The historical development of the Dutch posture-verb progressive construction: including a comparison with German
Okabe, A.

Citation

Version: Publisher's Version
License: Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from: https://hdl.handle.net/1887/3564457

Note: To cite this publication please use the final published version (if applicable).
Appendix A  Annotation and possible values per parameter for the Dutch database

This appendix describes the annotation of the Dutch data extracted from the corpora. The data are annotated in the database (cf. Okabe 2022) in terms of the data source (see (1) below), the inflection of the posture verb (2), the form of the connector (3), the syntactic and semantic features of the second verb (4), some features of the subject (5), the placement and some features of the object, if present (6), the presence/absence and the placement of the modifier (8), structural information (9), and other characteristics (10). Note that not all the information coded in the database is systematically discussed in the text, e.g. person and mood specification on the verb.

The specification of the data source includes the following four kinds of information:

(1) a. Name of the corpus: Corpus Gysseling,
Corpus Middelnederlands,
Corpus literair Nieuwnederlands
b. Name of the document: derived from the corpus

c. Publication year (in centuries): 13, 14, 15, 16, 17, and 18
d. Text type: prose, verse, and prose and verse for Middle Dutch;
prose, drama, non-fiction for Early Modern Dutch

Parameter (1a) has three options, as described in Chapter 2: ‘Corpus Gysseling’, ‘Corpus Middelnederlands’, and ‘Corpus literair Nieuwnederlands’. Parameter (1b) provides the name of the document from which each instance is extracted (this information is derived from the corpora). The publication year is also taken from the corpora and classified per century, i.e. 13th, 14th, 15th, 16th, 17th, and 18th century. This means, for example, that the publication years 1234, 1250, and 1289 would all be annotated as ‘13’ in the database. The text type (1d) also reflects the information given in the corpora. As illustrated in Chapter 2, the two Middle Dutch corpora are principally based on the binary classification of prose vs. verse, although there are two texts in the Corpus Middelnederlands which are marked as ‘prose and verse’. For the Corpus literair Nieuwnederlands, three categories are distinguished: prose, drama, and non-fiction. Hence, for both Middle Dutch and Early Modern Dutch, three values are available for the text type parameter.
The historical development of the Dutch posture-verb progressive construction

The information regarding the posture verb is annotated in the following way:

(2)  a. Posture verb: staan, zitten, liggen
    b. Person: 1st, 2nd, 3rd, na
    c. Number: singular, plural, na
    d. Tense/mood/finiteness: present, past, imperative, infinitive, subjunctive

Posture verb type (2a) has three options corresponding to the kind of posture verb in question: ‘staan’, which covers both the Middle Dutch form staen and the (Early) Modern Dutch form staan; ‘zitten’, which covers Middle Dutch sitten and (Early) Modern Dutch zitten; and ‘liggen’ for liggen. The inflection of the posture verb—that is, its person (2b), number (2c), and tense, mood, and finiteness (2d)—is annotated to enable the assessment of distributional deviations. For cases where person and number are not relevant, for example when the verb is in the infinitive, the instance is marked as ‘na’ (which stands for ‘not applicable’) in the database.

The connector type is annotated in terms of the observed form of the connector, as in (3).

(3)  Connector: en, ende, te, none

The syntactic and semantic features of the second verb are the most extensively annotated elements in the database, as summarized in (4).

(4)  a. Verb number: 2, 3, 4, 5
    b. Verb type: dictionary form of the second verb
    c. Conjugation: ++, +, --
    d. Dynamic: +, -
    e. Atelic: +, -
    f. Compatible with the posture: +, -
    g. No movement: +, -

First, the sequential number of each verb in the construction is annotated. In most cases, the verb is marked as ‘2’, meaning that the verb in question is the first verb following the posture verb in the construction. In some cases, the instance has three or more verbs in close vicinity, e.g. ?hij zat en at en drank ‘he sat and ate and drank’. In this case, two entries are created: one for the verb pair zitten and eten, where eten is annotated as ‘2’, and the other for
zitten and drinken, where drinken is annotated as ‘3’. Verb type (4b) provides the dictionary form of the second verb. The dictionaries consulted are the *Middelnederlandsch Woordenboek* for the instances extracted from the Middle Dutch corpora and the *Woordenboek der Nederlandsche Taal* for the instances from the Early Modern Dutch corpus. Some of the verbs are additionally marked with ‘/’ (e.g. /wachten) to exclude them from the analysis of the HTR of the second verb (cf. Appendix B). Conjugation (4c) reflects whether the posture and the second verb are conjugated or not. When a verb is conjugated, it is annotated as ‘+’, otherwise as ‘−’. Since there are two verbs and two options for each verb, four possibilities exist: ‘++’, ‘+-’, ‘-+’, and ‘−−’. Note, however, that cases where the posture verb is not conjugated while the second verb is conjugated (i.e. ‘−+’) are not included in the database (cf. section 2.2.3.). Features (4d-g) are associated with the semantic features of the second verb (cf. section 3.4.1.); whether each semantic feature is observed or not is marked as ‘+’ or ‘−’, respectively.

The characteristics of the nouns in the construction are annotated as shown in (5) and (6).

(5) Presence of the subject for the second verb: +, −

(6) a. Presence of the object: +, −
    b. Position of the object: 1, 2, 3, 4
    c. Extracted/preposed object: e, p, or none

The presence/absence of the subject and the object for the second verb is annotated as ‘+’ (present) or ‘−’ (absent) (5 & 6a). The position of the object (6b) is annotated by distinguishing four locations: before the posture verb [1], between the posture verb and the connector [2], between the connector and the second verb [3], and after the second verb [4]. This is shown schematically in (7).


As shown in (6c), instances with extracted and preposed objects are annotated as ‘e’ and ‘p’, respectively, and those with neither are marked by an empty cell.

The annotations related to the modifiers are summarized in (8).

(8) a. Position of the adverbial before the posture verb [1]: +, none
The historical development of the Dutch posture-verb progressive construction

b. Position of the adverbial between the posture verb and the connector [2]: +, none
c. Position of the adverbial between the connector and the second verb [3]: +, none
d. Position of the adverbial after the second verb [4]: +, none
e. Presence of the locative modifier: +, -
f. Presence of the temporal modifier: +, -
g. Negator for the posture verb: +, none
h. Negator for the second verb: +, none

First, the position of the adverbial is marked (from 8a to 8d) using the same schema as presented in (7). Second, the presence/absence of two types of modifiers, i.e. locative and durative temporal adverbials, is marked (8e & f). Third, the presence/absence of a negator is annotated (8g & h). When the negator is in the position to negate either a posture verb or second verb, it is marked as ‘+’.

The word order is annotated according to the position of the posture verb, as shown by (9).

(9) Location of the posture verb: nf, f

In this research, I distinguish instances where the posture verb is located in clause-non-final position and where it is located in the clause-final verbal complex (cf. section 3.4.). The former is annotated as ‘nf’ and the latter as ‘f’.

Finally, three types of additional information are annotated in the column named ‘Others’, where applicable. First, when a sentence shows the IPP effect (see section 1.2.2. for an explanation of the IPP effect), it is marked as ‘ipp’. Second, when the posture verb is used with a non-literal meaning (e.g. using liggen to mean ‘to stay’; cf. section 4.2.1.), the instance is annotated as ‘nl’. Lastly, when the posture verb and the second verb can be interpreted as disagreeing in number (cf. section 1.3.3. and 4.2.4.), it is marked as ‘nd’, which stands for ‘number disagreement’.

(10) Others: ipp, nl, nd
Appendix B  Dataset for the analysis of the hapax-token ratio

This appendix describes how the datasets were prepared for the analysis of the hapax-token ratio (HTR). In preparing the datasets, recurring Bible translations are first removed. The resulting dataset for the 13th and 14th centuries includes about 4.4 million tokens, that for the 15th and 16th centuries about 5.6 million tokens, and that for the 17th and 18th centuries around 6.5 million tokens. Subsequently, the datasets for the 15th and 16th centuries and the 17th and 18th centuries are reduced in size, so that all the datasets are comparable. The reduction is made by excluding some texts from the dataset. The texts to exclude are randomly chosen from each text genre per century in turn. A total of 39 texts are removed from the dataset for the 15th and 16th centuries and 65 texts from the dataset for the 17th and 18th centuries. The resulting datasets each consist of about 4.4 million tokens, as shown by Table 1.

<table>
<thead>
<tr>
<th></th>
<th>13th &amp; 14th</th>
<th>15th &amp; 16th</th>
<th>17th &amp; 18th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size (in number of words)</td>
<td>4,413,251</td>
<td>4,419,612</td>
<td>4,428,357</td>
</tr>
</tbody>
</table>

In the database, when the second verb (annotated as in (4b)) derives from the excluded texts, it is marked with ‘/’ (e.g. /eten) so that it can be distinguished and excluded from the HTR analysis.
Appendix C  Dataset for the diachronic development of the coordinating conjunction

This appendix describes how the data were prepared for the analysis of the coordinating conjunction. As outlined in 4.2.4., the analysis concerning the replacement of the connector *ende* by *en* (cf. Hypothesis 4) includes an investigation of the alternation of the coordinating conjunction between the forms *ende* and *en*. The development of the coordinating conjunction is examined by extracting data from randomly selected texts, which are evenly distributed across centuries and text genres. The names of the texts used are listed below in (1).

(1) a. *Spiegel historiael* (eerste partie/derde partie/vierde partie) for the early 14th century
b. *Stuttgarts leven van Jezus* for the early 14th century
c. *Roman van Walewein* for 1350 for the late 14th century
d. *Nieuwe Testament (Nndl. vert.)* for the late 14th century
e. *Brabantsche yeesten (boek 6)* for the early 15th century
f. *Historie van den grale* for the early 15th century
g. *Karel ende Elegast* for the late 15th century
h. *Historie van Reynaert die vos, Proza-Reinaert* for the late 15th century
i. *Devoot ende profitielyck boecxken, Liedboek van 1539* for the early 16th century
j. *Historie vanden vier heemskinderen* for the early 16th century
k. *Antwerpse blydschap* for the late 16th century
l. *Historie van Malegij* for the late 16th century
m. *Nederlandsche Historien* (boek 1-8) for the early 17th century
n. *Beschrijvinge der stad Leyden (fragment)* for the early 17th century
o. *Palamedes oft Vermoorde onnooselheyd* for the early 17th century
p. *Het leven van Joost van den Vondel* for the late 17th century
q. *Den vermakelyken avanturier (behalve laatste boek)* for the late 17th century
r. *Haagsche broeder-moord of dolle blydschap* for the late 17th century
s. *Vaderlandsche historie (t.e.m. boek 8 XIX)* for the early 18th century
t. *De Rotterdamsche Hermes* for the early 18th century
u. *Het wederzuyds huwelijksbedrog* for the early 18th century
v. *Het onseidbaar drietal redenwezens verlichting, deugd en tijd* for the late 18th century
w. *De vrouwelijke Cartouche* for the late 18th century
x. *De patriotten* for the late 18th century
The frequencies of the coordinating conjunction in the form of *ende* or *en* extracted from these texts are reported in 4.2.3.
Appendix D  Annotation and possible values per parameter for the German database

This appendix describes the annotation of the German data extracted from the corpus (cf. Okabe 2022). The German data are annotated largely in the same manner as the Dutch instances, as described in Appendix A. First, the name of the document and the publication year of the source document are specified in the following way:

(1) a. Name of the document: as given in the corpus

For posture verbs, the annotations cover the type of verb, the conjugation of the verb, and its person, number, and tense and mood, as shown in (2).

(2) a. Posture verb: stehen, dastehen, sitzen, dasitzen, liegen, daliegen
    b. Conjugation: +, -
    c. Person: 1st, 2nd, 3rd, na
    d. Number: singular, plural, na
    e. Tense/mood/finiteness: present, past, perfective, infinitive, subjunctive

Each instance in the database is marked in terms of what type of posture verb it includes (2a). As shown in Chapter 5, posture verbs with the particle da-, i.e. dastehen, dasitzen and daliegen, are analyzed separately from the verbs without the particle. For both the posture verb and the second verb, the presence or absence of conjugation is annotated as ‘+’ or ‘-’ respectively, in (2b). The annotation here is simpler than in Dutch, as inflectional mismatches between the verbs do not occur. Therefore, the ‘+’ or ‘-’ reflects whether or not both the posture verb and second verb are conjugated. The number and the tense/mood/finiteness are annotated in the same way as with the Dutch data (see Appendix A).

The syntactic and semantic features of the second verb are annotated with the parameters shown in (3).

(3) a. Verb number: 2, 3, 4
b. Verb type: dictionary form of the second verb
   c. Dynamic: +, -
   d. Atelic: +, -
   e. Compatible with the posture: +, -
   f. No movement: +, -

All the features are annotated in the same way as for the Dutch data (see Appendix A). The form of the connector is not annotated, since German only has one form: *und*.

As for the noun, the annotation covers the presence or absence of the subject for the second verb, the presence or absence of an object, and the presence of an extracted or preposed object.

\[(4)\]
   a. Presence of the subject for the second verb: +, -
   b. Presence of the object: +, -
   c. Extracted/preposed object: e, p, or none

Parameters \((4a-c)\) are annotated in the same manner as for Dutch (cf. Appendix A).

For the modifier, the annotation covers the position of the adverbials (if any), the presence or absence of certain types of modifiers, and the modification relation of the negator.

\[(5)\]
   a. Position of the adverbial before the posture verb \([1]\): +, none
   b. Position of the adverbial between the posture verb and the connector \([2]\): +, none
   c. Position of the adverbial between the connector and the second verb \([3]\): +, none
   d. Position of the adverbial after the second verb \([4]\): +, none
   e. Presence of the locative modifier: +, ++, -, na
   f. Presence of the temporal modifier: +, -
   g. Negator for the posture verb: +, none
   h. Negator for the second verb: +, none

All the points except for \((5e)\) are annotated in the same way as in Dutch (see Appendix A). With regard to \((5e)\), when an instance has one or more locative modifiers, it is marked as either ‘+’ or ‘++’. The difference between ‘+’ and ‘++’ is that the latter denotes deictic locative adverbs (e.g. *da* ‘there’, *hier* ‘here’) and the former covers the rest (cf. section 5.3.4.). The label ‘na’ is given when
The historical development of the Dutch posture-verb progressive construction

the instance includes one of the *da*-verbs, i.e. *dastehen, dasitzen,* or *daliegen,* and is not further modified for location.

The position of the posture verb is annotated by distinguishing cases where the posture verb is found in clause-non-final position versus clause-final position; these situations are annotated as ‘nf’ and ‘f’, respectively.

(6) Location of the posture verb: nf, f

Finally, cases of ‘subject lacking in finite clauses’ (SLF coordination, cf. 5.2.2.) are marked as ‘slf’ in the column ‘Others’.

(7) Others: slf, none
Bibliography

Corpus

Corpus literair Nieuwnederlands = Geleyn & Colleman (2015)

Dictionary


Literature

ANTHONISSEN, LYNN; ASTRID DE WIT; and TANJA MORTELHANS. 2019. (Inter)subjective uses of the Dutch progressive constructions. Linguistics 57 (5). 1111-1159.
The historical development of the Dutch posture-verb progressive construction


BOUZOUITA, MIRIAM; ANNE BREITBARTH; LIEVEN DANCKAERT; and MELISSA FARASYN. 2019. The determinants of diachronic stability. *The


The historical development of the Dutch posture-verb progressive construction


230 The historical development of the Dutch posture-verb progressive construction

KINN, TORODD; KRISTIAN BLESENIEUS; and PETER ANDERSSON. 2018. Posture, location, and activity in Mainland Scandinavian pseudocoordinations. CogniTextes 18. Online: https://journals.openedition.org/cognitextes/1158.


KUTSCHER, SILVIA, and EVA SCHULTZE-BERNDT. 2007. Why a folder lies in the basket although it is not lying: The semantics and use of German positional verbs with inanimate Figures. Linguistics 45 (5-6). 983-1028.


LÓPEZ-COUSO, MARÍA JOSÉ. 2016. Continuing the dialogue between corpus


NACCARATO, CHIARA. 2016. A corpus-based quantitative approach to the study of morphological productivity in diachrony: The case of samo-compounds in Russian. *A blend of MaLT: Selected contributions from the methods and linguistic theories symposium 2015*, ed. by HANNA CHRIST; DANIEL KLENOVŠAK; LUKAS SÖNNING; and VALENTIN WERNER, 133-152. Bamberg: University of Bamberg Press.


OKABE, AMI. 2022. Posture-verb constructions in Dutch and German.
The historical development of the Dutch posture-verb progressive construction

DataverseNL. doi: https://doi.org/10.34894/BQIQBU.


by JÖRG MEIBAUER; MARKUS STEINBACH; and HANS ALTMAANN, 356-371. Berlin: Mouton de Gruyter.


The historical development of the Dutch posture-verb progressive construction

WEIJNEN, A. A. 1971. Schets van de geschiedenis van de Nederlandse syntaxis. Assen: Van Gorcum & Comp. N.V.