

Normalization in context: Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

Universität Tübingen

NLP4CALL
7 November 2018

Introduction

Annotation

CREG

MTH

Difficult Cases

Inter-Annotator
Agreement

Conclusion

References



Normalization

Text normalization: converting text to a standard or canonical form for further use (search, analysis, text-to-speech, etc.)

Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

[CREG](#)

[MTH](#)

[Difficult Cases](#)

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)



Normalization

Text normalization: converting text to a standard or canonical form for further use (search, analysis, text-to-speech, etc.)

- ▶ historical corpora

wolde, woolde, wuld, wud, vwould → *would*

- ▶ user-generated content

tmrw → *tomorrow*

- ▶ L1: Litkey (German, school children)

Schuhle → *Schule*

- ▶ L2: Falko (German, university students)

Pflict → *Pflicht*

diese → *dieser*

(Archer et al. 2015; Han & Baldwin 2011; Laarmann-Quante et al. 2017; Reznicek et al. 2012)

Normalization in Context

How to normalize?

I don't know he live were.

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator
Agreement](#)

[Conclusion](#)

[References](#)



Normalization in Context

Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

How to normalize?

I don't know he live were.

And in context?

Translation task: *I don't know where he lives.*

from the Hiroshima English Learners' Corpus (Miura 1998)

⇒ Requires integration of bottom-up and top-down processing

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)



Target Hypotheses

- ▶ Minimal target hypothesis: Falko, ZH1 (Reznicek et al. 2012)
 - ▶ minimal form-based changes
⇒ grammaticality

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator
Agreement](#)

[Conclusion](#)

[References](#)



Target Hypotheses

- ▶ Minimal target hypothesis: Falko, ZH1 (Reznicek et al. 2012)
 - ▶ minimal form-based changes
⇒ grammaticality
- ▶ Meaning-based target hypothesis (Meurers 2015; Meurers & Dickinson 2017)
 - ▶ minimal changes that support meaning in context
⇒ grammaticality + meaning

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator
Agreement](#)

[Conclusion](#)

[References](#)



Normalization is Hard

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

- ▶ Normalization is hard
(e.g., Fitzpatrick & Seegmiller 2004; Lüdeling 2008; Tetreault & Chodorow 2008; Lee et al. 2009; Rosen et al. 2013; Dahlmeier et al. 2013)
 - ▶ Error annotation depends on normalization
(cf. Hirschmann et al. 2007)

Introduction

Annotation

CREG

MTH

Difficult Cases

Inter-Annotator
Agreement

Conclusion

References



Normalization is Hard

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

- ▶ Normalization is hard
(e.g., Fitzpatrick & Seegmiller 2004; Lüdeling 2008; Tetreault & Chodorow 2008; Lee et al. 2009; Rosen et al. 2013; Dahlmeier et al. 2013)
 - ▶ Error annotation depends on normalization
(cf. Hirschmann et al. 2007)
- ▶ Task context improves normalization (Lee et al. 2009, this work)

Introduction

Annotation

CREG

MTH

Difficult Cases

Inter-Annotator
Agreement

Conclusion

References



Corpus of Reading Exercises in German

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

CREG: German L2 learner corpus with explicit task contexts
(Ott et al. 2012)

- ▶ Learners: students in German classes at two American universities
- ▶ Tasks: reading comprehension exercises
 - ▶ Reading texts
 - ▶ Comprehension questions
 - ▶ Teacher-provided target answer(s)
 - ▶ Student answers
 - ▶ Teacher assessments of student answer meaning

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator
Agreement](#)

[Conclusion](#)

[References](#)



CREG Example

Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)

Question: Was sah der Mann, als er die Tür aufmachte?

'What did the man see when he opened the door?'

Student: Er sahe seiner Frau.

'He saw his wife.'

Teacher: Als er die Tür aufmachte, sah der Mann seine Frau.

'When he opened the door, the man saw his wife.'

Reading: Als er die Tür aufmachte (sie weinte dabei, die Tür), sahen ihm die blaßblauen Augen seiner Frau entgegen.

'When he opened the door (it creaked, the door), his wife's pale blue eyes awaited him.'

Meaning: Binary: *appropriate*, Detailed: *correct* (2x)



CREG-5K Meaning Assessments

Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

CREG-5K (5138 answers to 877 questions to 98 texts)

Binary	Approp.	Inapprop.
Detailed	(%)	(%)
Correct	76.9	0.0
Missing Concept	14.5	43.7
Extra Concept	6.2	3.2
Blend	2.4	50.2
Non-Answer	0.0	2.9

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)



CREG-5K Meaning Assessments

Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

CREG-5K (5138 answers to 877 questions to 98 texts)

Binary	Approp.	Inapprop.
Detailed	(%)	(%)
Correct	76.9	0.0
Missing Concept	14.5	43.7
Extra Concept	6.2	3.2
Blend	2.4	50.2
Non-Answer	0.0	2.9

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)



Meaning-Based Target Hypothesis Annotation

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

In the reading comprehension exercise context, a **meaning-based target hypothesis** (MTH) provides a grammatical version of the student answer that

- ▶ Preserves the meaning of the answer
- ▶ Respects the task context
- ▶ Makes minimal modifications

Introduction

Annotation

CREG

MTH

Difficult Cases

Inter-Annotator
Agreement

Conclusion

References



Multi-Layer Meaning-Based Target Hypotheses

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

Student: Er sahe seiner Frau.

Introduction

Annotation

CREG

MTH

Difficult Cases

Inter-Annotator
Agreement

Conclusion

References



Multi-Layer Meaning-Based Target Hypotheses

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

Introduction

Annotation

CREG

MTH

Difficult Cases

Inter-Annotator
Agreement

Conclusion

References

Student: Er sahe seiner Frau.

MTH1 (Non-Word): Er sah seiner Frau.



Multi-Layer Meaning-Based Target Hypotheses

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator
Agreement](#)

[Conclusion](#)

[References](#)

Student: Er sahe seiner Frau.

MTH1 (Non-Word): Er sah **seiner** Frau.

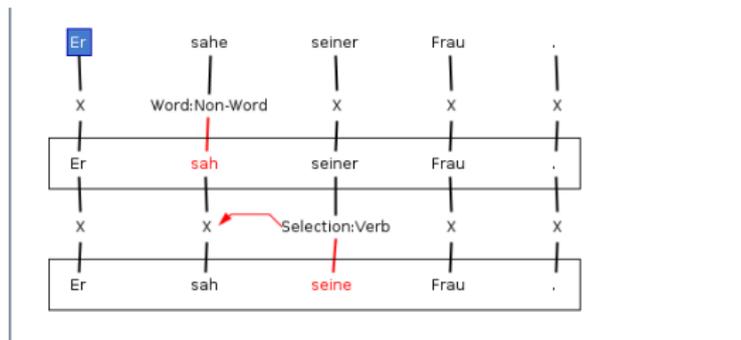


MTH2 (Full Norm.): Er sah **seine** Frau.



MTH Annotation: feat

feat: Flexible Error Annotation Tool developed for CzeSL
(Hana et al. 2012)



Er sahe seiner Frau.

Target Answer.

Als er die Tür aufmachte, sah der Mann seine Frau.

Question.

Was sah der Mann, als er die Tür aufmachte?

Reading Text.

Die drei dunklen Könige von Wolfgang Borchert Er tappte durch die dunkle Vorstadt. Die Häuser standen abgebrochen gegen den Himmel. Der Mond fehlte, und das Pflaster war erschrocken über den späten Schritt. Dann fand er eine alte Planke. Da trat er mit dem Fuß gegen bis eine Latte

Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

Introduction

Annotation

CREG

MTH

Difficult Cases

Inter-Annotator Agreement

Conclusion

References



Reading comprehension context

- ▶ Particular question types
- ▶ Lifting from reading text

General

- ▶ What are *minimal* changes?
- ▶ Interpretation of context

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator
Agreement](#)

[Conclusion](#)

[References](#)



Difficult Cases: Problematic Copied Material

Q: Was tat Herr Muschler, als seine Frau mit ihm zu sprechen versuchte?

'What was Herr Muschler doing while his wife was trying to talk to him?'

SA: Er sah **nicht nur** fern und die Zeitung.

*'He **not only** watched TV and the newspaper.'*

TA: Er sah fern, las die Zeitung, rauchte eine Zigarette und trank ein Glas Bier.

'He watched TV, read the newspaper, smoked a cigarette, and drank a glass of beer.'

RT: Herr Muschler sah **nicht nur** fern, sondern las außerdem noch die Zeitung.

*'He was **not only** watching TV but also reading the newspaper.'*

Difficult Cases: Copied + Answer Meaning

Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)

SA: Er arbeitet **mit dem BAföG**, und er hat Kindergeld
*'He works **with BAföG** and gets child benefits.'*

Q: Welche Einnahmen hat Johann Steffen? Nennen Sie zwei Dinge!
'What income sources does Johann Steffen have? Name two of them.'

RT: **Mit dem BAföG**, meiner Arbeit als wissenschaftliche Hilfe an der Uni und dem Kindergeld, das ich noch bis zum 25. Lebensjahr erhalte, komme ich auf 920 Euro im Monat.
*'**With BAföG**, my student assistant position at the university, and my child benefits, which I receive until I'm 25, my income is 920 euros per month.'*



Difficult Cases: Extra Concepts

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator
Agreement](#)

[Conclusion](#)

[References](#)

Q: Wo leben die meisten Amischen heute?

'Where do most Amish live today?'

TA: Heute leben die meisten Amischen in Ohio,
Pennsylvanien und Indiana.

*'Today most Amish live in Ohio, Pennsylvania, and
Indiana.'*

SA: Die meisten Amischen leben in Ohio, Pennsylvania,
und Indiana. **Es gibt auch ein paar in Yoder, Kansas.**

*'Most Amish live in Ohio, Pennsylvania, and Indiana.
There are also a few in Yoder, Kansas.'*



Difficulties for CREG MTHs

- ▶ Disagreements on interpretation of context:
 - ▶ No solution
- ▶ What is minimal?
 - ▶ *not only ... but also*
- ▶ Addressing answer meaning
 - ▶ *with BAFöG*

Harder for:

extra concept, missing concept, blend

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator
Agreement](#)

[Conclusion](#)

[References](#)

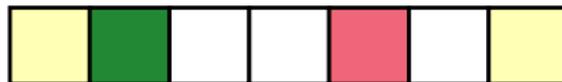


Inter-Annotator Agreement for Normalization

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

Ann1



Ann2



[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator
Agreement](#)

[Conclusion](#)

[References](#)

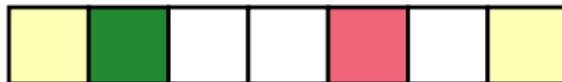


Inter-Annotator Agreement for Normalization

NUCLE (Dahlmeier et al. 2013)

- ▶ Normalization identification

Ann1



Ann2



Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)



Inter-Annotator Agreement for Normalization

NUCLE (Dahlmeier et al. 2013)

- Normalization identification

Ann1



Ann2



Ann1



Ann2



Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)

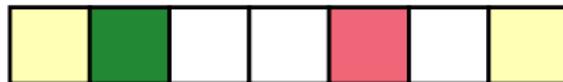


Inter-Annotator Agreement for Normalization

NUCLE (Dahlmeier et al. 2013)

- ▶ Error tag given identification

Ann1



Ann2



Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)

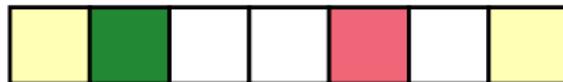


Inter-Annotator Agreement for Normalization

NUCLE (Dahlmeier et al. 2013)

- ▶ Error tag given identification

Ann1



Ann2



Ann1



Ann2



Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)

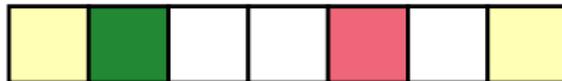


Inter-Annotator Agreement for Normalization

NUCLE (Dahlmeier et al. 2013)

- ▶ Error + normalization given identification

Ann1



Ann2



Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)

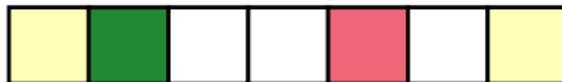


Inter-Annotator Agreement for Normalization

NUCLE (Dahlmeier et al. 2013)

- ▶ Error + normalization given identification

Ann1



Ann2



Ann1



Ann2



Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

[CREG](#)

[MTH](#)

[Difficult Cases](#)

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)

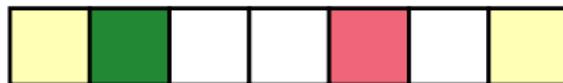


Inter-Annotator Agreement for Normalization

CzeSL (Rosen et al. 2013)

- ▶ Error tag identification

Ann1



Ann2



Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)

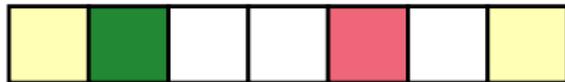


Inter-Annotator Agreement for Normalization

CzeSL (Rosen et al. 2013)

- ▶ Error tag identification

Ann1



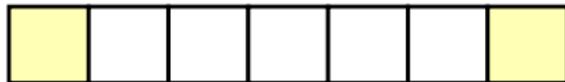
Ann2



Ann1



Ann2



Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)

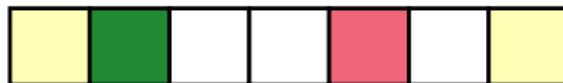


Inter-Annotator Agreement for Normalization

CzeSL (Rosen et al. 2013)

- ▶ Error tag identification

Ann1



Ann2



Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator
Agreement](#)

[Conclusion](#)

[References](#)



Inter-Annotator Agreement for Normalization

CzeSL (Rosen et al. 2013)

- ▶ Error tag identification

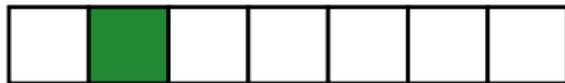
Ann1



Ann2



Ann1



Ann2



Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)

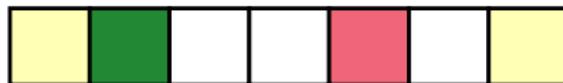


Inter-Annotator Agreement for Normalization

CzeSL (Rosen et al. 2013)

- ▶ Error tag identification

Ann1



Ann2



Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator
Agreement](#)

[Conclusion](#)

[References](#)

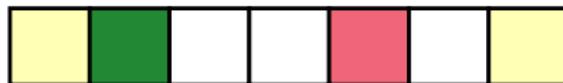


Inter-Annotator Agreement for Normalization

CzeSL (Rosen et al. 2013)

- ▶ Error tag identification

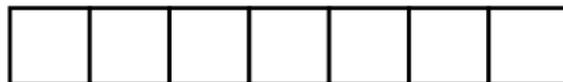
Ann1



Ann2



Ann1



Ann2



Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)



Problems with IAA for Normalization

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

Bollmann et al. (2016):

- ▶ Cohen's κ , Krippendorff's α
 - ▶ Expect finite/small set of annotations
 - ▶ Give higher weight to rare annotations
 - ▶ Miss similarities between normalizations

Introduction

Annotation

CREG

MTH

Difficult Cases

Inter-Annotator
Agreement

Conclusion

References



Problems with IAA for Normalization

Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

Bollmann et al. (2016):

- ▶ Cohen's κ , Krippendorff's α
 - ▶ Expect finite/small set of annotations
 - ▶ Give higher weight to rare annotations
 - ▶ Miss similarities between normalizations
- ▶ Special case: historical spelling normalization
 - ▶ Proposal: evaluate sets of character edit operations

Introduction

Annotation

CREG

MTH

Difficult Cases

Inter-Annotator Agreement

Conclusion

References



Problems with IAA for Normalization

Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

Bollmann et al. (2016):

- ▶ Cohen's κ , Krippendorff's α
 - ▶ Expect finite/small set of annotations
 - ▶ Give higher weight to rare annotations
 - ▶ Miss similarities between normalizations
- ▶ Special case: historical spelling normalization
 - ▶ Proposal: evaluate sets of character edit operations
- ▶ General case: modernization
 - ▶ ???

Introduction

Annotation

CREG

MTH

Difficult Cases

Inter-Annotator Agreement

Conclusion

References



Problems with IAA for Normalization

- ▶ Definition of units / items:



Normalization Evaluation: Alternatives

- ▶ Machine translation evaluation
- ▶ Paraphrase detection
 - ▶ Normalizations should be paraphrases of a target answer? ... each other?
- ▶ Plagiarism detection
 - ▶ Overlap between student answer and task context is expected
 - ▶ Normalizations should plagiarize from target answer and reading text

Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)



Normalization Evaluation: MT Metrics

Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

- ▶ Many common metrics (BLEU, TER, METEOR) don't correlate very well with human judgements
- ▶ One alternative: CharacTER (Wang et al. 2016)
 - ▶ Correlates better with human judgements for morphologically rich languages (German, Russian)

$$\text{CharacTER} = \frac{\text{word shift cost} + \text{char. edit distance}}{\# \text{ characters in the hypothesis sentence}}$$

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)



CharacT_{ER} for CREG MTH

Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

	MA Includes			
	Correct	Blend	Missing	Extra
# Answers	225	9	49	24
# Tokens	2914	157	652	455
κ , Norm. Id.	0.69	0.60	0.70	0.62
κ , Error Tag	0.47	0.43	0.49	0.29
CharacT _{ER}	0.10	0.12	0.12	0.14

- ▶ 250 appropriate answers, 2 annotators

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator Agreement](#)

[Conclusion](#)

[References](#)



Conclusion

- ▶ Normalization is hard
- ▶ Evaluation of normalization annotation is hard
 - ▶ Traditional IAA metrics are not appropriate
 - ▶ Sentence-level or text-level metrics from related tasks may be suitable

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

Introduction

Annotation

CREG

MTH

Difficult Cases

Inter-Annotator
Agreement

Conclusion

References



Preview: Word-Aligned Corpora in ANNIS

tok & tok & tok &

#1 ->align_f-g[type="good"] #2 &

#2 ->align_g-e[type="good"] #3

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator
Agreement](#)

[Conclusion](#)

[References](#)



Preview: Word-Aligned Corpora in ANNIS

tok & tok & tok &

#1 ->align_f-g[type="good"] #2 &

#2 ->align_g-e[type="good"] #3

This is an example .

this be a example .

DT VBZ DT NN .

Das ist ein Beispielsatz .

d sein ein Beispielsatz .

PDS VAFIN ART NN \$.

Xxx xxx xxx xxxxxxxxxxxxxxx .

Normalization in
context:
Inter-annotator
agreement for
meaning-based
target hypothesis
annotation

Adriane Boyd

[Introduction](#)

[Annotation](#)

CREG

MTH

Difficult Cases

[Inter-Annotator
Agreement](#)

[Conclusion](#)

[References](#)

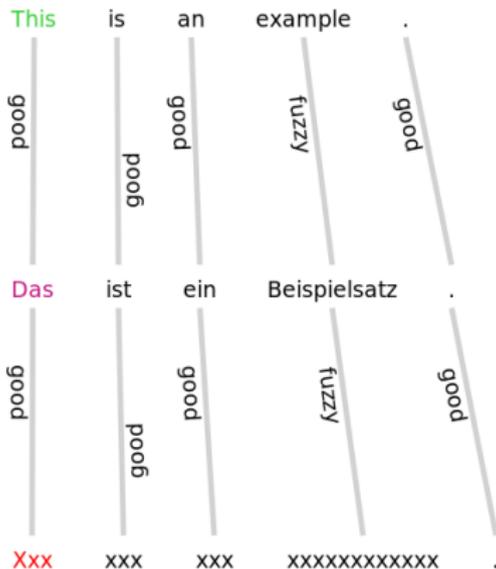


Preview: Word-Aligned Corpora in ANNIS

tok & tok & tok &

#1 ->align_f-g[type="good"] #2 &

#2 ->align_g-e[type="good"] #3



Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

Introduction

Annotation

CREG

MTH

Difficult Cases

Inter-Annotator Agreement

Conclusion

References



References

- Archer, D., M. Kytö, A. Baron & P. Rayson (2015). Guidelines for normalising Early Modern English corpora: Decisions and justifications. ICAME Journal 39.
- Bollmann, M., S. Dipper & F. Petran (2016). Evaluating Inter-Annotator Agreement on Historical Spelling Normalization. Proceedings of LAW X – The 10th Linguistic Annotation Workshop pp. 89–98.
<https://aclweb.org/anthology/W/W16/W16-1711.pdf>.
- Dahlmeier, D., H. T. Ng & S. M. Wu (2013). Building a Large Annotated Corpus of Learner English: The NUS Corpus of Learner English. In Proceedings of the Eighth Workshop on Innovative Use of NLP for Building Educational Applications. Association for Computational Linguistics, pp. 22–31.
<http://www.aclweb.org/anthology/W13-1703>.
- Fitzpatrick, E. & M. S. Seegmiller (2004). The Montclair electronic language database project. In U. Connor & T. Upton (eds.), Applied Corpus Linguistics: A Multidimensional Perspective, Amsterdam: Rodopi.
<http://purl.org/net/Fitzpatrick.Seegmiller-04.pdf>.
- Han, B. & T. Baldwin (2011). Lexical Normalisation of Short Text Messages: Makn Sens a #twitter. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. Association for Computational Linguistics, pp. 368–378.
<http://aclanthology.coli.uni-saarland.de/pdf/P/P11/P11-1038.pdf>.
- Hana, J., A. Rosen, B. Štindlová & P. Jäger (2012). Building a learner corpus. In N. C. C. Chair), K. Choukri, T. Declerck, M. U. Doğan, B. Maegaard, J. Mariani, J. Odiijk & S. Piperidis (eds.), Proceedings of the Eight International

- Conference on Language Resources and Evaluation (LREC'12). Istanbul, Turkey: European Language Resources Association (ELRA).
- Hirschmann, H., S. Doolittle & A. Lüdeling (2007). Syntactic annotation of non-canonical linguistic structures. In Proceedings of Corpus Linguistics 2007. Birmingham. <http://purl.org/net/Hirschmann.Doolittle.ea-07.pdf>.
- Laarmann-Quante, R., K. Ortman, A. Ehlert, M. Vogel & S. Dipper (2017). Annotating Orthographic Target Hypotheses in a German L1 Learner Corpus. Association for Computational Linguistics, pp. 444–456. <http://aclweb.org/anthology/W17-5051>.
- Lee, J., J. Tetreault & M. Chodorow (2009). Human Evaluation of Article and Noun Number Usage: Influences of Context and Construction Variability. In ACL 2009 Proceedings of the Linguistic Annotation Workshop III (LAW3). Association for Computational Linguistics. <http://aclweb.org/anthology/W/W09/W09-3010.pdf>.
- Lüdeling, A. (2008). Mehrdeutigkeiten und Kategorisierung: Probleme bei der Annotation von Lernerkorpora. In M. Walter & P. Grommes (eds.), Fortgeschrittene Lernervarietäten: Korpuslinguistik und Zweispracherwerbsforschung, Tübingen: Max Niemeyer Verlag, pp. 119–140.
- Meurers, D. (2015). Learner Corpora and Natural Language Processing. In S. Granger, G. Gilquin & F. Meunier (eds.), The Cambridge Handbook of Learner Corpus Research, Cambridge University Press, pp. 537–566.
- Meurers, D. & M. Dickinson (2017). Evidence and Interpretation in Language Learning Research: Opportunities for Collaboration with Computational Linguistics. Language Learning, Special Issue on Language learning research at the intersection of experimental, corpus-based and computational methods: Evidence and interpretation To appear. <http://purl.org/dm/papers/Meurers.Dickinson-17.html>.

Normalization in context:
Inter-annotator agreement for meaning-based target hypothesis annotation

Adriane Boyd

Introduction

Annotation

CREG

MTH

Difficult Cases

Inter-Annotator Agreement

Conclusion

References



- Miura, S. (1998). *Hiroshima English Learners' Corpus: English learner No. 2 (English I & English II)*. Department of English Language Education, Hiroshima University. <http://purl.org/icall/helc>.
- Ott, N., R. Ziai & D. Meurers (2012). Creation and Analysis of a Reading Comprehension Exercise Corpus: Towards Evaluating Meaning in Context. In T. Schmidt & K. Wörner (eds.), *Multilingual Corpora and Multilingual Corpus Analysis*, Amsterdam: Benjamins, Hamburg Studies in Multilingualism (HSM), pp. 47–69.
- Reznicek, M., A. Lüdeling, C. Krummes & F. Schwantuschke (2012). *Das Falko-Handbuch. Korpusaufbau und Annotationen Version 2.0*. <http://purl.org/net/Falko-v2.pdf>.
- Rosen, A., J. Hana, B. Štindlová & A. Feldman (2013). Evaluating and automating the annotation of a learner corpus. *Language Resources and Evaluation* pp. 1–28. <http://dx.doi.org/10.1007/s10579-013-9226-3>.
- Tetreault, J. & M. Chodorow (2008). Native Judgments of Non-Native Usage: Experiments in Preposition Error Detection. In *Proceedings of the workshop on Human Judgments in Computational Linguistics at COLING-08*. Manchester, UK: Association for Computational Linguistics, pp. 24–32. <http://www.ets.org/Media/Research/pdf/h4.pdf>.
- Wang, W., J.-T. Peter, H. Rosendahl & H. Ney (2016). CharacTer: Translation Edit Rate on Character Level. In *Proceedings of the First Conference on Machine Translation*. Berlin, Germany: Association for Computational Linguistics, pp. 505–510. <http://www.aclweb.org/anthology/W16-2342>.