** in Romania you don't pay with the Euro, you pay with the Leu.
- ?? **Although** my daughter attends university, my son is still in high school.
- \* **Although** we tried hard to arrive on time, the bus was terribly slow!
- **Although** his grades are OK, we really shouldn't hire him.

## A; instead B

- In Romania you don't pay with the Euro; **instead** you pay with the Leu.
- \* My daughter attends university; **instead** my son is still in high school.
- \* We tried hard to arrive on time, **instead** the bus was terribly slow!
- \* His grades are OK; **instead** we really shouldn't hire him.

# Some German contrastive connectives

- **aber** (*but*)  
„the prototypical adversative connective“
- **doch, jedoch** (*but*)  
equivalent to ‚aber‘?
- **hingegen** (*however, on the other hand*)
- **allerdings** (*though*)
- **dennoch** (*still*)

English translations from Collins dictionary

M. Stede: Kontrast im Diskurs. In: H. Blühdorn et al.: Brücken schlagen. Grundlagen der Konnektorensemantik. De Gruyter 2004

## Buscha: Lexikon deutscher Konjunktionen

- **aber** (adversativ)  
... Sie gibt einen Gegensatz an, der sich auf die Bedeutung der von den satzwertigen Teilen bezeichneten Sachverhalte bezieht und eine semantische Opposition zum Ausdruck bringt. Sie kann aber auch einen Ggs. angeben, der sich aus der Bewertung der Sachverhalte ergibt, in der Negation einer Erwartung besteht und eine Einschränkung darstellt.
- **doch** (adversativ)  
... ist bedeutungsgleich mit *aber*, gehört aber ... nur der gehobenen Stilebene an. Sie ist in literarischen und wissenschaftlichen Texten als stilistische Variante von *aber* üblich, in einer mündlichen Alltagskommunikation wirkt sie allzu gewählt (vor allem in der bedeutungsmäßig stärkeren Form *jedoch*).

# Test for substitutability

(Knott/Dale 1994)

1. Remove cue A; insert cue B into the clause, at any appropriate position.
2. If necessary, alter the punctuation.
3. If necessary, other cues in the text are changed (if dependent)

- A and B are substitutable if
  - new discourse describes the same events
  - new discourse achieves the same goals
- It does not matter if
  - stylistic change is introduced
  - cues are appropriate for text portions of different sizes
  - different amounts of background knowledge are assumed

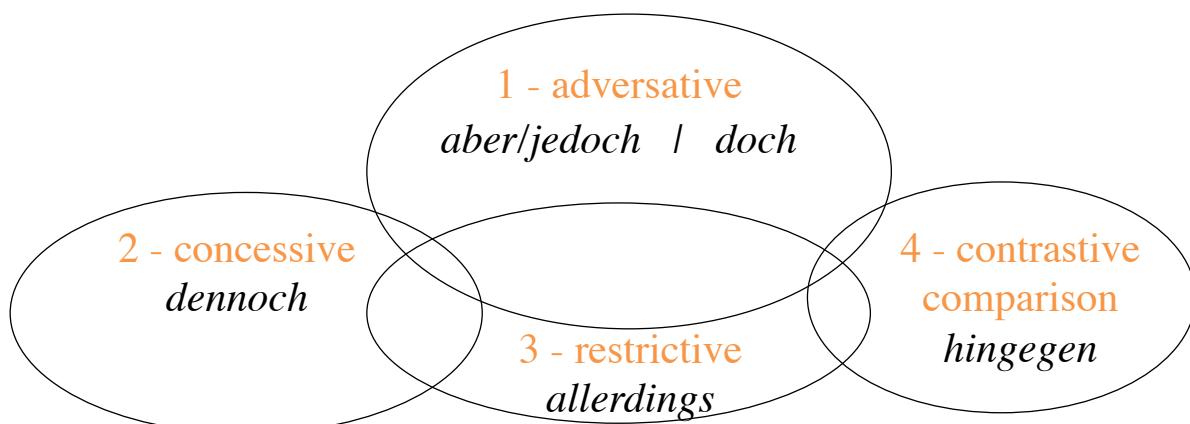
## Probing the Similarities

- **Replacement test:** In context X, which connective can replace the one that was originally used in X?
- **Data:** Newspaper text in COSMAS (IDS Mannheim)
- 50 instances per connective studied

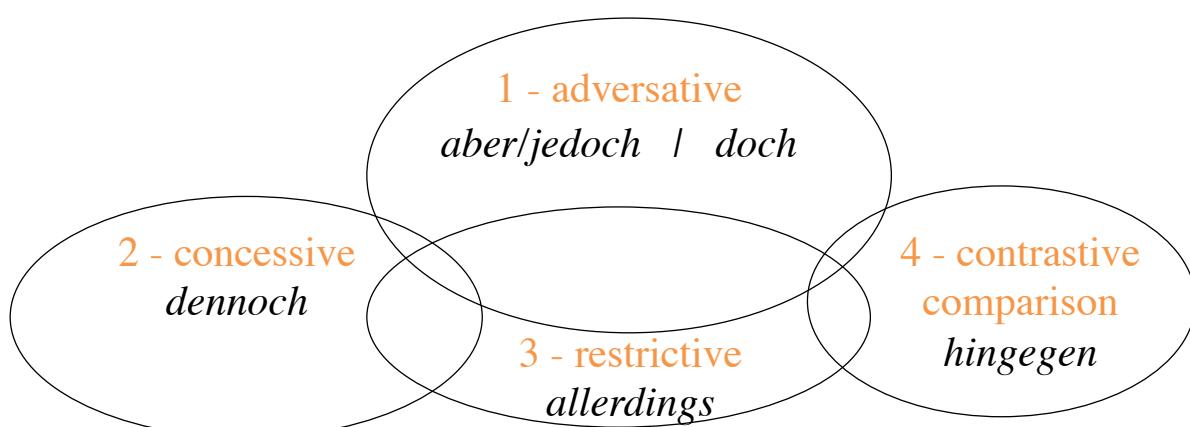
vierzehn Tage vorher, als die Lok 4029t am Zughaken hatte, lag das Übergewicht noch deutlich bei den Pessimisten. DOCH Leistung überzeugt.

**+aber, -allerdings, -dennoch, +DOCH, -hingegen, +jedoch**

# Result



# Result



1 – Dieser Computer ist nicht toll, aber recht preiswert.

1 – *This computer is not fabulous, but it is affordable.*

2 – Dieser Computer ist eher langsam. Dennoch werde ich ihn kaufen.

2 – *This computer is a bit slow. Still, I'm gonna buy it.*

3 – Dieser Computer ist ein Bestseller. Allerdings hat er etwas wenig Hauptspeicher.

3 – *This computer is a bestseller. It's a bit short of RAM, though.*

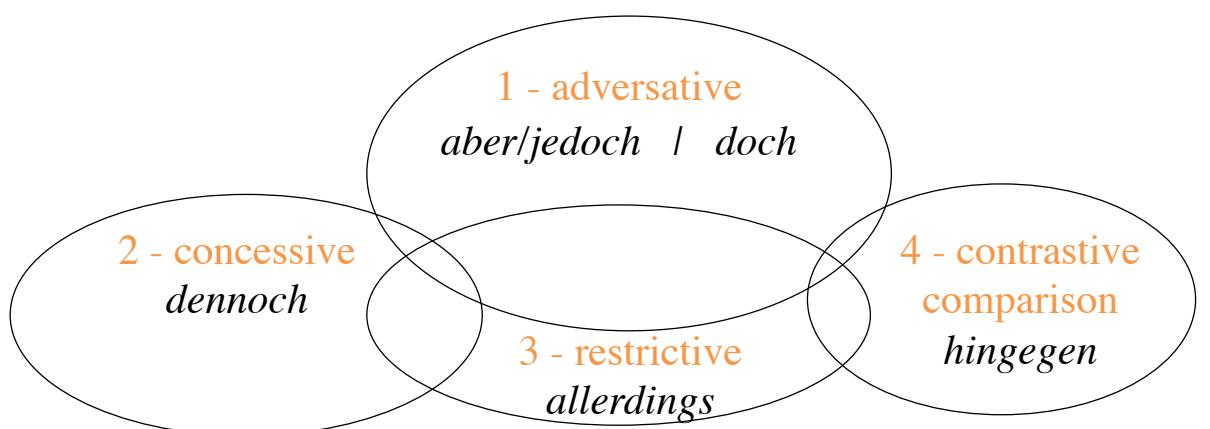
4 – Dieser Computer hat eine NVIDIA Grafikkarte. Hingegen hat der dort drüben eine GeForce.

4 – *This computer has an NVIDIA graphics card. While that one over there has a GeForce.*

# Overview

- Motivation: connective-lex.info
- Flashback: Some contrastive German connectives
- Towards a map
- Enhancing the connective lexicon
- Application: Argumentation mining
- Conclusions

## Add 'substitution'



1 – Dieser Computer ist nicht toll, aber recht preiswert.

1 – *This computer is not fabulous, but it is affordable.*

2 – Dieser Computer ist eher langsam. Dennoch werde ich ihn kaufen.

2 – *This computer is a bit slow. Still, I'm gonna buy it.*

3 – Dieser Computer ist ein Bestseller. Allerdings hat er etwas wenig Hauptspeicher.

3 – *This computer is a bestseller. It's a bit short of RAM, though.*

4 – Dieser Computer hat eine NVIDIA Grafikkarte. Hingegen hat der dort drüben eine GeForce.

4 – *This computer has an NVIDIA graphics card. While that one over there has a GeForce.*

5 – Dieser Computer hat keine Grafikkarte, sondern einen on-board Chip.

5 – *This computer doesn't have a graphics card but an on-board chip.*

# Concession (*dennoch, obwohl, ...*)

- **Protoypical connective:** *obwohl* (although)  
,Obwohl der Vortrag interessant war, bin ich dauernd eingeschlafen.'  
(*Although the talk was interesting, I fell asleep all the time.*)
- König 94: clauses describe „instances of situations that normally don't go together“
- if p, then normally not-q  
here, both p and q

## obwohl<sub>V2</sub>

- Standard *obwohl*:  
**subordinating conjunction**, requires verb-last clause
- In colloquial spoken language also:  
**discourse marker** joining to V2 main clauses
- *Ich kaufe mir eine Bratwurst; obwohl, ein Salat wäre noch besser.*  
,I'm gonna buy a sausage - **actually**, a salad would be even better.'
- Function: **restrict** or **cancel** the previous illocution

# Corpus evidence

- In speech,  $obwohl_{V2}$  joins two separate intonation units and is surrounded by pauses (likewise:  $wobei_{V2}$ )
- **Problem:** Corpora of colloquial German are small 😞
- **Proxy:** One-month German **Twitter** snapshot (Scheffler 2014)
  - 24 mio tweets, capturing around 90% of all German tweets sent in March 2013
- Use **pause markers** from Moraldo (2012): , / . / : / - / ...
- To estimate the **unmarked** V2 instances, manually inspect 200 *Obwohl* and 200 *obwohl* samples

## Frequencies

Twitter

	Total	V2	+ .	+ ... (2-5)	+ ,	+ :	+ -	no pause
obwohl	19.198		3%	27%	18%	3%	1%	47%
Obwohl	7.505		8%	24%	28%	6%	2%	33%
o Obw.	26.703	12.9%	6%	25%	24%	5%	2%	38%
Mor12	1.032	16.7%	6%	26%	25%	5%	4%	34%

# Frequencies

Twitter

	Total	V2	+ .	+ ... (2-5)	+ ,	+ :	+ -	no pause
obwohl	19.198	7.2%	3%	27%	18%	3%	1%	47%
Obwohl	7.505	27.4%	8%	24%	28%	6%	2%	33%
o Obw.	26.703	12.9%	6%	25%	24%	5%	2%	38%
Mor12	1.032	16.7%	6%	26%	25%	5%	4%	34%

## „Standard“ language: DWDS.de

- DWDS *Kernkorpus* (1900-1999)
- Fiction, Newspaper, Science, Gebrauchstexte
- Total: 121 million tokens
- Hits are filtered for parentheticals etc.

# Frequencies

Twitter

	Total	V2	+ .	+ ... (2-5)	+ ,	+ :	+ -	no pause
obwohl	19.198	7.2%	3%	27%	18%	3%	1%	47%
Obwohl	7.505	27.4%	8%	24%	28%	6%	2%	33%
o Obw	26.703	12.9%	6%	25%	24%	5%	2%	38%
Mor12	1.032	16.7%	6%	26%	25%	5%	4%	34%

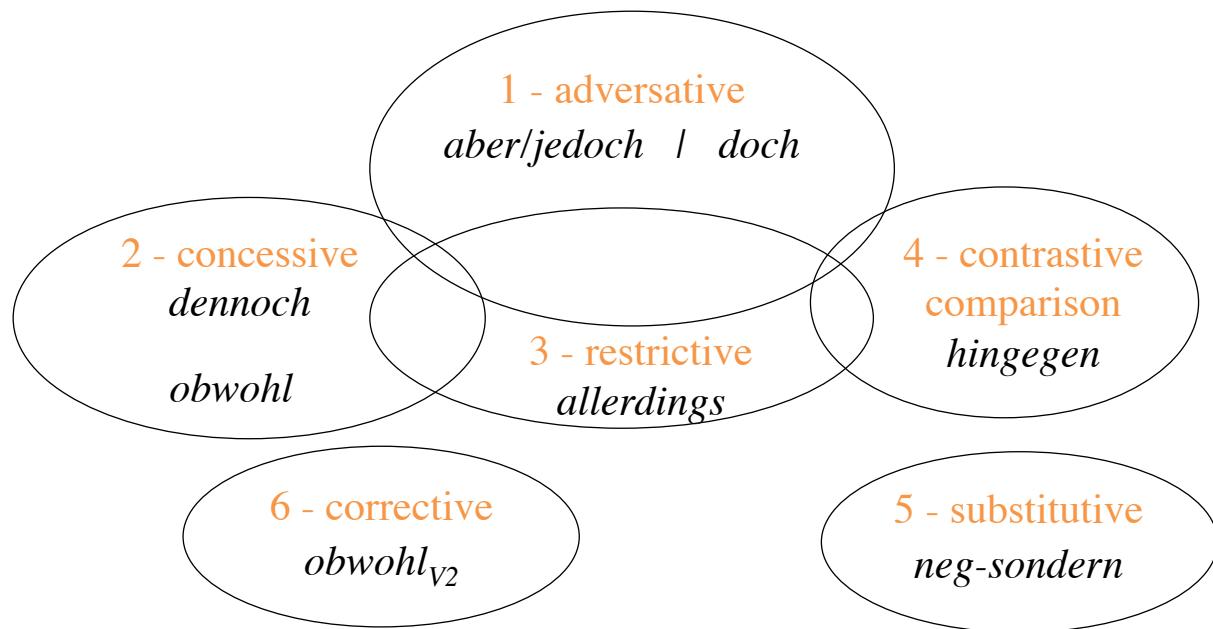
DWDS Kernkorpus

	Total		+ .	+ ... (2-5)	+ ,	+ :	+ -	other
o Obwohl	14.819		0 / 0%	0 / 0%	19 / 0.1%	3 / 0%	4 / 0%	99.8%
w Wobei	15.336		0 / 0%	0 / 0%	0 / 0%	0 / 0%	0 / 0%	100%

## Summary: obwohl<sub>V2</sub>

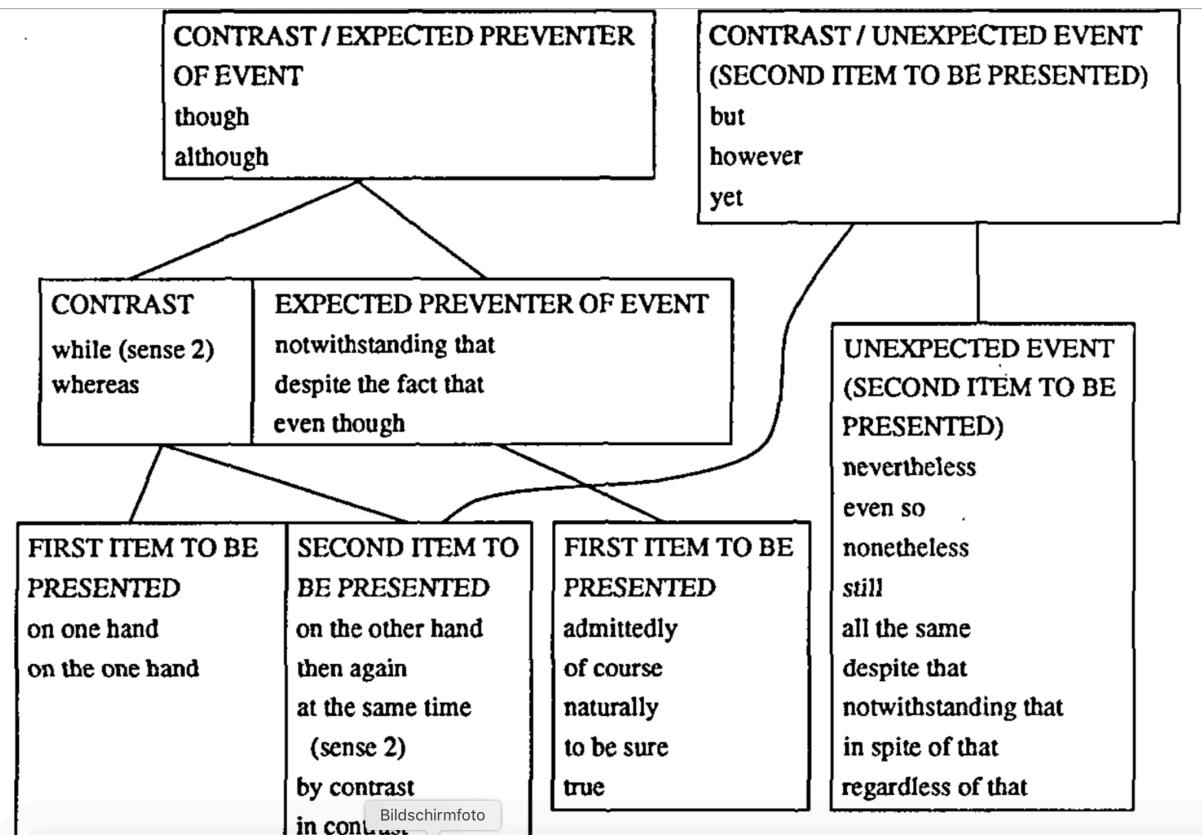
- A subordinating conjunction diachronically spun off a discourse marker
- **Change in meaning:** From concession to correction/restriction
- Same story, but less frequent: *wobei*
- Apparently (still) restricted to **colloquial language**
- In Twitter: 13-16% of *obwohl* are V2
- [!] Our definition of connective covers obwohl<sub>V2</sub>

# Add 'obwohl<sub>V2</sub>'



6 – OK, ich werde diesen Computer kaufen. Obwohl - lass nochmal die von Dell angucken.

6 – OK, I'm gonny buy this computer. Though - let's first take another look at the Dell machines.



# Compatibility

- General: *but, however, yet / though, although*
- Strict separation between
  - **contrastive comparison** (contrast: *while, whereas*)
  - **concessive** (exp. preventer of event: *even though, despite the fact that*)
- English *although* is more flexible than German *obwohl*; it also covers **restriction**
  - *Our university covers all the important areas of study, although there's no medicine*
  - *??Unsere Universität bietet alle wichtigen Fächer an, obwohl es keine Medizin gibt.*
  - *Unsere Universität bietet alle wichtigen Fächer an, allerdings gibt es keine Medizin.*

# Compatibility

- Some Differences:
  - **Substitution** is not covered in this part of the Knott/Dale taxonomy
  - Knott/Dale emphasize hierarchy (directed substitutability)
  - Knott/Dale distinguish *even though, despite the fact that* from *nevertheless, despite that* (partly on the grounds of linear order)

# Overview

- Motivation: connective-lex.info
- Flashback: Some contrastive German connectives
- Towards a map
- Enhancing the connective lexicon
- Application: Argumentation mining
- Conclusions



Connective-Lex.info   [About / Imprint](#)   [Help](#)

**Lexicon Selection**

<input type="checkbox"/> Arabic	Arabic	about
<input type="checkbox"/> DiMLex-Bangla	Bangla	about
<input type="checkbox"/> CzeDlex	Czech	about
<input type="checkbox"/> DiMLex	German	about
<input type="checkbox"/> DisCoDict	Dutch	about
<input checked="" type="checkbox"/> Eng-DiMLex	English	about
<input type="checkbox"/> LDM-PT	Portuguese	about
<input type="checkbox"/> LexConn	French	about
<input type="checkbox"/> LICO	Italian	about

**Search Options**

Filter: Word

**Syntactic Category**

cco  csu  adv  prep  other

**Discourse Relation (PDTB3)**

COMPARISON  
 Contrast  
 Similarity  
 Concession (  Arg1-/  Arg2-as-denier )

CONTINGENCY  
 Cause (  Reason/  Result )  
 Condition (  Arg1-/  Arg2-as-cond )  
 Negative condition (  Arg1-/  Arg2-as-negcond )  
 Purpose (  Arg1-/  Arg2-as-goal )

EXPANSION  
 Conjunction  
 Disjunction  
 Equivalence  
 Instantiation  
 Level-of-detail (  Arg1-/  Arg2-as-detail )  
 Substitution (  Arg1-/  Arg2-as-subst )  
 Exception (  Arg1-/  Arg2-as-except )  
 Manner (  Arg1-/  Arg2-as-manner )

TEMPORAL  
 Synchronous  
 Asynchronous (  Precedence/  Succession )  
 +Belief    +SpeechAct    Other

**COMPARISON:Concession:Arg2-as-denier**  
*Ms. Levine had never been fired, although she had stopped working at the restaurant*

**COMPARISON:Concession:Arg1-as-denier**  
*Although started in 1965, Wedtech didn't really get rolling until 1975*

**prep**  
**COMPARISON**  
*Mr. Morishita's main business certainly appears to be thriving, although he won't disclose numbers*

**EXPLANATION:Exception**  
*Although the Treasury will announce details of the November refunding tomorrow, it could be delayed if Congress and President Bush fail to increase the Treasury's borrowing capacity*

**COMPARISON:Concession**  
*Worth watching, although the music has lasted better than the plot or the humor*

**COMPARISON:Contrast**  
*And Beauregard was mentioned twice -- although very briefly and in passing  
 Although viewership soars when big news breaks, it ebbs during periods of calm  
 He doesn't fully succeed, although Mr. Straszheim lists an array of interest-rate scenarios*

**COMPARISON:Concession:Arg2-as-denier**  
*Ms. Levine had never been fired, although she had stopped working at the restaurant*

**COMPARISON:Concession:Arg1-as-denier**  
*Although started in 1965, Wedtech didn't really get rolling until 1975*

though	Variants	Synonyms	English – Eng-DiMLex 3
other			
<b>COMPARISON</b> <i>But even though companies are paying more taxes, many are still paying less than the statutory rate</i>			
<b>COMPARISON:Concession</b> <i>teachers for the most part teach in English, even though it is a foreign language for most Philippine children</i>			
<b>COMPARISON:Contrast</b> <i>Giuliani has always been pro-choice, even though he has personal reservations    resulting in a two-year plan to reduce the deficit by more than \$76 billion -- even though the deficit actually rose by nearly \$12 billion during that period    and nuclear power (even though it would counteract the greenhouse effect)</i>			
<b>COMPARISON:Concession:Arg2-as-denier</b> <i>I'm more favorably impressed by "no-load life," even though it turns out to be low-load life</i>			
<b>COMPARISON:Concession:Arg1-as-denier</b> <i>that hung over parts of the factory, even though exhaust fans ventilated the area</i>			

# connective-lex

- Features in the public version
  - syntactic category cco, csu, adv, prep, other
  - PDTB relation(s)
  - argument order
  - orthographic variants
  - non-connective readings
- Potential new features
  - more fine-grained syntax
  - modifiability
  - argument linking (semantic roles: causing / caused, etc.)
  - sem/prag function

## DiMLex-Ger: Substitution

anstatt (dass)

anstelle dessen

sondern

statt

stattdessen

vielmehr

# DiMLex-Ger: Contrast

aber	doch	unterdessen
alldieweil	einerseits	wiederum
allein	freilich	wogegen
allerdings	gleichzeitig	wohingegen
andererseits	hingegen	während
bloß dass	immerhin	währenddessen
dafür	indessen	zugleich
dafür dass	jedoch	zum anderen
dagegen	nur	zum einen
dahingegen	nur dass	
dementgegen	sosehr	
demgegenüber	und	

# DiMLex-Ger: Concession

aber	entgegen	obzwar
abgesehen davon	freilich	sosehr
allerdings	gleichwohl	trotz
als ob	immerhin	trotzdem
angesichts	jedoch	unbeschadet dessen
bloß	mangels	ungeachtet dessen,
bloß dass	mithin	wenn .. auch
dabei	nichtsdestominder	wenngleich
davon abgesehen, dass	nichtsdestoweniger	wiederum
dennoch	obgleich	wobei
dessen ungeachtet	obschon	wohingegen
doch	obwohl	zwar

# Simplifications

- For present purposes of sem/prag grouping,
  - ignore prepositions
  - fuse quasi-synonyms (*obwohl*, *obschon*, *obzwar*)
    - dialect
    - register
  - fuse entries that are semantically very similar, yet they differ syntactically

## DiMLex-Ger: Substitution

anstatt

anstatt dass

anstelle dessen

sondern

statt

stattdessen

vielmehr

# DiMLex-Ger: Contrast

aber	doch	unterdessen
alldieweil	einerseits	wiederum
allein	freilich	wogegen
allerdings	gleichzeitig	wohingegen
andererseits	hingegen	während
bloß dass	immerhin	währenddessen
dafür	indessen	zugleich
dafür dass	jedoch	zum anderen
dagegen	nur	zum einen
dahingegen	nur dass	
dementgegen	sosehr	
demgegenüber	und	

# DiMLex-Ger: Concession

aber	entgegen	obzwar
abgesehen davon	freilich	sosehr
allerdings	gleichwohl	trotz
als ob	immerhin	trotzdem
angesichts	jedoch	unbeschadet dessen
bloß	mangels	ungeachtet dessen
bloß dass	mithin	wenn .. auch
dabei	nichtsdestominder	wenngleich
davon abgesehen, dass	nichtsdestoweniger	wiederum
dennoch	obgleich	wobei
dessen ungeachtet	obschon	wohingegen
doch	obwohl	zwar

# DiMLex-Ger: subset

## Substitution

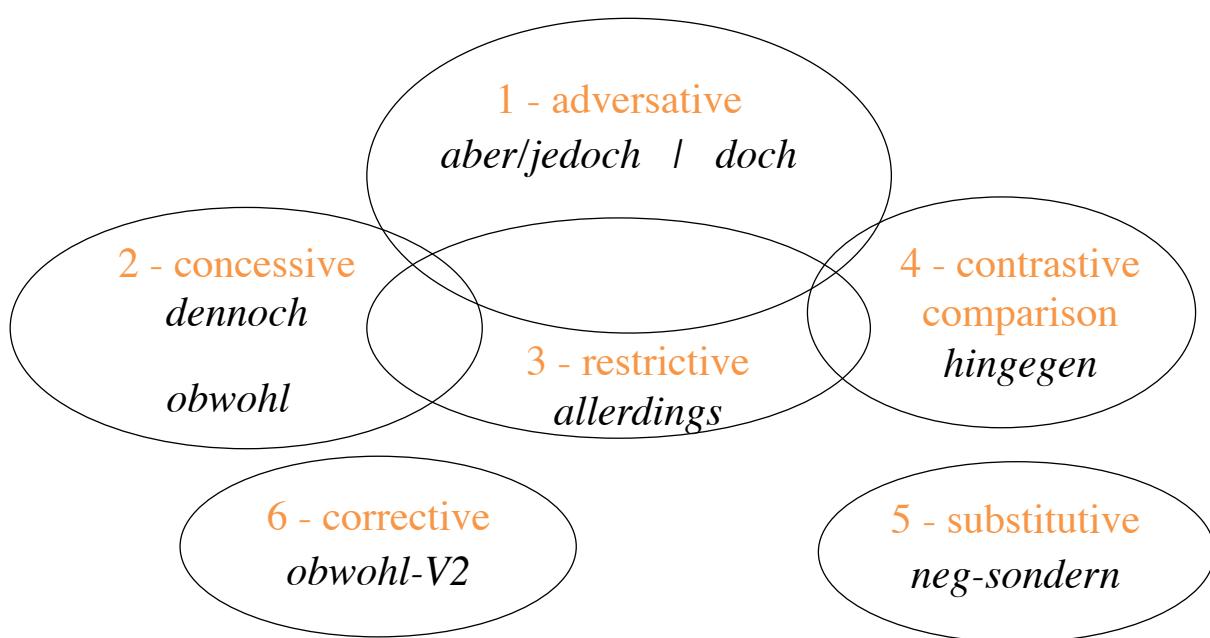
statt  
sondern  
vielmehr

## Contrast

aber  
während  
allein  
allerdings  
zum einen/anderen  
nur  
dafür  
hingegen  
zugleich  
freilich  
immerhin  
indessen  
sosehr  
und  
wiederum

## Concession

aber  
abgesehen davon  
allerdings  
als ob  
bloß  
dabei  
dennoch  
freilich  
immerhin  
mithin  
obwohl  
sosehr  
wenn .. auch  
wiederum  
wobei  
wohingegen  
zwar



# DimLex-Ger: subset

## Substitution

statt - 5  
sondern - 5  
vielmehr - 5

1 - adversative

2 - concessive

3 - restrictive

4 - contr. comparison

5 - substitutive

6 - corrective

(n): only in specific  
contexts: evaluative, ..

## Contrast

aber - 1, 2  
während - 4  
allein - 3  
allerdings - 2, 3  
zum einen/anderen - 4  
nur - 3  
dafür - (5)

hingegen - 4  
zugleich - (3, 4)

freilich - 1, 3

immerhin - (1)

indessen - 4

sosehr - (2, 4)

und - 4

wiederum - (4)

## Concession

aber - 1, 2  
abgesehen davon - (2)  
allerdings - 2, 3  
als ob - (2)  
bloß - 2, 3  
dabei - 2  
dennoch - 2  
freilich  
immerhin - (2)  
mithin  
obwohl - 2, 6  
sosehr - (2, 4)  
wenn .. auch - (2)  
wiederum  
wobei - 2, 3, 6  
zwar - 2

(using the 30 sample instances compiled by Scheffler/Stede 16)

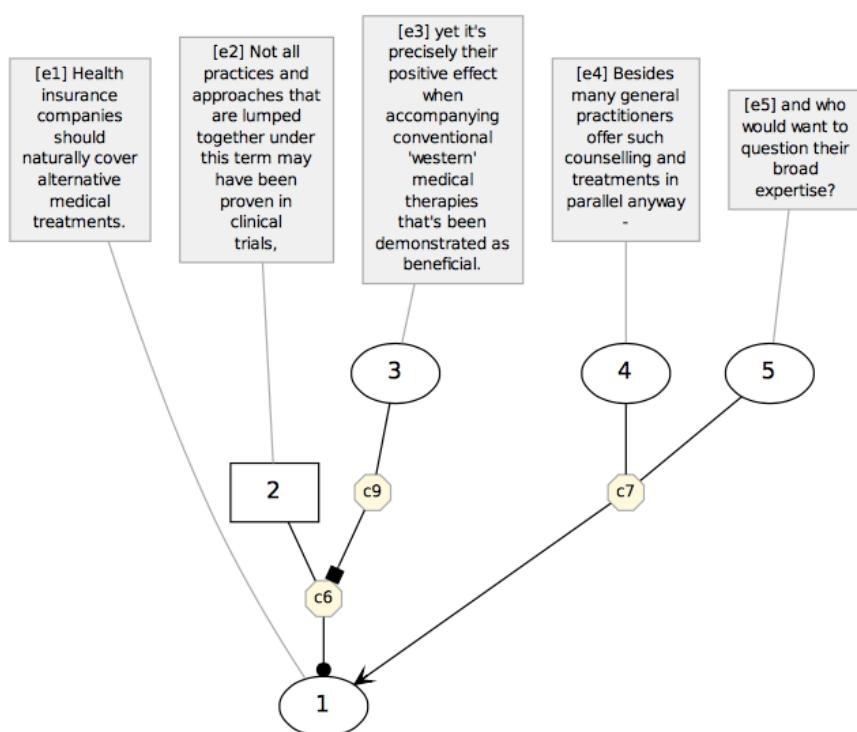
# connective-lex

- Features in the public version
  - syntactic category  
cco, csu, adv, prep, other
  - PDTB relation(s)
  - argument order
  - orthographic variants
  - non-connective readings
- Potential new features
  - more fine-grained syntax
  - modifiability
  - argument linking  
(semantic roles: causing / caused, etc.)
  - sem/prag function
    - adversative
    - concessive
    - restrictive
    - contr. comparison
    - substitutive
    - corrective
  - register, dialect

# Overview

- Motivation: connective-lex.info
- Flashback: Some contrastive German connectives
- Towards a map
- Enhancing the connective lexicon
- Application: Argumentation mining
- Conclusions

## Application: Argumentation Mining



# Application: Argumentation Mining

- The „argumentative microtext“ corpus (Peldszus/Stede 15) has parallel annotations: discourse structure (RST, SDRT) and argumentation structure
- We measured the potential benefit of discourse structure information for argumentation structure parsing (Peldszus/Stede 16, Hewett et al. 19)
- Error analyses indicate that the coarse granularity of additive, causal, contrastive RST relations limits the utility for understanding argumentation

Should the bachelor thesis be more important for the overall BSc grade?

The idea of the bachelor degree is to put together a cumulative grade from various individual assignments in order to generate a final grade. This is also more akin to the real working world later, where performance isn't singular, but constant and has to be delivered over longer periods. That's why it makes sense that the bachelor thesis only makes up a relatively small portion of the final grade (15 out of 120 credits). However, the effort required for the individual equally weighted modules differs strongly. It would therefore be wise to adjust the required effort for final theses accordingly.

Sollte die Abschlussarbeit in der BSc Gesamtnote höher gewichtet werden?

Die Idee des Bachelorstudiums ist, die Gesamtnote aus zahlreichen Einzelleistungen zusammenzusetzen, um die Endnote zu generieren. Dies entspricht auch eher der späteren Arbeitswelt, in der Leistung nicht singulär, sondern regelmäßig und über längere Zeit hinweg erbracht werden soll. Daher macht es Sinn, dass die Bachelorarbeit nur einen relativ kleinen Teil der Endnote ausmacht (15 von 120 SP). Allerdings differiert der unterschiedliche Aufwand für die einzelnen, teils gleichwertigen Module stark. Es wäre also sinnvoll, den Aufwand der Abschlussarbeit(en) entsprechend anzupassen.

# Summary

- **connective-lex.info**: from breadth to depth
- There are “**discourse marker connectives**” that should be covered by a connective lexicon ( $obwohl_{V2}$ ,  $wobei_{V2}$ ,  $weil_{V2}$ )
- For CONTRAST: An inventory of **six bottom-up induced functions** can enrich the lexicon, providing more fine-grained categories than the coherence relations

# Outlook

- **Intersubjectivity**: Do annotators agree on assigning the 6 functions?
- Can **argumentation mining** benefit from such finer-grained functions/relations?
- Representation of **cross-lingual mappings** in connective-lex
- More granularity for other **semantic groups** of connectives?

- Thank you!

- aber only:

Vierzehn Tage vorher, als die Lok 4029t am Zughaken hatte, lag das Übergewicht noch deutlich bei den Pessimisten. Doch Leistung überzeugt.

# obwohl<sub>V2</sub> into the lexicon?

- obwohl<sub>V2</sub> (and wobei<sub>V2</sub>), though limited to spontaneous, conceptually-oral language, belongs into connective-lex, as it clearly denotes a two-place relation
- (Contrary to Danlos et al. 2018, there is overlap between discourse markers and connectives)

## Granularity: CONTRAST

<ul style="list-style-type: none"><li>• <b>PDTB</b>: Substitution, Contrast, Concession</li><li>• <b>RST</b>: Contrast (multinuclear), Antithesis, Concession</li><li>• <b>SDRT/Lexconn</b>: Contrast (coord), Concession (subord), Opposition</li><li>• <b>SDRT/STAC</b>: Contrast, Correction</li></ul>	<ul style="list-style-type: none"><li>• adversative</li><li>• concessive</li><li>• restrictive</li><li>• contr. comparison</li><li>• substitutive</li><li>• corrective</li></ul> <ul style="list-style-type: none"><li>• 31 connectives</li><li>• 75 connectives</li></ul>
---	--

## TODO - DiMLex-Eng: Substitution

but

instead (of)

rather (than)

much as

quite the contrary

## TODO - DiMLex-Eng: Contrast

although

as

earlier

though

nonetheless

neither .. nor

but

meanwhile

conversely

nevertheless

yet

on the contrary

if

rather

on the one/other hand

previously

in fact

instead of

and

besides

by the way

then

nor

at the same time

while

whereas

still

in/by comparison

instead

in/by contrast

on the other hand

in the end

## TODO - DiMLex-Eng: Concession

although	as	in spite of
though	nonetheless	
but	meanwhile	
nevertheless	yet	
if	as if	
when	in the end	
and	regardless	
then	much as	
while	even so	
however	despite	
after	even though	
still	even if	