



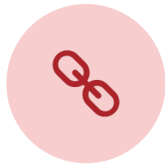
University of Antwerp
| TRICS | Translation, Interpreting
and Intercultural Studies

Annotation von Kohäsion in L2- Korpora

Präsentation des Forschungsprojektes “Academic literacy in German as a foreign language: A corpus-based contrastive analysis of cohesion in written learner language.”

Helena Wedig

Überblick des heutigen Vortrages



Was ist Kohäsion?



Einfluss der
Erstsprache auf den
Fremdsprachenerwerb



Bestehende Studien &
Korpora



Unser Vorhaben



Forschungsfragen



Belgisches
Deutschkorpus als
Datengrundlage



Annotation



Erhebung eines neuen
Korpus



Zeitliche Planung und
Ausblick



Was ist Kohäsion?

- Textkriterium, welches die Art betrifft, „wie die Komponenten des Oberflächentextes, d.h. die Wörter, wie wir sie tatsächlich hören oder sehen, miteinander verbunden sind“ (de Beaugrande/Dressler 1981: 3–4)
- Unterscheidung von 5 Kategorien (nach Halliday & Hasan, 1976)
 - Konnektive
 - (Ko-)referenz
 - Substitution
 - Ellipsen
 - Lexikalische Kohäsion

Einfluss der Erstsprache auf den Fremdsprachenerwerb

- **Bisherige Untersuchungen haben gezeigt, dass die Erstsprache eines Lernenden den Fremdsprachenerwerb beeinflusst**
- **Um Unterschiede zwischen deutschen Muttersprachlern und DaF-Lernenden herauszustellen, kann auf die Contrastive Interlanguage Analysis (CIA) zurückgegriffen werden**
 - CIA kontrastiert vergleichbare L1- und L2 Korpora und stellt Unterschiede und Gemeinsamkeiten fest
- **Unser Fokus liegt dabei auf Kohäsion**

Bestehende Studien & Korpora – L2 Deutsch

- **Belz (2005)** stellt deutlichen Mindergebrauch von Pronominaladverbien bei Lernenden heraus
 - Telecorp-Korpus besteht aus Texten von amerikanischen DaF-Studierenden und deutschen Muttersprachlern
- **Strobl (2020)** findet ebenfalls einen Mindergebrauch von Pronominaladverbien bei Lernenden
 - Folgt der Methodik von Belz (2005) mit dem Falko-Aufsatzkorpus als Datengrundlage
- **Walter (2007)** stellt fest, dass Lernende häufiger auf subordinierende Konjunktionen zurückgreifen (Falko-Zusammenfassungskorpus)

Bestehende Studien & Korpora – Übersetzung: Deutsch – Englisch

- **German English Contrast in Cohesion (GECCo) Korpus** umfasst deutsche und englische Texte verschiedener Register
- Texte wurden mit Konnektiven, (Ko-)referenz, Substitution, Ellipsen und lexikalischer Kohäsion annotiert
- Deutsch zeigt ein höheres Maß an Variation bei kohäsiven Mitteln im Vergleich zum Englischen
- Deutsch ist u.a. flexibler in Bezug auf die Verwendung von kohäsiven Ellipsen
- Deutsch enthält explizitere semantische und logische Beziehungen z.B. Konjunktionen und Koreferenzen

Bestehende Studien & Korpora – Übersetzung: Deutsch – Niederländisch

- **Van de Velde (2011)** stellt präferierte Nutzung von Koreferenz im Deutschen fest
- **Dendooven (2018)** fand Tendenz zu unterschiedlichen Realisierungen von Koreferenz z.B. bedingt durch grammatische Einschränkungen (e.g., *der Stuhl, auf dem er sitzt* vs. *de stoel waarop hij zit*) und sprachspezifische Vorlieben (e.g., *man* vs. *je*)



Unser Vorhaben

- Analyse des Einflusses der Erstsprache NL auf den - Fremdspracherwerb des Deutschen in Bezug auf die Verwendung von Kohäsionsmitteln
- Via vergleichender Analyse dreier Korpora
 - Belgisches Deutschkorpus (Beldeko) – L2 Deutsch L2, L1 Niederländisch
 - German Summary Corpus L1 (GerSumCo L1) – L1 Deutsch
 - German Summary Corpus L2 (GerSumCo L2) – L2 Deutsch, heterogene L1



Forschungsfragen

Wie verwenden DaF-Lerner beim Schreiben wissenschaftlicher Texte Kohäsionsmittel, im Vergleich zu erstsprachlichen Schreibern?

Zeigen sich besondere Ausprägungen bzgl. Kohäsion in den DaF-Texten von Lernenden mit Niederländischer L1?

Welche theoretischen Schlüsse lassen sich bzgl. der Rolle der L1 im Hinblick auf die Lernendensprache von DaF-Lernenden ziehen?



Beldeko als Datengrundlage

- 301 Texte verfasst von 115 Studierenden der Universität Ghent (Ø 230 Wörter)
- Synthesen zweier populärwissenschaftlicher Texte zu vier ausgewählten Themengebieten (Kiezdeutsch, Mundartdebatte in der Schweiz, Viadrinisch, Varianten-Wörterbuch des Deutschen)
- Texterstellung in 60min mittels eines PCs (mit Nutzung aller vorhandenen Hilfsmittel)



Beldeko als Datengrundlage

„Kiezdeutsch ist ein typischer Dialekt von den deutschen Jugendlichen . Diese Jugendsprache hat sich in urbanen Wohngebieten entwickelt , wo Menschen von verschiedenen Kulturen zusammenleben . Kiezdeutsch ist jedoch kein typisch deutsches Phänomen , auch in den Niederlanden , Dänemark und in Schweden gibt es solche Jugendsprachen . Aber zwischen diesen Jugendsprachen besteht einen gemeinsamen Punkt ; sie werden von verschiedenen Jugendlichen einheimischer und nicht-einheimischer Herkunft gesprochen . [...]“



Annotation

- Vorab: Erstellen einer Zielhypothese, um Annotation zu erleichtern
- Semi-automatische Annotation kohäsiver Mittel
 - Zu Beginn: Konnektive und (Ko-)referenz
 - Anschließend: Substitutionen, Ellipsen und lexikalische Kohäsion



Annotation

- Vorannotation greift teilweise auf bestehende Datenbanken und Algorithmen zurück
- Manuelle Annotation wird durch studentische Hilfskräfte durchgeführt
 - Mittels der Annotationsplattform Inception

Automatische Vorannotation

- Verwendung von Python zur Entwicklung einer eigenen Annotationspipeline
- Vorverarbeitung durch ParZu (Sennrich et al., 2009) – Anreicherung mit PoS-Tags, Lemmata und Abhängigkeiten
 - Nutzung von ParZu wird von CorZu erfordert
- Annotation mit koreferenziellen Relationen via CorZu (Tuggener, 2016)
 - Enthält keine Spezifizierung der einzelnen Relationen (wie z.B. kataphorisch vs. anaphorisch)

Automatische Vorannotation

- Annotation von Konnektiven mithilfe eines Abgleiches mit DimLex (Stede, 2002; Scheffler & Stede, 2016)

The screenshot shows the Connective-Lex.info website interface. At the top, there is a navigation bar with 'Connective-Lex.info', 'About / Imprint', and 'Help'. Below this is a 'Lexicon Selection' section with a dropdown menu set to 'all' and a list of lexicons with checkboxes and 'about' links. The 'DiMLex German' option is selected. The main content area displays search results for the word 'aber', indicating 274 matching results from 1 lexicon. The results are categorized by part of speech (adv) and ordering (post), with specific semantic relations listed: COMPARISON:Concession:Arg2-as-denier, COMPARISON:Concession:Arg1-as-denier, COMPARISON:Contrast, and EXPANSION:Conjunction. An example sentence is provided: 'Mein Auto faehrt nicht besonders schnell, aber das Auto von Juergen.'

Connective-Lex.info About / Imprint Help

Lexicon Selection all

- Arabic Arabic about
- DiMLex-Bangla Bangla about
- CzeDLex Czech about
- DiMLex German about
- DisCoDict Dutch about
- Eng-DiMLex English about
- LDM-PT Portuguese about
- LexConn French about
- LICO Italian about
- Naija-Lex Nigerian Pidgin about
- TCL Turkish about
- UK_DiMLex Ukrainian about

Search Options

Found 274 matching results from 1 lexicon.
DiMLex: 274

aber German DiMLex k1

adv Ordering: post

- COMPARISON:Concession:Arg2-as-denier
- COMPARISON:Concession:Arg1-as-denier
- COMPARISON:Contrast
- EXPANSION:Conjunction

Example: Mein Auto faehrt nicht besonders schnell, aber das Auto von Juergen.

- Annotation von lexikalischer Kohäsion z.B. Hyper-, Hypero- und Synonymen durch Abgleich mit GermaNet (Hamp & Feldweg, 1997; Henrich & Hinrichs, 2010)

```

#Text=Kiezdeutsch ist jedoch kein typisch deutsches Phänomen , auch in den Niederlanden , Dänemark und in Schweden gibt es solche Jugendsprachen .
3-1 192-203 Kiezdeutsch NN N Kiezdeutsch - - - - - - - -
3-2 204-207 ist VAFIN V sein - - - - - - - -
3-3 208-214 jedoch ADV ADV jedoch - - 01\_Comparison * * - - - -
3-4 215-219 kein PIAT ART ART keine - - - - - - - -
3-5 220-227 typisch ADJD ADV typisch - - - - - - - -
3-6 228-237 deutsches ADJA ADJA deutsch - - - - - - - -
3-7 238-246 Phänomen NN N Phänomen - - - - - - - -
3-8 247-248 , $, $, , - - - - - - - -
3-9 249-253 auch ADV ADV auch - - 03\_Expansion * * - - - -
3-10 254-256 in APPR PREP in - - - - - - - -
3-11 257-260 den ART ART die - - - - - - - -
3-12 261-273 Niederlanden NE N Niederlande - - - - - - - -
3-13 274-275 , $, $, , - - - - - - - -
3-14 276-284 Dänemark NE N Dänemark - - - - - - - -
3-15 285-288 und KON KON und - - - - - - - -
3-16 289-291 in APPR PREP in - - - - - - - -
3-17 292-300 Schweden NE N Schweden - - - - - - - -
3-18 301-305 gibt VVFIN V geben - - - - - - - -
3-19 306-308 es PPER PRO es - - - - - - - -
3-20 309-315 solche PIDAT ART solche - - - - *->1-1 *[1] - - - -
3-21 316-330 Jugendsprachen NN N Jugendsprache - - - - *->1-1 *[1] - - - -

```

Automatische Vorannotation - Datenformat

- Abspeicherung in verschiedenen Formaten möglich wie z.B. CONLL oder auch WebAnno TSV 3.2-Format

Planung der manuellen Annotation – Beldeko

- **Wird durch zwei SHKs durchgeführt**
 - Deutschlevel C1 bzw. Erstsprache
 - Mit linguistischem Hintergrund
- **Erfolgt in mehreren Zyklen, um Qualität via Inter-Annotator Agreement zu gewährleisten**
 - Parallele Annotation von 30 Texten
 - Anpassung der Guidelines
 - Anschließend eigenständige Annotation von jeweils 100 Texten
- **Fokus liegt zuerst auf Konnektiven**

HWedig: AliG - Annotation/HG_2013_01_T1.tsv 1-13 / 13 sentences [doc 1 / 20]

Layer Connectives

Annotation

1 Kiezdeutsch ist ein typischer Dialekt von den deutschen Jugendlichen .

2 Diese Jugendsprache hat sich in urbanen Wohngebieten entwickelt , wo Menschen von verschiedenen Kulturen zusammenleben .

3 Kiezdeutsch ist jedoch kein typisch deutsches Phänomen , auch in den Niederlanden , Dänemark und in Schweden gibt es solche Jugendsprachen .

01_Comparison 03_Expansion

Inception als Annotationsplattform

- Erlaubt die Annotation einzelner Token, Spannen sowie Relationen
- Daten werden auf Uni-eigenem Server gespeichert
- Ermöglicht die Kuratierung der Ergebnisse aller Annotierenden

Connective-Lex.info About / Imprint Help

- LDM-PT Portuguese about
- LexConn French about
- LICO Italian about
- Naija-Lex Nigerian Pidgin about
- TCL Turkish about
- UK_DiMLex Ukrainian 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 (Arg1-/ Arg2-as-instance)
 - Level-of-detail (Arg1-/ Arg2-as-detail)
 - Substitution (Arg1-/ Arg2-as-subst)
 - Exception (Arg1-/ Arg2-as-excpt)
 - Manner (Arg1-/ Arg2-as-manner)
- TEMPORAL
 - Synchronous
 - Asynchronous (Precedence/ Succession)

Found **274** matching results from **1** lexicon.
DiMLex: **274**

aber German
DiMLex k1

adv Ordering: post

- [COMPARISON:Concession:Arg2-as-denier](#)
- [COMPARISON:Concession:Arg1-as-denier](#)
- [COMPARISON:Contrast](#)
- [EXPANSION:Conjunction](#)

Example: Mein Auto faehrt nicht besonders schnell, aber das Auto von Juergen.

abgesehen davon German
DiMLex k200

adv Ordering: post

- [EXPANSION:Conjunction](#)
- [EXPANSION:Exception:Arg1-as-excpt](#)
- [COMPARISON:Concession:Arg1-as-denier](#)
- [COMPARISON:Concession:Arg2-as-denier](#)

abgesehen davon, dass German
DiMLex k201

csu Ordering: ante, post, insert

- [EXPANSION:Conjunction](#)
- [EXPANSION:Exception:Arg2-as-excpt](#)

Variants

- abgesehen davon , dass
- abgesehen davon , daß

alldieweil German
DiMLex k2

DimLex als Hilfestellung für die Annotierenden

Guidelines zu Konnektiven basierend auf PDTB3

1 Connectives

These Guidelines help to annotate Connectives in German texts. To annotate connectives, we will annotate the data according to the following procedure:

1. Identify all the existing connectives
2. Search for possible groups (multi-word connectives)
3. Evaluate the meaning given the context

1.1 How to identify Arguments

To identify arguments in the text given, it is helpful to first highlight the two ideas/sentences that are connected by the connective at hand. Afterwards the following rules help to evaluate the arguments (see Guidelines PTB3):

- The arguments to Inter-S discourse relations remain labelled by position: Arg1 is first (lefthand) argument and Arg2, the second (righthand) argument.
- With Intra-S coordinating structures, the arguments are also labelled by position: Arg1 is first argument and Arg2, the second one, independent of which argument(s) have attached coordinating conjunction(s).
- With Intra-S subordinating structures, Arg1 and Arg2 are determined syntactically. The subordinate structure is always labelled Arg2, and the structure to which it is subordinate is labelled Arg1.

1.2 How to differential between connectives with and without cohesive meaning

In order to ease the annotation process, we defined two criteria to distinguish between connectives with and without a cohesive function:

1. Cohesive connectives do not connect two VPs or two NPs with each other, they connect ideas.
2. The two connected arguments have to be recognizable.

Guidelines zu Konnektiven basierend auf PDTB3

2 Categories

We annotate the texts with the following relations:

01_Comparison This tag is used when the relation between Arg1 and Arg2 highlights prominent differences between two situations e.g., in the form contrast or concession.

Examples: allein, dafür, einerseits, nur dass, trotz, unbeschadet dessen, dass ... , wengleich, zwar

”**Dafuer, dass** er nichts gelernt hat, ist die Klausur sehr gut ausgefallen.”

02_Contingency This tag indicates that Arg1 and 2 are causally or conditionally related.

Examples: angesichts, also, um...zu, sofern, es sei denn, andernfalls

”Die Rechnung muss bezahlt werden, **andernfalls** wird der Strom abgestellt.”

03_Expansion This tag is used if the relation between Arg1 and Arg2 expands the discourse or move its narrative or exposition forward.

Examples: im Übrigen, davon abgesehen, ansonsten, anstatt, statt, anstatt dass

”Ich mag keine Fische, **davon abgesehen** mag ich alle Tiere.”

04_Temporal This tag is used if the connective indicates a temporal overlap, a preceding or a following event.

Examples: derweilen, gleichzeitig, indessen, während, anschließend, bevor, bis, danach, hieraufhin, kaum, seit, zuerst

”**Bevor** ich dieses Cabrio kaufe, muss der Haendler im Preis nachlassen.”

Guidelines zu Konnektiven basierend auf PDTB3

2.2 Error Tags

sem_error The connective is used with a wrong meaning in mind.

*”Während meiner Erasmuszeit hatte ich tatsächlich mehr Zeit zu feiern und Menschen kennenzulernen, aber ich glaube, dass auch das wichtige Erfahrungen sind, die man nur auf Erasmus lernen kann. **So** sind die Kurse an der Universität einfacher, aber lernt man die Fremdsprache viel besser beherrschen.”*

→ Here, the connective suggests the continuity of a thought, while, instead, a new topic is introduced.

syn_error The connective is used with in the right semantic setting, but it doesn't fit syntactically.

*”Das wichtigste, was man lernt, bezieht sich auf die Gelassenheit, **worüber** Gantenbrink spricht.”*

→ Here, the antecedence of the connective ”worüber” is not clear. Instead ”von der” should be used.

sem_syn_error Neither the syntactic position nor the meaning are used appropriately.

redundant The connective is redundant.

*”Gute Kompetenzen in der Standardsprache sind also notwendig um sich erfolgreich in der heutigen komplexen Gesellschaft bewegen zu können **und** die Umgangssprache kann in entspannenden Kontexten verwendet werden.”*

→ Here, the connective ”und” is not required, instead it should be two separate sentences.

missing A connective is missing here. To annotate this tag, mark the span of words that would surround the needed connective.

*”Ein Dialekt ist nicht nur eine Sprachvarietät und nicht nur ein Medium, das die Kommunikation zwischen Menschen ermöglicht. Ein Dialekt ist Teil der Kultur, Teil einer Persönlichkeit und sogar Teil einer Familie. Wenn man Niederländisch hört, dann hört man Niederländisch. Wenn man aber seinen eigenen Dialekt hört, dann hört man sein Zuhause. Obwohl **Dialekte eine** sehr spezielle Bedeutung haben, verschwinden trotzdem die meisten alten Dialekte.”*

→ Here, the last sentence requires a connective between ”Dialekte” and ”eine”, indicating a conclusion e.g. ”Also”.

style The connective doesn't feel right to use in this context (”Ausdruck”).

*”Der innovative Charakter wird erneut wahrgenommen beim Entstehen neuer Fremdwörter aus dem Türkischen und dem Arabischen. **So** wird Kiezdeutsch oft als eine Drohung betrachtet, weil sich die Menschen um einen heftigen Einfluss der türkischer Sprache fürchten.”*

→ Here, the connector ”So” indicates a connection between the two sentences, even though it does not fit contextually.

Important In addition to annotating the error tag, the specification of the target connective is needed.

Methodik zur Auswertung

- **Anwendung einer Multifactorial Prediction and Deviation Analysis (MuPDA) (Wulff & Gries, 2015)**
 - Beinhaltet die Anpassung eines Regressionsmodell an die L1 Texte, um L1-Nutzungsweisen vorauszusagen
 - Versucht anschließend die L2 Nutzung anhand des berechneten Modells vorauszusagen (entweder via Wahrscheinlichkeiten oder Kategorien)
 - Stellt unterschiedliche Nutzung heraus und evaluiert, ob Korrelationen zu finden sind
- **Im Anschluss erfolgt eine qualitative Analyse der Ergebnisse z.B. mit Fokus auf erkannte starke Abweichungen**

Erhebung eines neuen Korpus



Ziel

German Summary Corpus (GerSumCo) als
Sammlung von deutschen Syntheseschreibungen

Verwendung des Korpus als Referenz zu den DaF-
Texten im Beldeko

Bisheriger Plan: 100 Texte je Subkorpus, die
unter Beldeko-Bedingungen erhoben werden



Aufteilung in zwei Subkorpora

GerSumCo L1: Enthält Synthesen von L1
Deutsch-Schreibenden

GerSumCo L2: Enthält Synthesen von DaF-
Schreibenden mit heterogenen L1-Hintergrund

Erhebung eines neuen Korpus

- **Falko-Zusammenfassungskorpus ist aufgrund der Überlappungen der L1 und DaF-Texte mit den Originalen nicht für unseren Forschungszweck geeignet**
 - Neues Korpus erlaubt uns die Erhebung neuer Daten unter Beldeko-Bedingungen
- **Zusammenarbeit mit Hamburg – bisherige Erwartungen**
 - Dr. Patrick Grommes wird Datenerhebung im Rahmen seiner Seminare durchführen
 - Zielgruppe: Studierende der Germanistik und des Grundschullehramtes
 - Bisher geplant: Erfassung in drei Seminaren mit jeweils 15-20 Studierenden im SoSe 2022
 - Fokus auf ein Themengebiet: Viadrinisch
- **Mitarbeit weiterer Universitäten erwünscht, um Korpus stetig zu erweitern!**

Zeitliche Planung und Ausblick

Year	20	2021				2022				2023				2024		
Quarter	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3
WP LIT																
WP 1.1																
WP 1.2																
WP 1.D																
WP 2																
WP 2.D																
WP 3.1																
WP 3.2																
WP 3.D																
WP 4																
WP 4.D																
WP 5																
WP 5.D																
WP 6																
WP PHD																

- WP LIT: Acquire theoretical background through systematic literature review
- WP 1.1: Compile corpus and perform pilot annotation - Beldeko
- WP 1.2: Annotate Beldeko
- WP 1.D: Disseminate (preliminary) results
- WP 2: Collect data GerSumCo
- WP 2.D: Publish GerSumCo
- WP 3.1: Perform pilot annotation GerSumCo
- WP 3.2: Annotate GerSumCo
- WP 3.D: Disseminate annotation and/or preliminary results
- WP 4: Conduct CIA 1 (GerSumCo L1 vs. GerSumCo L2)
- WP 4.D: Disseminate results CIA 1
- WP 5: Conduct CIA 2 (GerSumCo L1 vs. Beldeko)
- WP 5.D: Disseminate results CIA 2
- WP 6: Contrast and combine results of WP 4 and WP 5
- WP PHD: Finalize (article-based) PhD

Promotionskommission



Carola Strobl

- Professorin für Angewandte Sprachwissenschaft und Deutsch
- Institut für Angewandte Sprachwissenschaft, Übersetzer und Dolmetscher
- Universität Antwerpen



Jim Ureel

- Professor für Angewandte Sprachwissenschaft
- Institut für Angewandte Sprachwissenschaft, Übersetzer und Dolmetscher
- Universität Antwerpen



Tanja Mortelmans

- Professorin für Sprachwissenschaft und Deutsch
- Institut für Linguistik
- Universität Antwerpen



Hagen Hirschmann

- Wissenschaftlicher Mitarbeiter
- Institut für deutsche Sprache und Linguistik
- Humboldt-Universität zu Berlin



Sarah Bernolet

- Professorin für Sprachwissenschaft und Niederländisch
- Institut für Linguistik
- Universität Antwerpen



Fragen?

Danke für Ihre Aufmerksamkeit!

Literatur

- **Appel, R., & Szeib, A. (2018).** Linking adverbials in L2 English academic writing: L1-related differences. *System*, 78, 115–129.
- **Belz, J. A. (2005).** Corpus-driven characterizations of pronominal da-compound use by learners and native speakers of German. *Die Unterrichtspraxis/Teaching German*, 38(1), 44–60.
- **Boeckmann, K. B. (2010).** *Grundlagen des Erst- und Fremdsprachenerwerbs: Ergänzungsheft*. Langenscheidt.
- **De Beaugrande, R. & Dressler, W. (1981).** *Einführung in die Textlinguistik (28)*. Max Niemeyer Verlag.
- **Dendooven, F. (2018).** Die Übersetzung von Koreferenzmitteln: Eine Studie auf Basis eines deutsch-niederländischen Übersetzungskorpus von Museumstexten [Unpublished master's thesis]. Ghent University.
- **Hansen-Schirra, S., Neumann, S., & Steiner, E. (2007).** Cohesive explicitness and explicitation in an English-German translation corpus. *Languages in Contrast*, 7(2), 241–265.
- **Hamp, B. & Feldweg, H. (1997).** GermaNet - a Lexical-Semantic Net for German. Proceedings of the ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications.
- **Hasan, R., & Halliday, M. A. (1976).** Cohesion in english. *Londra, Longman*.
- **Henrich, V. & Hinrichs, E. (2010).** GernEdiT - The GermaNet Editing Tool. Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC 2010), 2228-2235.
- **Klie, J. C., Bugert, M., Boulosa, B., de Castilho, R. E., & Gurevych, I. (2018).** The inception platform: Machine-assisted and knowledge-oriented interactive annotation. In Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations (pp. 5–9). Association for Computational Linguistics.
- **Kunz, K., Degaetano-Ortlieb, S., Lapshinova-Koltunski, E., Menzel, K., & Steiner, E. (2017).** English- German contrasts in cohesion and implications for translation. In G. De Sutter, M.-A. Lefer, & I. Delaere (Eds.), *Empirical translation studies: New methodological and theoretical traditions*. (pp. 265–312). Mouton, De Gruyter.

Literatur

- **Kunz, K., Lapshinova-Koltunski, E., Martínez Martínez, J., Menzel, K. & Steiner, E. (2021).** GECCo - German-English Contrasts in Cohesion: Insights from corpus-based studies of languages, registers and modes. De Gruyter Mouton.
- **Lapshinova-Koltunski, E., & Kunz, K. (2014).** Annotating cohesion for multilingual analysis. Proceedings of the 10th Joint ISO-ACL SIGSEM Workshop on Interoperable Semantic Annotation.
- **Scheffler, T. & Stede, M. (2016).** Adding Semantic Relations to a Large-Coverage Connective Lexicon of German. Proceedings of LREC.
- **Sennrich, R., Schneider, G., Volk, M., & Warin, M. (2009).** A new hybrid dependency parser for German. Proceedings of the German Society for Computational Linguistics and Language Technology, 115-124.
- **Stede, M. (2002).** DiMLex: A Lexical Approach to Discourse Markers. In: A. Lenci, V. Di Tomaso (Eds.), Exploring the Lexicon - Theory and Computation. Edizioni dell'Orso.
- **Strobl, C. (2020).** Darum sind Pronominaladverbien eine Herausforderung für Deutschlerner: Eine korpusbasierte kontrastive Interimssprachenanalyse hierzu. Germanistische Mitteilungen, 45 (1), 89-111.
- **Tuggener, D. (2016).** Incremental coreference resolution for German (Doctoral dissertation, University of Zurich).
- **Van de Velde, M. (2011).** Explizierung und Implizierung im Übersetzungspaar Deutsch Niederländisch: Eine quantitative Untersuchung. In P. A. Schmitt, S. Herold, & A. Weilandt (Eds.), Translationsforschung (pp. 865–884). Peter Lang.
- **Walter, M. (2007).** Hier wird die Wahl schwer, aber entscheidend: Konnektorenkontraste im Deutschen. In H.-J. Krumm (Ed.), Theorie und Praxis - Österreichische Beiträge zu Deutsch als Fremdsprache (Vol. 10, pp. 145–161). StudienVerlag.
- **Webber, B., Prasad, R., Lee, A., & Joshi, A. (2019).** The penn discourse treebank 3.0 annotation manual. University of Pennsylvania.
- **Wulff, S., & Gries, S. T. (2015).** Prenominal adjective order preferences in Chinese and German L2 English: A multifactorial corpus study. Linguistic Approaches to Bilingualism, 5(1), 122-150.