

The loss of the MHG bipartite negation marker *ne ... niht*

Diachronic and diatopic variation in ReM and CAO

Daniel Hrbek¹ Oliver Schallert²

¹Osnabrück University
daniel.hrbek@uni-osnabrueck.de

²Ludwig Maximilian University of Munich
oliver.schallert@lmu.de

2 July 2022



- ① Point of departure
- ② Negation: diachronic aspects
Jespersen's Cycle
Open issues

③ The empirical base

The ReM

The CAO

④ Results

Theoretical background

ReM data

CAO-data

- **Julia Hertel** (née Schüler) (Saarland University) for sharing work from her PhD thesis with us.
- **Carsten Becker** (University of Marburg) for the fruitful work on MHG charters we have done together.

- (1) a. *die **nemugin** iz uernemen*
they NEG=could it hear
(Wiener Physiologus 149va,5)
- b. *erne wil dich **nit** lazen*
he=NEG wants you NEG let
(Rolandslied 0a,8969)
- c. *Sí *ſ*prachen **nih**t an dem hilígen tag*
they talked NEG on the sacred day
(Oberaltaicher Evangelistar 25ba,37)

- In all Old Germanic languages, we find a preverbal negation particle *ni/ne* that derives from PTG **ni* (cf. Eythórsson 1995, Eythórsson 2002):

- (2) a. *ni was wulþag* (Gothic)
NEG was glorious
(2nd Letter Corinthians 3,10)
- b. *ni láz thir nan ingárgan* (OHG)
NEG let you him escape
(Otfrid IV 37,11)
- c. *út þú ne komir / órum hǫllum frá*
out you NEG come our.DAT.PL hall.DAT.PL from
(Old Norse) (Vafprúðnismál 7)

- Grimm (1890: 690) already noted that the early Germanic dialects are strikingly similar in terms of their negation patterns:

NI war die ursprüngliche und wahre negation; in der goth sprache hat sie noch den weitesten spielraum, in den übrigen nimmt sie allmählich ab, wiewohl auf verschiedene weise; heutzutage ist sie vor dem verbo überall verschwunden und den partikeln gewichen, die anfangs bloss zu ihrer verstärkung hinter das verbum gestellt wurden und zum theil mit ihr selbst zusammengesetzt sind.

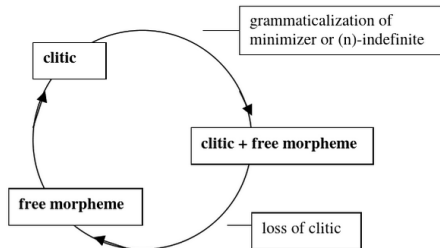
Jespersen's Cycle consists of three stages which can be observed:

- **Stage I:** the single preverbal element OHG *ni* gets (phonologically) weakened → MHG *ne/en*
- **Stage II:** now weakened *ne* is insufficient to express negation on its own and is therefore strengthened by an additional element MHG *niht* (< OHG *niouuiht* 'nothing') → **bipartite negation marker** *ne ... niht*
- **Stage III:** the original element falls victim to complete loss → MHG *niht* expresses (sentential) negation on its own

Negation: diachronic aspects

Jespersen's Cycle

Jespersen's Cycle:



	stage I: clitic	stage II: clitic + free morpheme	stage III: free morpheme
German	<i>nis<u>aga</u></i>	<i>ih <u>ensage</u> <u>niht</u></i>	<i>ich <u>sage</u> <u>nicht</u></i>
English	<i>ic <u>ne</u> <u>secge</u></i>	<i>I <u>ne</u> <u>seye</u> <u>not</u></i>	<i>I <u>say</u> <u>not</u></i>
French	<i>jeo <u>ne</u> <u>di</u></i>	<i>je <u>ne</u> <u>dis</u> <u>pas</u></i>	<i>je <u>dis</u> <u>pas</u></i>

Figure: *Jespersen's Cycle* (Jäger 2008: 15)

- For a long time, *Jespersen's Cycle* was regarded to hold for West Germanic in its entirety.
- In particular, it also applies to Low German (Breitbarth 2014).
- Therefore, MHG was assumed to represent stage II: *ne ... niht* as the main strategy of negation.
- However, recent studies for MHG (Jäger 2008; Pickl 2017) reveal that the situation in MHG is significantly more complex.

- Jägers (2008) data suggest that MHG finished *Jespersen's Cycle* at the end of the 13th century – at the latest.
- Pickl's (2017) study on sermons support Jäger's findings for Upper German → *niht* as the main negation strategy already by 1200.
- However, Witzhausen (2019) (*ne* only) and Schüler (2016)/Hertel (2022), who analyzed MHG charters (incl. Western Central German dialects) both observe a high amount of areal variation.

	Bipartite negation	Single <i>niht</i>	Total
Cologne	64 (94,1%)	4 (5,9%)	68 (100%)
Regensburg	6 (5,5%)	103 (94,5%)	109 (100%)
Zürich	1 (1,3%)	78 (98,7%)	79 (100%)
total	71 (27,6%)	186 (72,4%)	256 (100%)

Table: Variation in sentential negation in MHG charters by 1280 (Schüler 2016: 98)

Previous research shows an unclear picture:

- So far, we know that *Jespersen's Cycle* in MHG is subject to strong areal variation.
- While UG enters stage III by 1250, CG remains a stage II language by the end of MHG.
- Triggers for *Jespersen's Cycle* are still unclear.
 - Previous studies suggest syntactic (Breitbarth 2014; Jäger 2008) as well as phonological (Hertel 2022) factors.
- Both accelerating and inhibiting factors need further investigation.

We will present results from our two corpus-based studies:

- **Referenzkorpus Mittelhochdeutsch** 'Reference Corpus of MHG' (ReM) (1050–1350)
- **Corpus der altdutschen Originalurkunden bis zum Jahre 1300** 'Corpus of Old German Original Charters until the year 1300' (CAO)

Referenzkorpus Mittelhochdeutsch (ReM) :

- Freely available since 2016, contains texts from the entire Middle High German period (1050–1350)
- ~ 2 millions tokens, PoS-tagged
- **Balanced** in terms of text genres → text type-specific influences can be controlled for.
 - However, in **diachronic and diatopic matters**, the ReM is **not well-balanced**: Western Central and Eastern Upper German predominate; early texts are scarce.
- ☞ East Central dialects are excluded for the most part (relatively young, only small text portions available from this region.

(Klein et al. 2016)

- Although it uses ANNIS3 (Krause and Zeldes 2016), the ReM **lacks syntactic annotations**. → limited use for complex syntactic phenomena
- Furthermore: Crashes/overloads and occasional errors/bugs
- Queries for negation patterns are possible by indirect means (e. g. with distance parameters)
- Sample of 501 instances of *ne ... niht* (using *R*)
- Results presented here taken from Hrbek (2021)

Some background:

- *Corpus der altdeutschen Originalurkunden* „Corpus of Old German Original Charters until 1300“ (CAO) (Wilhelm et al. 1932–2004).
 - Paper edition comprising 5 Volumes plus supplementary materials; completed 2004.
 - An electronic version is available online, however, with only rudimentary search functions. [\(Link\)](#)

- Founded by Friedrich Wilhelm (1882–1932);
- Total of 4617 charters; 4289 in German (to a lesser extent also other languages are included, e. g. Latin, Middle Dutch).
- Covers a period from 1200–1300; however, only 0.2 % of the materials date before 1250 (Ganslmayer 2009: 42).

	until 1279	1280–1289	1290–1299	Total
No.	621	1095	2901	4617

Table: Extant charters w. r. t. different decades

Pros and cons:

- Issue date and place are usually mentioned;
 - The materials are lexicographically analyzed (i. e. lemmatized) via the *Wörterbuch der mittelhochdeutschen Urkundensprache* (WMU).
 - Clear dominance of Upper German regions; see the map on the next slide.
- ➡ New studies demonstrate its validity for investigating areal variation in MHG (Becker and Schallert 2021, 2022a,b; Hertel 2022).

The empirical base

The CAO

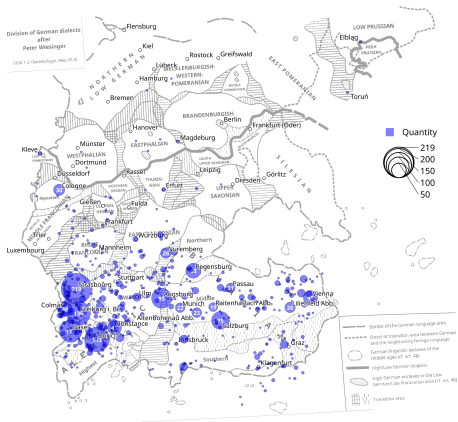


Figure: Number of charters per place (background map adapted from Schmidt, Herrgen, and Kehrein, 2008–)

A more detailed look on factors influencing negation:

- bipartite negation marker *ne ... niht* mainly attested in Western Central German texts
- early loss (or no use at all) in Upper German texts
- early loss in V1 clauses, preservation in embedded clauses (with Vend)
- (morphologically) complex verbs with prefixes (as well as participles formed with *ge-*) show an accelerated loss of *ne/en*

(Behaghel 1918; Gärtner 1977; Pickl 2017; Schüler 2016, 2017; Hrbek 2021)

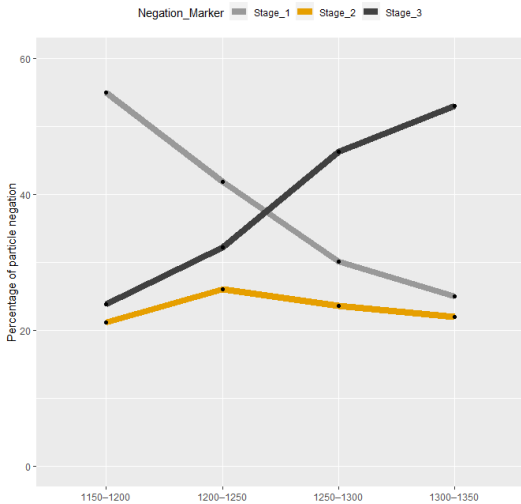
Prosody-based approach (Hertel 2022):

- *ne* (reduced variant of OHG *ni*) as a weak functional word gets reduced and falls victim to /ə/-deletion quickly.
- In V1 clauses with *ne*, the first syllable is unstressed – no trochee possible!
 - In MHG, the trochee becomes the preferred foot structure.
 - e. g. *wehsele* > *wechsel*∅ ‘exchange’
 $((\sigma_s \sigma_r \sigma_r)_F)_\omega > ((\sigma_s \sigma_r \emptyset)_F)_\omega$
- The same holds true for prefixed verbs or participles with *ge-*!
- VF clauses, on the other hand, preserve the MHG trochee → last resort for *ne ... niht*

- *ne* underwent a *lexical split* creating a homophone *ne*₂ with exceptive meaning – the ReM does not differentiate between *ne*₁ and *ne*₂
- Ignoring diatopic variation for the time being, our results confirms the findings of Jäger (2008) and Pickl (2017): *ne* gets replaced by *niht* early on.

Our results

ReM data



- MHG can be broken down into four major variants:
 - West and East Upper German
 - West and East Central German (the latter from 1250 onwards)
- Negation patterns vary considerably in the different regions.

	East Upper German		West Upper German	
	<i>ne ... niht</i>	<i>niht</i>	<i>ne ... niht</i>	<i>niht</i>
1050–1100	0	0	0	0
1100–1150	0	0	0	0
1150–1200	785	686	47	72
1200–1250	154	459	113	115
1250–1300	101	611	45	415
1300–1350	59	869	56	310
total	1099	2625	261	912

Table: Phase II und III in Upper German

	East Central German		West Central German	
	<i>ne ... niht</i>	<i>niht</i>	<i>ne ... niht</i>	<i>niht</i>
1050–1100	0	0	(2)	(2)
1100–1150	0	0	(2)	(2)
1150–1200	0	0	101	289
1200–1250	0	0	235	47
1250–1300	113	95	413	199
1300–1350	141	537	676	527
total	254	632	1427	1064

Table: Phase II and III in Central German

- **Upper German:**

- Rapid change from bipartite to postverbal *niht* between 1150 and 1200.
- After 1250, almost every negated clause contains **only *niht*** → transition to stage III

Western Central German:

- **Stable bipartite negation** until the end of the MHG period.
- The transition from stage II to III negation seems to take place after 1350. *ne ... niht* stays the **main strategy** for
- Fluctuations due to two single texts with an unusual profile (potential influence from West Upper German) expressing negation.

☞ Evidence for a stable stage II in MHG!

Verb orders with bipartite negation:

- **V1 clauses are rare** – this is due to prosodic reasons, as we discussed earlier.
 - **V2** is the **most frequent** verb order, but its rate is **decreasing** by comparison with V-end (or V-later).
 - **V-end** seems to have a **preservative effect** (cf. Behaghel 1918: 245).
- ➡ No significant areal variation in this respect (matching the results of Hertel 2022).

	V1	V2	V-end V-later
1150–1200	11	124	16
1200–1250	3	59	21
1250–1300	14	44	48
1300–1350	14	90	57
gesamt	42	317	142

Table: Bipartite Negation by verb order

Phenomena:

- Graphical variants of the **clitic negation** norm. *ne*, *en*, etc.
 - Some variants (not exhaustive) are given in (1).
 - Spellings with *i* are indicative of a stage predating phonological weakening.
- **Simple negation** (*en* or *niht*) vs. **discontinuous negation**.
 - Some variants in (2).

(1) *en(-)*, *ên(-)*, *em-*, *jn(-)*, *in(-)*, *ien*, *n-*, *-n* (Kirschstein et al. 2003: 1292–1293)

(2) *niht(-)*, *níht*, *nîht*, *njcht*, *neⁱcht*, *nieht*, *nivht*, *nũht*,
nuht/nvht (Kirschstein et al. 2003: 1312–1315)

Queries:

- We transferred the electronic version of the CAO to an SQL-database. With the aid of the *RNNTagger* (Schmid 2020), we annotated the data for POS and morphosyntactic information.
- This enabled us to query directly for the relevant lemmas *ne* and *niht* via the tag PTKNEG (STTS tagset).
- Random sample of 10% of all hits for negation structures (with *R*), manually corrected
- Graphical variants were annotated exhaustively for vowel quality.

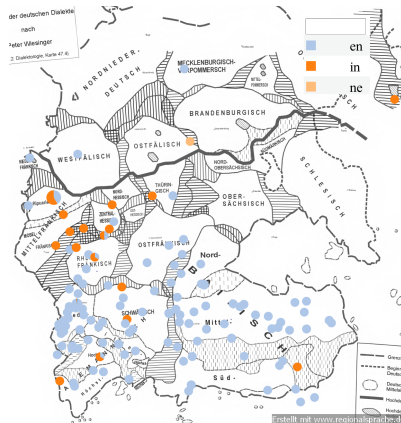


Table: Graphical variants of the clitic negation

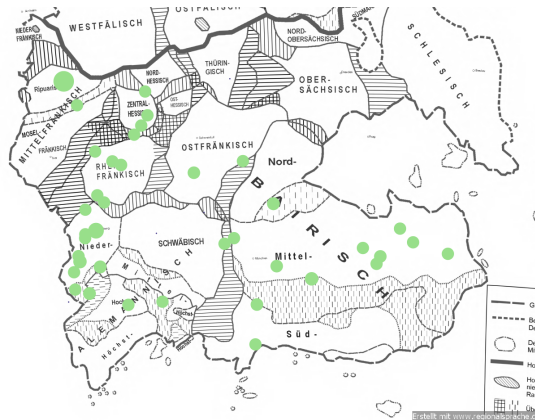


Table: Bipartite negation (*ne ... niht*) in the CAO

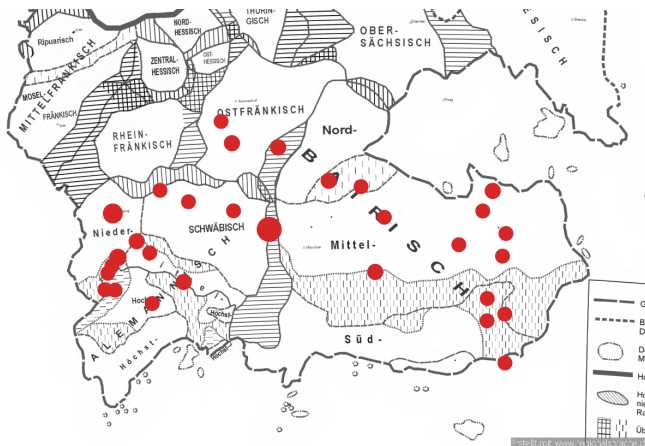





Table: Simplex negation (*niht*) in the CAO







Diatopic, diachronic and syntactic variation observed during loss of the MHG bipartite negation marker, using the ReM and the CAO...

- **Upper German:** short use and early loss
- **Western Central German:** long use and delayed loss
- **Across all dialects:** early loss in V1 clauses due to prosodic reasons





Thanks for listening!









-  Becker, Carsten and Oliver Schallert (2021). “Charters as a source for the dialectology of Middle High German”. In: *North-Western European Language Evolution [NOWELE]* 74.2: 199–241.
-  — (2022a). “Areale Variation im Bairischen des 13. Jahrhunderts: Eine Innen- und Außenperspektive”. In: *Struktur von Variation zwischen Individuum und Gesellschaft. Akten der 14. Bayerisch-Österreichischen Dialektologietagung, Salzburg, 7.–9.11.2019*. Ed. by Philip C. Vergeiner, Stephan Elspaß, and Dominik Wallner. (Zeitschrift für Dialektologie und Linguistik – Beihefte). Steiner.
-  — (2022b). “Urkunden als Quelle (schreib-)sprachlandschaftlicher Variation: Bestandsaufnahmen und Perspektiven”. In: *Zeitschrift für deutsches Altertum und deutsche Literatur [ZfdA]* 151.




-  Behaghel, Otto (1918). “Die Verneinung in der deutschen Sprache”. In: *Wissenschaftliche Beihefte zur Zeitschrift des Allgemeinen Deutschen Sprachvereins* 38/40: 225–252.
-  Boor, Helmut de et al., eds. (2004). *Corpus der altdeutschen Originalurkunden bis zum Jahr 1300*. Vol. 5. *Nachtragsurkunden 1261–1297*. Found. by Friedrich Wilhelm. Cont. by Helmut de Boor et al. Berlin: Erich Schmidt.
-  Breitbarth, Anne (2014). *The History of Low German Negation*. Oxford [u.a.]: Oxford University Press.
-  Eythórsson, Thórhallur (1995). *Verbal syntax in the early Germanic languages*. Cornell University [PhD thesis].
-  — (2002). “Negation in C: The Syntax of Negated Verbs in Old Norse”. In: *Nordic Journals of Linguistics* 25: 190–224.
-  Ganslmayer, Christine (2009). *Adjektivderivation in der Urkundensprache des 13. Jahrhunderts*. (Studia Linguistica Germanica; 97). Berlin, Boston: de Gruyter.

-  Gärtner, Kurt (1977). "Zur Negationspartikel *ne* in den Handschriften von Wolframs ‚Willehalm‘. Die mit *ne* und *niht* verneinten Sätze". In: *Wolfram-Studien* 4. Ed. by Werner Schröder. Berlin: Schmidt: 81–103.
-  Grimm, Jacob (1890). *Deutsche Grammatik, Band 3 [1831]*. Gütersloh: C. Bertelsmann.
-  Hertel, Julia (2022). *Zur Negationssyntax im Mittelhochdeutschen*. Saarland University [PhD thesis].
-  Hrbek, Daniel (2021). *Untersuchungen zu Negationsstrukturen im Mittelhochdeutschen*. University of Munich [M.A.-thesis].
-  Jäger, Agnes (2008). *History of German negation*. (Linguistik Aktuell; 118). Amsterdam, Philadelphia: Benjamins.
-  Kirschstein, Bettina et al., eds. (2003). *Wörterbuch der mittelhochdeutschen Urkundensprache auf der Grundlage des Corpus der altdeutschen Originalurkunden bis zum Jahr 1300*. Vol. 2. *hinnen dar – swester*. Berlin: Erich Schmidt.

-  Klein, Thomas et al. (2016). *Referenzkorpus Mittelhochdeutsch (1050–1350), Version 1.0*. Tech. rep. url: <https://www.linguistics.ruhr-uni-bochum.de/rem/index.html>.
-  Krause, Thomas and Amir Zeldes (2016). “ANNIS3: A new architecture for generic corpus query and visualization”. In: *Digital Scholarship in the Humanities* 31. url: <http://dsh.oxfordjournals.org/content/31/1/118>.
-  Pickl, Simon (2017). “Neues zur Entwicklung der Negation im Mittelhochdeutschen. Grammatikalisierung und Variation in oberdeutschen Predigten”. In: *Beiträge zur Geschichte der deutschen Sprache und Literatur* 139. de Gruyter: 1–46.
-  — (2021). “Ein korpuslinguistischer Beitrag zu Herkunft und Entwicklung des negativen Indefinitums *kein*”. In: *Zeitschrift für Germanistische Linguistik*. 49 (1): 84–120.

-  ReN-Team (2019). *Referenzkorpus Mittelniederdeutsch/Niederrheinisch (1200–1650), Version 1.0*. Tech. rep. url: <https://www.linguistics.ruhr-uni-bochum.de/rem/index.html>.
-  Schmid, Helmut (2020). “Deep Learning-Based Morphological Taggers and Lemmatizers for Annotating Historical Texts”. In: *Proceedings of Digital Access to Cultural Heritage*. New York, NY: Association for Computing Machinery: 133–137. doi: <https://doi.org/10.1145/3322905.3322915>.
-  Schmidt, Jürgen E., Joachim Herrgen, and Roland Kehrein, eds. (2008–). *Einteilung der deutschen Dialekte nach Peter Wiesinger*. Abzeichnung der Karte 47.4 in Wiesinger (1983). url: <https://regionalsprache.de> (visited on 11/19/2020).

-  Schüler, Julia (2016). “Alte und neue Fragen zur mittelhochdeutschen Negationssyntax”. In: *Syntax aus Saarbrücker Sicht I*. Ed. by Augustin Speyer et al. (Zeitschrift für Dialektologie und Linguistik – Beihefte; 165). Stuttgart: Steiner: 91–107.
-  — (2017). “Negationsstrukturen in den Kölner Urkunden des 13. Jahrhunderts im Vergleich”. In: *Rheinische Vierteljahresblätter* 81: 1–23.
-  Schwarz, Christian (2019). “Wortbildungswandel im frühen Deutsch. Vergleichende Analysen der Referenzkorpora Altdeutsch und Mittelhochdeutsch”. In: *Jahrbuch für Germanistische Sprachgeschichte* 10 (1): 241–260.

-  Wiesinger, Peter (1983). “Die Einteilung der deutschen Dialekte”. In: *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung*. Ed. by Werner Besch et al. Vol. 2. (Handbooks of Linguistics and Communication Science; 1). Berlin and New York, NY: de Gruyter: 807–900. doi: 10.1515/9783110203332-003.
-  Wilhelm, Friedrich, ed. (1932). *Corpus der altdeutschen Originalurkunden bis zum Jahr 1300*. Vol. 1. 1200–1282. Lahr: Schauenburg.
-  Witzenhausen, Elisabeth (2019). *Negation – Exception – Contrast. The post-cyclical development of ne/en in Middle High German, Middle Low German and Middle Dutch*. Universiteit Gent [PhD thesis].

- Due to the lack of traditional texts (not only within the ReM), ECG dialects can only be investigated from the second half of the 13th century.
 - ECG dialects are the result of the German colonization of the East (late 10th – 13th century)
 - All in all, we can observe two periods of 50 years each: (i) 1250–1300 and (ii) 1300–1350.
 - Therefore, the results should be treated with caution.
- What we can observe: During the **second half of the 13th century**, the **bipartite negation marker is (still) in its peak**, but (most probably) already declining.
- However, this doesn't last for long: Within just 50 years, **the tide turns completely**: During the 14th century, ***niht*** rises quickly and **becomes the main negation strategy** (by far!) and displaces *ne ... niht*.

	Eastern Central German	
	Bipartite	Postverbal
1150–1200	–	–
1200–1250	–	–
1250–1300	113	95
1300–1350	141	537
Total:	254	632

Table: Frequencies of stage II and III in comparison (ECG only)

- As we have already discussed, the position of the finite verb does not (significantly) vary between the four major MHG dialect groups.
- Nevertheless, there is one conspicuous feature that is noticeable in Eastern Central German: We can observe **a (slightly) higher frequency of verb-initial clauses** in ECG.
 - While VF clauses are rare (even during the last 50 years), V1 clauses are pretty common.
 - However, this may be due to the **conjunct *vnde*** which may occur in the form of *en* or *in*.
 - In some cases, we can clearly decide whether *ne* represents 'and' or 'not' – but in the majority of instances, we can't.

Appendix

Jespersen's Cycle in Eastern Central German (ECG)

	Verb-initial	Verb-second	Verbs-later/-final
Eastern Upper German	15	142	26
Western Upper German	2	27	12
Eastern Central German	12	20	6
Western Central German	13	128	98
Total	42	317	142

Table: Verb position with the bipartite negation in the four major dialect groups

- As we said earlier, the ReM is an annotated and PoS-tagged corpus consisting of approx. two million tokens.
- Although it exists since 2016, there are hardly any studies based on/using the ReM so far (Pickl 2021; Schwarz 2019; Witzhausen 2019).
- It makes use of ANNIS 3 (Krause and Zeldes 2016) and (in theory) allows for complex queries with e.g. syntactic boundaries, clause types and so on.
- Due to its features and its broad well-balanced texts, it can be used for diatopic and diachronic studies on MHG morphology and/or syntax.
- It is part of a larger series of reference corpora for Elder High (OHG, MHG, ENHG) and Low German (OS, MLG).

- With its sisters (*Referenzkorpus Altdeutsch* (ReA); *Referenzkorpus Mittelniederdeutsch/Niederrheinisch* (ReN; ReN-Team 2019), it forms the joint network *Deutsch Diachron Digital* (DDD)).
- However, from the perspective of a syntactician, there is **one major problem**:
- Unlike the others, the ReM **lacks syntactic annotations**: While we have the commands (e.g. parameter *bound_sent*), we can't use them because the underlying information is missing.
- For our study, we would have used sentence boundaries to exclude hits in which the bipartite negation marker is spread over two clauses.

- Instead, we decided to work with distance parameters (e.g. no more than five token between the two negative elements) – definitely not the best option, but a sufficient and easy work-around.
- Although it works and has an acceptable precision, a lot of additional manual work is necessary.