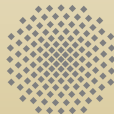


Logical metonymies and qualia structures: an annotated database of logical metonymies for German

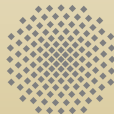
Alessandra Zarcone and Stefan Rüd
Institut für Maschinelle Sprachverarbeitung, Stuttgart
{ zarconaa@ims.uni-stuttgart.de,
ruedsn@studenten.ims.uni-stuttgart.de }

LREC 2012
Istanbul, 21-27 May 2012



Outline

- 1 Logical metonymies and qualia structures**
 - Logical Metonymies and Long Forms
 - Previous work on English
- 2 An annotated database of logical metonymies**
 - Extraction
 - Annotation
 - Evaluation
- 3 Conclusions**



Logical metonymies and qualia structures

Covert events

Logical metonymy:

begin the newspaper → begin **reading** the newspaper
enjoy the beer → enjoy **drinking** the beer



Covert Events (CE)

- events that are not realized on the surface, but are understood
- influence reading times, and are available for inference
- Qualia-determined range [Pustejovsky, 1995]:
CEs are retrieved from the *qualia* structure of the noun's lexical entry
→ range: 1-2 CEs, e.g. *newspaper* → writing or reading

Logical metonymies and qualia structures

Covert events

Logical metonymy:

begin the newspaper → begin **reading** the newspaper
enjoy the beer → enjoy **drinking** the beer



Covert Events (CE)

- events that are not realized on the surface, but are understood
- influence reading times, and are available for inference
- **Qualia-determined range [Pustejovsky, 1995]:**
CEs are retrieved from the *qualia structure* of the noun's lexical entry
→ range: 1-2 CEs, e.g. *newspaper* → writing or reading

Logical metonymies and qualia structures

Covert events

Logical metonymy:

begin the newspaper → begin **reading** the newspaper
enjoy the beer → enjoy **drinking** the beer



Covert Events (CE)

- events that are not realized on the surface, but are understood
- influence reading times, and are available for inference
- **Qualia-determined range [Pustejovsky, 1995]:**
CEs are retrieved from the *qualia structure* of the noun's lexical entry
→ range: 1-2 CEs, e.g. *newspaper* → writing or reading

Logical metonymies and qualia structures

Covert events

Logical metonymy:

begin the newspaper → begin **reading** the newspaper
enjoy the beer → enjoy **drinking** the beer



Covert Events (CE)

- events that are not realized on the surface, but are understood
- influence reading times, and are available for inference
- **Qualia-determined range [Pustejovsky, 1995]:**
CEs are retrieved from the *qualia structure* of the noun's lexical entry
→ range: 1-2 CEs, e.g. *newspaper* → writing or reading

Logical metonymies and qualia structures

Covert events

Logical metonymy:

begin the newspaper → begin **reading** the newspaper
enjoy the beer → enjoy **drinking** the beer



Covert Events (CE)

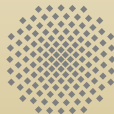
- events that are not realized on the surface, but are understood
- influence reading times, and are available for inference
- **Qualia-determined range [Pustejovsky, 1995]:**
CEs are retrieved from the *qualia structure* of the noun's lexical entry
→ range: 1-2 CEs, e.g. *newspaper* → writing or reading

Logical metonymies and qualia structures

Qualia structures

Qualia Structures (QS)

- **Agentive quale (AQ):** the event which brings about the object
 - **Telic quale (TQ):** the purpose of the object
-
- *The author began the book*
→ **writing the book (AQ)**
 - *The student began the book*
→ **reading the book (TQ)**

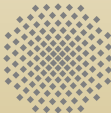


Logical metonymies and qualia structures

Qualia structures

Qualia Structures (QS)

- **Agentive quale (AQ)**: the event which brings about the object
 - **Telic quale (TQ)**: the purpose of the object
-
- *The author began the book*
→ **writing** *the book* (AQ)
 - *The student began the book*
→ **reading** *the book* (TQ)



Logical metonymies and qualia structures

Qualia roles as clusters of meanings

Buch (101241): 5006 lesen, 3468 schreiben, 1561 geben, 1092 kaufen, 1018 veröffentlichen, 893 empfehlen, 619 machen, 581 finden, 566 nehmen, 464 legen, 435 vorstellen, 385 kennen, 370 lassen, 357 bestellen, 350 finden, 331 verfassen, [...] , 162 verschlingen, 161 aufschlagen, 160 nennen, 154 durchlesen...

- sets of interpretations:

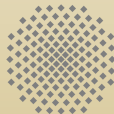
"in most cases not one verb, but a family of verbs is needed"
[Vendler, 1968, Lapata and Lascarides, 2003]

- *finish the book* → *writing, editing, publishing..*

- qualia as clusters of meanings → concepts, not single predicates

AQ: schreiben, veröffentlichen, kaufen, geben, ...

TQ: lesen, verschlingen, durchlesen, ...



Logical metonymies and qualia structures

Qualia roles as clusters of meanings

Buch (101241): 5006 lesen, 3468 schreiben, 1561 geben, 1092 kaufen, 1018 veröffentlichen, 893 empfehlen, 619 machen, 581 finden, 566 nehmen, 464 legen, 435 vorstellen, 385 kennen, 370 lassen, 357 bestellen, 350 finden, 331 verfassen, [...] , 162 verschlingen, 161 aufschlagen, 160 nennen, 154 durchlesen...

- **sets of interpretations:**

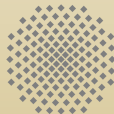
"in most cases not one verb, but **a family of verbs** is needed"
[Vendler, 1968, Lapata and Lascarides, 2003]

- *finish the book* → *writing, editing, publishing..*

- qualia as **clusters of meanings** → concepts, not single predicates

AQ: *schreiben, schreiben, schreiben, schreiben, schreiben, ...*

TQ: *lesen, verschlingen, durchlesen, ...*



Logical metonymies and qualia structures

Qualia roles as clusters of meanings

Buch (101241): 5006 **lesen**, 3468 **schreiben**, 1561 geben, 1092 kaufen, 1018 **veröffentlichen**, 893 empfehlen, 619 **machen**, 581 finden, 566 nehmen, 464 legen, 435 vorstellen, 385 kennen, 370 lassen, 357 bestellen, 350 finden, 331 **verfassen**, [...] , 162 **verschlingen**, 161 aufschlagen, 160 nennen, 154 **durchlesen**...

- **sets of interpretations:**

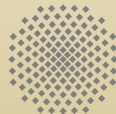
"in most cases not one verb, but **a family of verbs** is needed"
[Vendler, 1968, Lapata and Lascarides, 2003]

- *finish the book* → *writing, editing, publishing*..

- qualia as **clusters of meanings** → concepts, not single predicates

AQ: **schreiben, veröffentlichen, machen, verfassen**,...

TQ: **lesen, verschlingen, durchlesen**,...



Logical metonymies and qualia structures

Type-clash and qualia structures

- **CE undergeneration**

- *My goat eats anything. He really enjoyed your book*
→ **eating** *the book* (AQ? TQ?)
- *The publisher began a series of books*
→ **publishing** *a series of books* (AQ?)

Question #1

To what extent can covert events in logical metonymies be accounted for by qualia structure?

Logical metonymies and qualia structures

Type-clash and qualia structures

- CE undergeneration

- *My goat eats anything. He really enjoyed your book*
→ **eating** *the book* (AQ? TQ?)
- *The publisher began a series of books*
→ **publishing** *a series of books* (AQ?)

Question #1

To what extent can covert events in logical metonymies be accounted for by qualia structure?

Logical metonymies and qualia structures

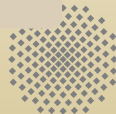
Type-clash and qualia structures

- CE undergeneration

- *My goat eats anything. He really enjoyed your book*
→ **eating** *the book* (AQ? TQ?)
- *The publisher began a series of books*
→ **publishing** *a series of books* (AQ?)

Question #1

To what extent can covert events in logical metonymies be accounted for by qualia structure?



Logical metonymies and qualia structures

Logical Metonymies and Long Forms

Logical Metonymies

Raucher können mit bestem Blick über die Stadt ihre Zigaretten geniessen

"Smokers can enjoy their cigarettes with the best view of the city"

- ① → **smoke** ("rauchen") *Zigaretten geniessen*
- ② *cigarettes, smoke* ("Zigaretten, rauchen") → **TQ**

Long Forms

In dieser Zeit begann er, seine berühmten großformatigen Aquarelle zu malen

"In this period he began **to paint** his famous largeformat watercolors"

- ① *watercolors, paint* ("Aquarelle, malen") → **AQ**

Logical metonymies and qualia structures

Logical Metonymies and Long Forms

Logical Metonymies

Raucher können mit bestem Blick über die Stadt ihre Zigaretten geniessen

“Smokers can enjoy their cigarettes with the best view of the city”

- ① → **smoke** (“rauchen”) *Zigaretten geniessen*
- ② *cigarettes, smoke* (“Zigaretten, rauchen”) → **TQ**

Long Forms

In dieser Zeit begann er, seine berühmten großformatigen Aquarelle zu malen

“In this period he began **to paint** his famous largeformat watercolors”

- ① *watercolors, paint* (“Aquarelle, malen”) → **AQ**

Logical metonymies and qualia structures

Logical Metonymies and Long Forms

Logical Metonymies

Raucher können mit bestem Blick über die Stadt ihre Zigaretten geniessen

“Smokers can enjoy their cigarettes with the best view of the city”

- ① → **smoke** (“rauchen”) *Zigaretten geniessen*
- ② *cigarettes, smoke* (“Zigaretten, rauchen”) → **TQ**

Long Forms

In dieser Zeit begann er, seine berühmten großformatigen Aquarelle zu malen

“In this period he began **to paint** his famous largeformat watercolors”

- ① *watercolors, paint* (“Aquarelle, malen”) → **AQ**

Logical metonymies and qualia structures

Logical Metonymies and Long Forms

Question #1

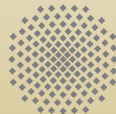
To what extent can **covert events** in **logical metonymies** be accounted for by qualia structure?

Question #2

To what extent can **events** in **long forms** be accounted for by qualia structure?



A corpus study of German verbs



Logical metonymies and qualia structures

Logical Metonymies and Long Forms

Question #1

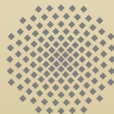
To what extent can **covert events** in **logical metonymies** be accounted for by qualia structure?

Question #2

To what extent can **events** in **long forms** be accounted for by qualia structure?



A corpus study of German verbs



Logical metonymies and qualia structures

Logical Metonymies and Long Forms

Question #1

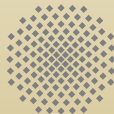
To what extent can **covert events** in **logical metonymies** be accounted for by qualia structure?

Question #2

To what extent can **events** in **long forms** be accounted for by qualia structure?



A corpus study of German verbs



Logical metonymies and qualia structures

Previous work on English

Similar studies for English [Briscoe et al., 1990] and [Verspoor, 1997]:

Briscoe et al. : LOB corpus (1 million words)

enjoy, prefer, finish, start, begin, miss and regret

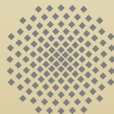
on average 17% CEs not solvable with qualia

Vespoor : BNC corpus (100 million words)

begin, finish, begin on

begin and *finish*: 5% CEs not solvable with qualia

begin on: 64% CEs not solvable with qualia



Logical metonymies and qualia structures

Previous work on English

Similar studies for English [Briscoe et al., 1990] and [Verspoor, 1997]:

Briscoe et al. : LOB corpus (1 million words)

enjoy, prefer, finish, start, begin, miss and regret

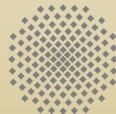
on average 17% CEs not solvable with qualia

Vespoor : BNC corpus (100 million words)

begin, finish, begin on

begin and *finish*: 5% CEs not solvable with qualia

begin on: 64% CEs not solvable with qualia



Logical metonymies and qualia structures

Previous work on English

Similar studies for English [Briscoe et al., 1990] and [Verspoor, 1997]:

Briscoe et al. : LOB corpus (1 million words)

enjoy, prefer, finish, start, begin, miss and regret

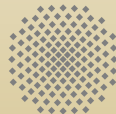
on average 17% CEs not solvable with qualia

Vespoor : BNC corpus (100 million words)

begin, finish, begin on

begin and *finish*: 5% CEs not solvable with qualia

begin on: 64% CEs not solvable with qualia

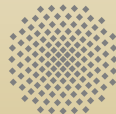


An annotated database of logical metonymies

A corpus study of German verbs:

anfangen (mit), “start (with)”, *aufhören (mit)*, “stop (with)”, *beenden*
“finish”, *beginnen (mit)*, “begin (with)”, *geniessen*, “enjoy”

- dependency parsed version of the deWaC corpus
(1.7 billion words, [Baroni et al., 2008])
- two annotators, native speakers of German
- only sentences with artifact obj.

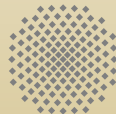


An annotated database of logical metonymies

A corpus study of German verbs:

anfangen (mit), “start (with)”, *aufhören (mit)*, “stop (with)”, *beenden* “finish”, *beginnen (mit)*, “begin (with)”, *geniessen*, “enjoy”

- dependency parsed version of the deWaC corpus (1.7 billion words, [Baroni et al., 2008])
- two annotators, native speakers of German
- only sentences with artifact obj.

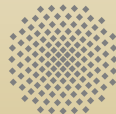


An annotated database of logical metonymies

A corpus study of German verbs:

anfangen (mit), “start (with)”, *aufhören (mit)*, “stop (with)”, *beenden* “finish”, *beginnen (mit)*, “begin (with)”, *geniessen*, “enjoy”

- dependency parsed version of the deWaC corpus (1.7 billion words, [Baroni et al., 2008])
- two annotators, native speakers of German
- only sentences with artifact obj.



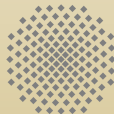
An annotated database of logical metonymies

Logical metonymies

Question #1

To what extent can **covert events** in **logical metonymies** be accounted for by qualia structure?

- *Raucher können mit bestem Blick über die Stadt ihre Zigaretten geniessen*
"Smokers can enjoy their cigarettes with the best view of the city"
- CE annotation
Zigaretten geniessen → **rauchen**
"enjoy cigarettes" → **smoking**
- Annotation of qualia overlap
Zigaretten, smoke ("cigarettes, smoking") → **TELIC QUALE**
- 2661 metonymies annotated



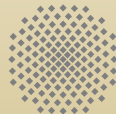
An annotated database of logical metonymies

Logical metonymies

Question #1

To what extent can **covert events** in **logical metonymies** be accounted for by qualia structure?

- *Raucher können mit bestem Blick über die Stadt ihre Zigaretten geniessen*
“Smokers can enjoy their cigarettes with the best view of the city”
- CE annotation
Zigaretten geniessen → **rauchen**
“enjoy cigarettes” → **smoking**
- Annotation of qualia overlap
Zigaretten, smoke (“cigarettes, smoking”) → **TELIC QUALE**
- 2661 metonymies annotated



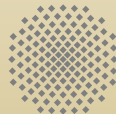
An annotated database of logical metonymies

Logical metonymies

Question #1

To what extent can **covert events** in **logical metonymies** be accounted for by qualia structure?

- *Raucher können mit bestem Blick über die Stadt ihre Zigaretten geniessen*
“Smokers can enjoy their cigarettes with the best view of the city”
- CE annotation
Zigaretten geniessen → **rauchen**
“enjoy cigarettes” → **smoking**
- Annotation of qualia overlap
Zigaretten, smoke (“cigarettes, smoking”) → **TELIC QUALE**
- 2661 metonymies annotated



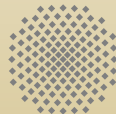
An annotated database of logical metonymies

Logical metonymies

Question #1

To what extent can **covert events** in **logical metonymies** be accounted for by qualia structure?

- *Raucher können mit bestem Blick über die Stadt ihre Zigaretten geniessen*
“Smokers can enjoy their cigarettes with the best view of the city”
- CE annotation
Zigaretten geniessen → **rauchen**
“enjoy cigarettes” → **smoking**
- Annotation of qualia overlap
Zigaretten, smoke (“cigarettes, smoking”) → **TELIC QUALE**
- 2661 metonymies annotated



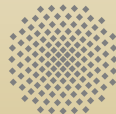
An annotated database of logical metonymies

Logical metonymies

Question #1

To what extent can **covert events** in **logical metonymies** be accounted for by qualia structure?

- *Raucher können mit bestem Blick über die Stadt ihre Zigaretten geniessen*
“Smokers can enjoy their cigarettes with the best view of the city”
- CE annotation
Zigaretten geniessen → **rauchen**
“enjoy cigarettes” → **smoking**
- Annotation of qualia overlap
Zigaretten, smoke (“cigarettes, smoking”) → **TELIC QUALE**
- 2661 metonymies annotated



Covert events in the corpus

Long forms

Question #2

To what extent can **events** in **long forms** be accounted for by qualia structure?

- *In dieser Zeit begann er, seine berühmten großformatigen Aquarelle zu malen*
“In this period he began **to paint** his famous largeformat watercolors”
- Annotation of qualia overlap
watercolors, paint (“Aquarelle, malen”) → **AGENTIVE QUALE**
- 1886 long forms annotated



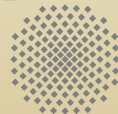
Covert events in the corpus

Long forms

Question #2

To what extent can **events** in **long forms** be accounted for by qualia structure?

- *In dieser Zeit begann er, seine berühmten großformatigen Aquarelle zu malen*
“In this period he began **to paint** his famous largeformat watercolors”
- Annotation of qualia overlap
watercolors, paint (“Aquarelle, malen”) → **AGENTIVE QUALE**
- 1886 long forms annotated



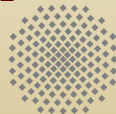
Covert events in the corpus

Long forms

Question #2

To what extent can **events** in **long forms** be accounted for by qualia structure?

- *In dieser Zeit begann er, seine berühmten großformatigen Aquarelle zu malen*
“In this period he began **to paint** his famous largeformat watercolors”
- Annotation of qualia overlap
watercolors, paint (“Aquarelle, malen”) → **AGENTIVE QUALE**
- 1886 long forms annotated



Covert events in the corpus

Long forms

Question #2

To what extent can **events** in **long forms** be accounted for by qualia structure?

- *In dieser Zeit begann er, seine berühmten großformatigen Aquarelle zu malen*
“In this period he began **to paint** his famous largeformat watercolors”
- Annotation of qualia overlap
watercolors, paint (“Aquarelle, malen”) → **AGENTIVE QUALE**
- 1886 long forms annotated



An annotated database of logical metonymies

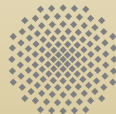
Method

Method:

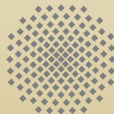
	Logical metonymies	Long forms
1	Extraction	
2	CE annotation	—
3	Annotation of qualia overlap	
Tagset	AQ TQ OTHER Q UNDET CTXT INSUFF	

database available at:

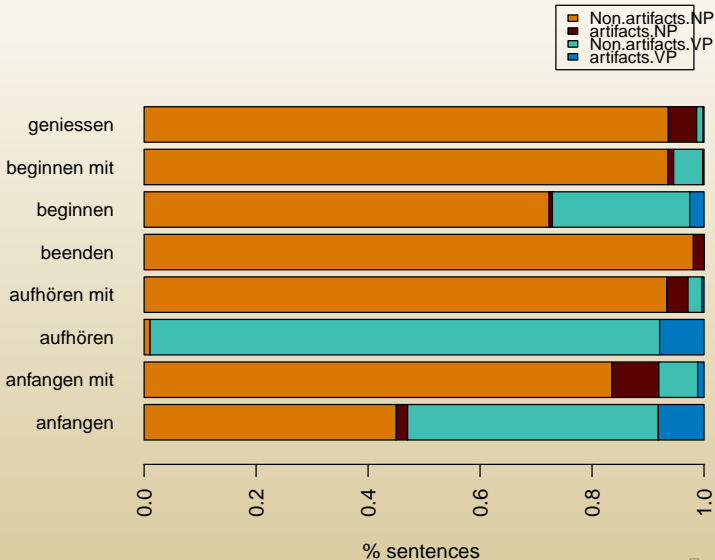
<http://www.ims.uni-stuttgart.de/~zarconaa/data.html.en>



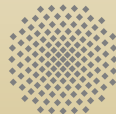
Extraction



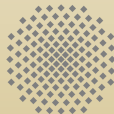
Extraction



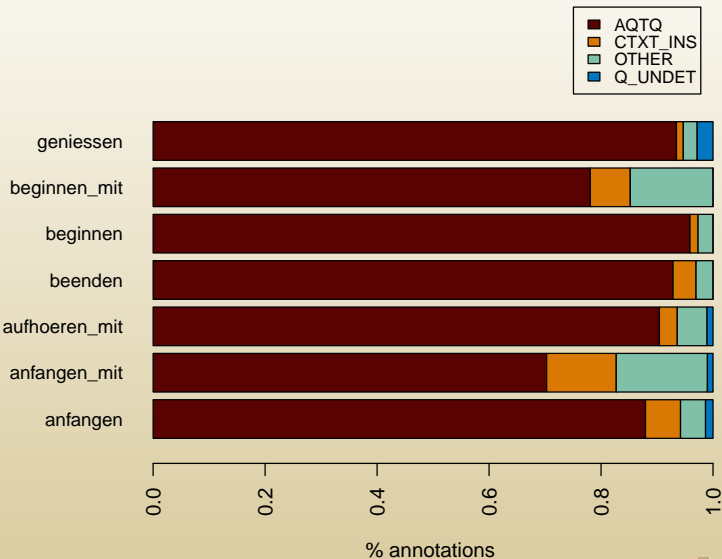
Only
logical metonymies
and
long forms
annotated



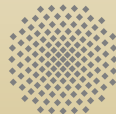
Annotation of qualia overlap



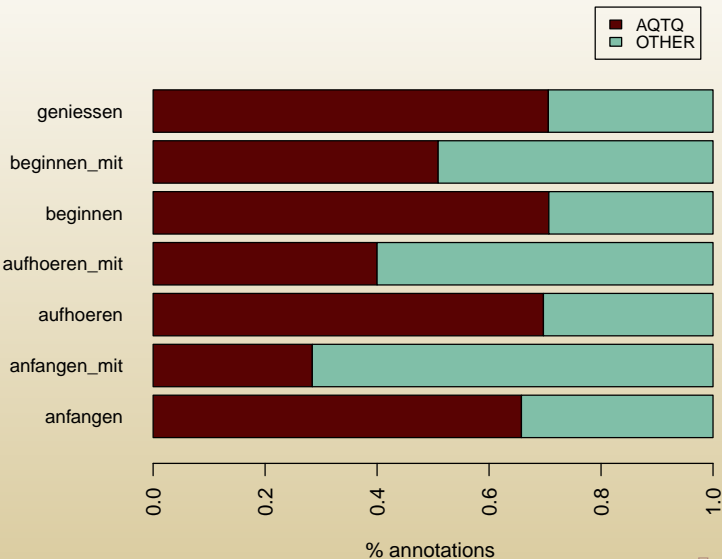
Annotation



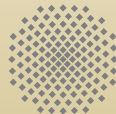
Only
logical metonymies
and
long forms
annotated



Annotation



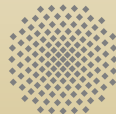
Only
logical metonymies
and
long forms
annotated



Examples

Verb tendencies: AQ vs. TQ.:

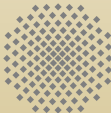
- *anfangen* and *beginnen*: AQ preference (54%-88%)
- *anfangen mit* and *beginnen mit*: TQ preference (31%-63%)
- *Geniessen*, logical metonymy: strong TQ preference (90%-97%)
- *Geniessen*, long form:
Ich habe es wirklich genossen, diesen Film zu drehen.
I really enjoyed making this film.



Examples

Verb tendencies: AQ vs. TQ.:

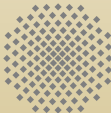
- *anfangen* and *beginnen*: AQ preference (54%-88%)
- *anfangen mit* and *beginnen mit*: TQ preference (31%-63%)
- *Geniessen*, logical metonymy: strong TQ preference (90%-97%)
- *Geniessen*, long form:
Ich habe es wirklich genossen, diesen Film zu drehen.
I really enjoyed making this film.



Examples

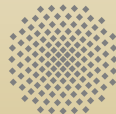
Verb tendencies: AQ vs. TQ.:

- *anfangen* and *beginnen*: AQ preference (54%-88%)
- *anfangen mit* and *beginnen mit*: TQ preference (31%-63%)
- *Geniessen*, logical metonymy: strong TQ preference (90%-97%)
- *Geniessen*, long form:
Ich habe es wirklich genossen, diesen Film zu drehen.
I really enjoyed making this film.



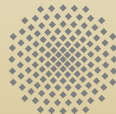
Examples

- Logical metonymies vs. long forms
(*anfangen, aufhören mit, beenden, beginnen, geniessen*)
 - Logical metonymies: majority of QS interpretation (AQ or TQ)
 - Long forms: higher percentages of non-QS interpretations
- Context-based interpretation:
 - *Wir haben mit einem traditionellem Brett angefangen und es lief recht gut.*
We started with a traditional board and it went quite well.



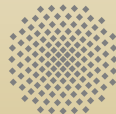
Examples

- Logical metonymies vs. long forms
(*anfangen, aufhören mit, beenden, beginnen, geniessen*)
 - Logical metonymies: majority of QS interpretation (AQ or TQ)
 - Long forms: higher percentages of non-QS interpretations
- Context-based interpretation:
 - *Wir haben mit einem traditionellem Brett angefangen und es lief recht gut.*
We started with a traditional board and it went quite well.



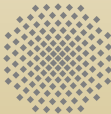
Examples

- Logical metonymies vs. long forms
(*anfangen, aufhören mit, beenden, beginnen, geniessen*)
 - Logical metonymies: majority of QS interpretation (AQ or TQ)
 - Long forms: higher percentages of non-QS interpretations
- Context-based interpretation:
 - *Wir haben mit einem traditionellem Brett angefangen und es lief recht gut.*
We started with a traditional board and it went quite well.

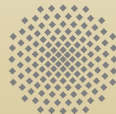


Examples

- Logical metonymies vs. long forms
(*anfangen, aufhören mit, beenden, beginnen, geniessen*)
 - Logical metonymies: majority of QS interpretation (AQ or TQ)
 - Long forms: higher percentages of non-QS interpretations
- Context-based interpretation:
 - *Wir haben mit einem traditionellem Brett angefangen und es lief recht gut*
We started with a traditional board and it went quite well.



Evaluation



Evaluation

- two expert annotators, whole dataset
Krippendorff's α for the annotated database was .71 for the long forms and .6 for the metonymies (substantial agreement)
- crowdsourcing inter-annotator evaluation on AMT
24 unique annotators, random sample
(100 logical metonymies, 100 long forms)
 - long forms:
does the underlined event depict the purpose of the object, or does it depict the coming about of the object?
 - logical metonymies:
which implicit event is involved? does it depict the purpose of the object, or the coming about of the object?



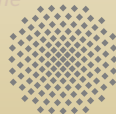
Evaluation

- two expert annotators, whole dataset
Krippendorff's α for the annotated database was .71 for the long forms and .6 for the metonymies (substantial agreement)
- crowdsourcing inter-annotator evaluation on AMT
24 unique annotators, random sample
(100 logical metonymies, 100 long forms)
 - long forms:
does the underlined event depict the purpose of the object, or does it depict the coming about of the object?
 - logical metonymies:
which implicit event is involved? does it depict the purpose of the object, or the coming about of the object?



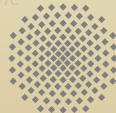
Evaluation

- two expert annotators, whole dataset
Krippendorff's α for the annotated database was .71 for the long forms and .6 for the metonymies (substantial agreement)
- crowdsourcing inter-annotator evaluation on AMT
24 unique annotators, random sample
(100 logical metonymies, 100 long forms)
 - long forms:
does the undelined event depict the purpose of the object, or does it depict the coming about of the object?
 - logical metonymies:
which implicit event is involved? does it depict the purpose of the object, or the coming about of the object?



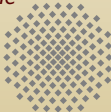
Evaluation

- two expert annotators, whole dataset
Krippendorff's α for the annotated database was .71 for the long forms and .6 for the metonymies (substantial agreement)
- crowdsourcing inter-annotator evaluation on AMT
24 unique annotators, random sample
(100 logical metonymies, 100 long forms)
 - long forms:
does the undelined event depict the purpose of the object, or does it depict the coming about of the object?
 - logical metonymies:
which implicit event is involved? does it depict the purpose of the object, or the coming about of the object?



Evaluation

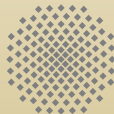
- two expert annotators, whole dataset
Krippendorff's α for the annotated database was .71 for the long forms and .6 for the metonymies (substantial agreement)
- crowdsourcing inter-annotator evaluation on AMT
24 unique annotators, random sample
(100 logical metonymies, 100 long forms)
 - long forms:
does the undelined event depict the purpose of the object, or does it depict the coming about of the object?
 - logical metonymies:
which implicit event is involved? does it depict the purpose of the object, or the coming about of the object?



Evaluation

	Logical Metonymies	Long Forms
(sample) ANN 1 - ANN 2	0.57	0.65
AMT	0.5	0.42
ANN1 - ANN2 - AMT	0.59	0.5

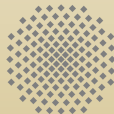
- Long forms have a higher number of non-QS events
- non-expert annotators seem to have bigger difficulties in classifying these, and provide fewer OTHER responses for long forms



Evaluation

	Logical Metonymies	Long Forms
(sample) ANN 1 - ANN 2	0.57	0.65
AMT	0.5	0.42
ANN1 - ANN2 - AMT	0.59	0.5

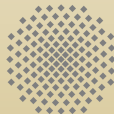
- Long forms have a higher number of non-QS events
- non-expert annotators seem to have bigger difficulties in classifying these, and provide fewer OTHER responses for long forms



Evaluation

	Logical Metonymies	Long Forms
(sample) ANN 1 - ANN 2	0.57	0.65
AMT	0.5	0.42
ANN1 - ANN2 - AMT	0.59	0.5

- Long forms have a higher number of non-QS events
- non-expert annotators seem to have bigger difficulties in classifying these, and provide fewer OTHER responses for long forms



Conclusions

Question #1

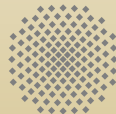
To what extent can **covert events** in **logical metonymies** be accounted for by qualia structure?

Question #2

To what extent can **events** in **long forms** be accounted for by qualia structure?

Briscoe et al. : LOB corpus (1 million words)
17% CEs not solvable with qualia

Vespoor : BNC corpus (100 million words)
begin and *finish*: 5% CEs not solvable with qualia
begin on: 64% CEs not solvable with qualia



Conclusions

Question #1

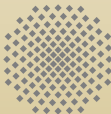
To what extent can **covert events** in **logical metonymies** be accounted for by qualia structure?

Question #2

To what extent can **events** in **long forms** be accounted for by qualia structure?

Briscoe et al. : LOB corpus (1 million words)
17% CEs not solvable with qualia

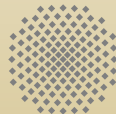
Vespoor : BNC corpus (100 million words)
begin and *finish*: 5% CEs not solvable with qualia
begin on: 64% CEs not solvable with qualia



Conclusions

our study : deWaC corpus (1.7 billion words) - German language

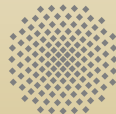
- **logical metonymies**: 81%-98% QS-interpretations for *anfangen, beenden, beginnen* and *geniessen*
- **long forms**: stronger tendency for non-QS events (conversational maxims [Grice, 1975, Lapata et al., 2003, Egg, 2004]: implicit CE → basic/default QS-interpretation; explicit event → non-QS interpretation)
- a database of 4547 annotated contexts (2661 metonymies and 1886 long forms) freely available for scientific research purposes
- database available at:
<http://www.ims.uni-stuttgart.de/~zarconaa/data.html.en>



Conclusions

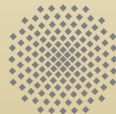
- our study** : deWaC corpus (1.7 billion words) - German language
- **logical metonymies**: 81%-98% QS-interpretations for *anfangen, beenden, beginnen* and *geniessen*
 - **long forms**: stronger tendency for non-QS events (conversational maxims [Grice, 1975, Lapata et al., 2003, Egg, 2004]: implicit CE → basic/default QS-interpretation; explicit event → non-QS interpretation)

- a database of 4547 annotated contexts (2661 metonymies and 1886 long forms) freely available for scientific research purposes
- database available at:
<http://www.ims.uni-stuttgart.de/~zarconaa/data.html.en>



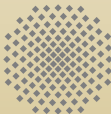
Conclusions

- our study** : deWaC corpus (1.7 billion words) - German language
- **logical metonymies**: 81%-98% QS-interpretations for *anfangen, beenden, beginnen* and *geniessen*
 - **long forms**: stronger tendency for non-QS events (conversational maxims [Grice, 1975, Lapata et al., 2003, Egg, 2004]: implicit CE → basic/default QS-interpretation; explicit event → non-QS interpretation)
- a database of 4547 annotated contexts (2661 metonymies and 1886 long forms) freely available for scientific research purposes
 - database available at:
<http://www.ims.uni-stuttgart.de/~zarconaa/data.html.en>



Acknowledgements

The research for this paper was funded by the German Research Foundation (DFG) as part of the SFB 732 - project D6 at the University of Stuttgart





Baroni, M., Bernardini, S., Ferraresi, A., and Zanchetta, E. (2008).
The WaCky Wide Web: A collection of very large linguistically processed web-crawled corpora.
Language Resources and Evaluation, 43(3):209–226.



Briscoe, T., Copestake, A., and Boguraev, B. (1990).
Enjoy the paper: Lexical semantics via lexicology.
In *Proceedings of the 13th COLING*, volume 2, pages 42–47.



Egg, M. (2004).
Metonymie als Phaenomen der Semantik-Pragmatik-Schnittstelle.
Metaphorik, 6:37–53.



Grice, H. P. (1975).
Logic and conversation.
In et al., P. C., editor, *Syntax and Semantics*, volume 3. Academic Press, New York.



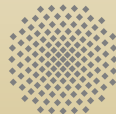
Lapata, M., Keller, F., and Scheepers, C. (2003).
Intra-sentential context effects on the interpretation of logical metonymy.
Cognitive Science, 27(4):649–668.



Lapata, M. and Lascarides, A. (2003).
A probabilistic account of logical metonymy.
Computational Linguistics, 29(2):263–317.



Pustejovsky, J. (1995).
The Generative Lexicon.
MIT Press, Cambridge, MA.





Vendler, Z. (1968).

Adjectives and Nominalizations.

Mouton, The Hague, The Netherlands.



Verspoor, C. M. (1997).

Contextually-dependent lexical semantics.

Doctoral dissertation, University of Edinburgh.

