

MANUAL FOR THE USE OF NEHOL: THE NEGERHOLLANDS DATABASE*

Robbert van Sluijs, with the assistance of Margot van den Berg, Pieter Muysken, and Cefas van Rossem

Languages in Contact (LinC) Group
Centre for Language Studies

Radboud University Nijmegen



* This database has been based on earlier material from an earlier NWO grant to Hans den Besten and Pieter Muysken for the digitalization of Negerhollands texts, carried out by Cefas van Rossem and Hein van der Voort. The present database was supported by a CLARIN-NL grant to Pieter Muysken and Margot van den Berg, with Robbert van Sluijs as project researcher.

CONTENTS

1. Introduction
 - 1.1 The NEHOL project
 - 1.2 Negerhollands and a short history of the Danish West Indies
2. The structure of the NEHOL database
 - 2.1 Instructions on how to use the database
 - 2.1.1 Accessing the datafiles
 - 2.1.2 Searching the datafiles
 - 2.2 Subcorpora
 - 2.3 Sources/Sessions
 - 2.3.1 Annotated sources
 - 2.3.2 Other sources
 - 2.3.3 Complete overview of NEHOL sources and their references
 - 2.4 Metadata
 - 2.4.1 Metadata files
 - 2.4.2 References of the metadata
 - 2.4.3 Diplomatic editing codes
3. Annotations
 - 3.1 Introduction
 - 3.1.1 The annotation layers
 - 3.1.2 Justification of the classification used for annotating
 - 3.2 Open class categories
 - 3.2.1 Nouns
 - 3.2.2 Verbs
 - 3.2.3 Adjectives
 - 3.2.4 Adverbs
 - 3.2.4.1 Intensifiers
 - 3.2.5 Reduplication
 - 3.2.6 Interjection
 - 3.2.7 Ideophone
 - 3.3 Closed class categories
 - 3.3.1 Adpositions
 - 3.3.2 Conjunctions
 - 3.3.2.1 Subordinating conjunctions
 - 3.3.2.2 Relative pronouns
 - 3.3.2.3 Coordinating conjunctions
 - 3.3.3 Copula
 - 3.3.4 Determiners
 - 3.3.5 Quantifiers
 - 3.3.6 Numerals
 - 3.3.7 Focus marking
 - 3.3.8 Inflection
 - 3.3.9 Negation
 - 3.3.10 Pronouns
 - 3.3.10.1 Personal and possessive pronouns
 - 3.3.10.2 Demonstrative pronouns
 - 3.3.10.3 Indefinite pronouns
 - 3.3.10.4 Reflexive pronouns

3.3.10.5 Reciprocal pronouns

3.3.11 Proform

3.3.12 TMA marking

3.4 Other categories

3.4.1 Abbreviations

3.4.2 Foreign words

3.4.3 Formulaic expressions

3.4.4 Punctuation

3.4.5 Unclassified words

4. References

1. Introduction

1.1 The NEHOL project

This is a manual written for users of the Negerhollands database 'NEHOL', a project financed by CLARIN-NL. The goal of this project is to make the data from the Dutch-lexifier Creole language Negerhollands publicly available in a digitally accessible and searchable database, in the same format as the parallel SUCA (SURiname Creole Archive) corpus, coordinated by Margot van den Berg (Radboud University Nijmegen). The NEHOL project is coordinated by Pieter Muysken (Radboud University Nijmegen), and technically supported by Daan Broeder and Paul Trilsbeek from TLA ('The Language Archive') at the MPI for Psycholinguistics in Nijmegen.

The majority of the data included in the database have been digitalized and carefully edited in a previous NWO financed project from the first part of the '90s, carried out by Cefas van Rossem and Hein van der Voort. This mostly concerns the eighteenth and nineteenth century documents written by Moravian missionaries and discovered by dr. Peter Stein in the Archiv der Brüder-Unität in Herrnhut (Germany). Other documents digitalized in the previous project mainly consist of primers, hymn books and documents from the Danish Lutheran mission. This project also resulted in the publication of *Die Creol Tael* (Van Rossem & Van der Voort 1996), a comprehensive book discussing and exemplifying the broad variety of Negerhollands sources and their different varieties of the language.

All documents included in the NEHOL project have been checked with copies of the original manuscripts for errors, as it was not feasible to do this for all the data in the previous project. Additional sources have been digitalised for NEHOL, of which the 1926 publication by De Josselin de Jong, containing 58 pages of basilectal Negerhollands data plus a wordlist, is the most important. The main feature of NEHOL is that a considerable and representative part of the database is annotated, which enables the user to make advanced search queries.

The annotations consist of three lines: 1) a morphological line, containing a standardised lemma for each lexical and functional item to overcome spelling variation in search queries; 2) a line with glosses, providing information on the meaning of content words and the grammatical properties of function words; 3) a line containing Part of speech tags. The annotations are discussed in detail in section 3 of the manual.

1.2 Negerhollands and a short history of the Danish West Indies

Negerhollands is the extinct creole language that used to be spoken in the Danish West Indies, from 1917 on known as the US Virgin Islands. In spite of its name, 17th century Zeelandic, Zealand-Flemish and West-Flemish (South-Western Dutch dialects) are hypothesized to be the main superstrate languages. Other superstrate influences are Standard Dutch, English, Portuguese/Spanish (possibly via Papiamentu), French, and Danish. Negerhollands is special in that it is one of very few Dutch-based creoles (Berbice Dutch is a notable other), but also one of which we have a considerable amount of data available, collected in this database. The exact substrate languages of the 17th century slaves that created Negerhollands have not been uncontroversially identified, but they probably belonged to the Kwa language family spoken on the African West Coast in and around current day Ghana. For an overview of the grammatical properties of Negerhollands, I refer the reader to Van Sluijs (2013).

Negerhollands was the result of the inhumanly cruel reality of colonial slavery in the Caribbean. It was the language of the slave population of the plantations on the islands of St. Thomas and St. John, and to a lesser extent on St. Croix, where an English creole was spoken on a large number of plantations from the beginning. An alternative name that is used is Virgin Islands Dutch Creole.

The Danish started their colony in 1672 on St. Thomas. Although the official positions were filled with Danishmen, the colonists were very heterogeneous in nationality. St. Thomas was a plantation colony and the majority of the plantation owners were Dutch, which resulted in Dutch being an important lingua franca on St. Thomas. This explains why the creole language that emerged among the plantation slaves had Dutch or Flemish/Zelandic varieties of Dutch as its main lexifier. In 1717 the Danish occupied the neighbouring island St. John, and in 1733 they bought St. Croix from the French.

The Moravian mission on St. Thomas starts in 1732. Converting the slaves was not easy and in addition the missionaries were opposed by the colonists who saw religious activity as only keeping the slaves from their work. The missionaries eventually succeeded in building a congregation of slaves by addressing them in their own language: Negerhollands. However, not all of the missionaries were able to master a variety of Negerhollands. Friedrich Martin, for example is mentioned as introducing Dutch as a lingua franca to communicate with the slaves (Stein 2010:212-213, footnote 16). The slave letters included in the database are all the product of writing lessons by the Moravian missionaries.

The Danish Lutheran church was much later in establishing a mission in the Danish colony to convert the slaves, which started only in 1756. The Danish missionaries were very productive in printing primers and hymn books, of which a number is included in NEHOL.

The Dutch Creole languages was however not limited to the slaves. A considerable part of the white community of the colony spoke an acrolectal variety of Negerhollands, which was much more like Dutch. This is not surprising as the white community always spoke other European languages beside their ancestral native language. A few data of this acrolectal Negerhollands is present in NEHOL as well.

In 1848 slavery was abolished in all of the colony. Many of the former slaves took to the towns, where English was the main language. On St. Croix, English had always been an important language and the slaves with English owners spoke an English creole. From the turn of the 18th to the 19th century, English creole started to be more and more common on St. Thomas and St. John as well, at the expense of Negerhollands. But Negerhollands was able to hold out for long. This is proven by the fact that the Dutch scholar De Josselin de Jong was able to collect mostly folk stories by native speakers of Negerhollands on St. Thomas and St. John in 1922 and 1923. In 1936, Frank Nelson even collected words and sentences of Negerhollands on St. Croix. These two sources have been included in NEHOL, including the long version of Frank Nelson's field notes (see Van Rossem, to appear).

Even in the 1970s, there were speakers of Negerhollands to be found. Anne Victoria Adams Graves and Gilbert Sprauve have worked with them. Gilbert Sprauve and his then student Robin Sabino have made recordings of the last speaker, Mrs. Alice Stevens in the 1980s. She passed away in 1987, and so did Negerhollands.

2. The structure of the NEHOL database

2.1 Instructions on how to use the database

2.1.1 Accessing the datafiles

The NEHOL database is hosted by the MPI TLA ('The Language Archive'). It is accessible to anyone via the Internet without restrictions. The webaddress to access the NEHOL database is:

http://corpus1.mpi.nl/ds/imdi_browser/

Within the list of available corpora, look for NEHOL. Click on the round symbol before the name, not on the name itself. Now the available subcorpora will reveal themselves to you. Again, click on the round symbol before the name of the subcorpus you want to explore or get access to. Then, all the sessions in the subcorpus will appear. To access the metadata, this time click on the name of the session. To access the data itself, click on the round symbol and subsequently click on the name of the desired datafile. On the right side of the webbrowser, a screen will appear. The first piece of information given is a URL. Click on this URL to gain access to the file, which you can open or save to your computer.

The two files from the subcorpus 'Negerhollands basilectal data' 1926 DJDJ texts and 1926 DJDJ wordlist make use of special symbols. To guarantee that these are correctly rendered on your screen, make sure that you have the font 'Junicode' installed on your computer. It can be downloaded for free from the following URL:

<http://sourceforge.net/projects/junicode/files/>

Once you have installed the font, follow the following steps to access these files with the symbols rendered correctly. Click on the name of the desired datafile and then click on the URL that appears above on the right side of the screen. Now, do not open the file but save it to your computer. Open the file with a program that allows you to read it in Junicode - a UTF-8 font. An example of such a program for Windows is Microsoft Word.

2.1.2 Searching the datafiles

A metadata search can be performed by right-clicking on the name of the desired session and choosing 'metadata search'. This search can be performed on multiple sessions by holding the control button on your keyboard, while clicking all desired sessions (which can stem from different subcorpora).

To perform a search query on the data, right-click on the name of the desired datafile and choose 'annotation content search'. This will direct you to the TLA program Trova. For information on how to use this program, I direct you to the program's manual:

<http://www.mpi.nl/corpus/html/trova/index.html>

A downloadable PDF-version of the manual and more information can be found via the following URL:

<http://tla.mpi.nl/tools/tla-tools/trova/>

In order to be able to perform useful search queries, it is best to get acquainted with the make-up of the annotation of the particular datafiles, e.g. by first browsing them. Detailed information on the annotations can be found in section 3 of this manual.

2.2 Subcorpora

The **NEHOL database** consists of four subcorpora. The classification of a given dataset into one of these subcorpora is determined on the basis of the language variety of Negerhollands used in the data. A full reference list of the sources contained in the database is given in section 2.3.3.

1) Negerhollands acrolectal data

The acrolectal variety of Negerhollands was spoken by a part of the inhabitants of the Virgin Islands of European descent. It was their native language, most likely alongside a European language (such as Dutch or Danish). There is one source in the database that contains data of this variety: **1770 MAGENS**.

2) Negerhollands basilectal data

The basilectal variety is the language as spoken by the slaves in the initial Danish colony and their descendants in the 19th and 20th century. AMAGENS contains non-native Negerhollands, but it is included in this subcorpus because it is clearly the basilectal variety that the author uses. De Josselin de Jong (1926 DJDJ texts) is the most voluminous source of basilectal data and consists of narratives.

3) Negerhollands missionary data

The majority of the data contains the missionary variety of Negerhollands. This subcorpus portrays a broad array of different kinds of sources (e.g. Bible translations, primers, religious literature, mission reports, letters, psalms) that each highlight a different aspect of the Negerhollands language variety of the missionaries. While some of the sources are direct translations of a source document in a European language (German or Danish), other sources show the missionaries construct a new, original text in Negerhollands. The conversations from Oldendorp (**OLDGESPR**) are also included in this subcorpus, because we cannot verify whether they reflect the basilectal variety from the 18th century or whether they have been constructed by missionaries themselves. The same counts for the two short conversations from **Grammatik_conversation**. We do not know how these texts have been constructed, only that the manuscript was written by the Herrnhuter (so German) missionaries. The language used does differ considerably from the language the other missionary documents are written in and seems much more like native Negerhollands, be it acrolectal or basilectal. In order to avoid having to create a fifth subcorpus for sources of which it is not clear what variety of Negerhollands the source represents, these two sources have been included in the subcorpus of missionary Negerhollands, because they were written or at least written down by missionaries. It is up to the user him or herself to judge in what way to use these sources.

4) Slave letters

The slave letters are included in a separate subcorpus, because they cannot be categorised as fitting into one of the other subcorpora. These letters have either been written by slaves - with possibly a few slaves having learnt how to write, who have written down letters others dictated to them – or missionaries writing the letters dictated to them by some of the slaves. It is clear that the variety of Negerhollands used in the letters aims at an acrolectal variety of Negerhollands or sometimes even Dutch. This will be at least partially because it is the missionaries who taught the slaves how to write and they will have aimed for a more Dutch-like standard.

2.3 Sources/Sessions

2.3.1 Annotated sources

The different sources can be found under a header that is called a *session*. The session contains metadata about the source, the author or other contributors, and the content of the source. Each source is represented by either a .tbt file or a .pdf file and an .xml file. Sources with a .tbt file (toolbox text) have been annotated in the free software program Toolbox (<http://www.sil.org/computing/toolbox/>). The annotations are discussed in section 3. The data have not been translated, but when the source text contains a translation, this translation has been included. The selection of sources annotated is intended to be representative of the variety of sources in the NEHOL database. Table 1 shows which sessions have been annotated:

Session	Fully glossed and tagged	Subcorpus
1926 DJDJ texts	yes	basilectal data
PONTOPPI	yes	basilectal data
AUER74	yes	missionary data
AUER84	yes	missionary data
OLDGESPR	yes	missionary data
PONTOPPI_Evangelium	yes	missionary data
Preface Böhner2	yes	missionary data
Zinzendorf Farewell Letter (3.1.1.)	yes	missionary data
Plicht (3.3.5.)	only half	missionary data
Böhner2 (3.2.2.)	partly: 322A1, 322A2, 322B, and 322C.	missionary data
Loretz (3.3.6.)	old glosses from the NWO project	missionary data
Slave letter to the Danish King (3.1.1.)	yes	slave letters
Slave letter to the Danish Queen (3.1.1.)	yes	slave letters

Table 1. Overview of the sources that have been annotated and their location in the corpus.

The annotated file from the session Loretz (3.3.6.) is called EHELEHRE.txt. Because it was annotated in the earlier project, there is no .tbt-file. It contains only one annotation line, a gloss line.

2.3.2 Other sources

Sources that do not have a .tbt file (toolbox text file) have not been annotated. There is only a pdf-file and an xml-file for these sources. The xml-files are fully searchable as well, the search possibilities are only less refined. However, in the case of digitalised manuscripts, the pdf-files contain a special diplomatic editing code system that is absent in the tbt files. The pdf-files can be searched with the simple search function in Adobe Acrobat Reader. The diplomatic editing codes have been entered by Cefas van Rossem and Hein van der Voort to provide information on corrections, revisions, and added footnotes and their location in the original manuscript. The diplomatic editing codes will be explained in section 2.4.3.

2.3.3 Complete overview of NEHOL sources and their references

Below is the complete overview of sources included in NEHOL with the reference to the original source and if relevant a reference to the source the data have been taken from. The sources are listed in the order in which they occur in the database. This is in alphabetical order based on the name of the session that contains the source. The classification in the different subcorpora as discussed in section 2.2 is retained here as well. Headed under the name of the subcorpora are the different sessions with their name in bold. Indented below the session name, the different

files are listed that are included under that session. If needed, information is given on the content of the files. It is suggested to read Van Rossem & van der Voort (1996) as a reference work on the variety of Negerhollands sources and additional information on particular ones.

A general note on the numerical codes that appear in some of the session names: they refer to the code given to the sources by the one who discovered them in the archives of the Moravian Brethren, dr. Peter Stein (see Stein 1986).

Acrolectal data:

1770 MAGENS

- MAGENS1770.pdf

Magens, Jochum Melchior. 1770. *Grammatica over det Creolske sprog, som burges paa de trende Danske Eilande, St. Croix, St. Thomas og St. Jans i America*. Sammenskrevet og opsat af en paa St. Thomas indföd Mand, pp. 34-36, 52-80. Copenhagen: Gerhard Giese Salikrath.

taken from:

Hesseling, Dirk Christiaan. 1905. *Het Negerhollands der Deense Antillen*. Bijdrage tot de geschiedenis der Nederlandse taal, pp. 134-182. Leiden: A.W. Sijthoff.

Basilectal data:

1926 DJDJ texts

- DJDJtekst.tbt

De Josselin de Jong, Jan Petrus Benjamin. 1926. *Het huidige Negerhollandsch: teksten en woordenlijst*, pp. 11-68. Amsterdam: Koninklijke Akademie van Wetenschappen.

- INLEIDING.pdf

The file 'INLEIDING.PDF' is the introduction to the texts. It is in Dutch and contains background information on the different types of texts, the informants, and the spelling convention used plus the pronunciation it is intended to represent.

De Josselin de Jong, Jan Petrus Benjamin. 1926. *Het huidige Negerhollandsch: teksten en woordenlijst*, pp. 5-10. Amsterdam: Koninklijke Akademie van Wetenschappen.

1926 DJDJ wordlist

- DJDJwoordenlijst.tbt

De Josselin de Jong, Jan Petrus Benjamin. 1926. *Het huidige Negerhollandsch: teksten en woordenlijst*, pp. 70-107. Amsterdam: Koninklijke Akademie van Wetenschappen.

AMAGENS

- AMAGENS.pdf

- AMAGENS_NHonly.pdf

The first file contains remarks in German by Schuchardt added to the letter, while the second contains only the letter in Negerhollands.

Schuchardt, Hugo. 1914. Zum Negerholländischen von St. Thomas. In *Tijdschrift voor Nederlandsche Taal- en Letterkunde* 33, 123-135.

Greider

- GREIDER.pdf

Hesseling, Dirk Christiaan. 1905. *Het Negerhollands der Deense Antillen*. Bijdrage tot de geschiedenis der Nederlandse taal, pp. 34. Leiden: A.W. Sijthoff.

Nelson

- NELSON_rawnotes.pdf

Read Den Besten & Van Rossem (to appear) and Van Rossem (to appear) for an extensive discussion on the word list.

Manuscript reference:

Nelson, Frank G. 1936. *Words and short texts in Negerhollands, gathered in St. Thomas, June 1326*.

PONTOPPI

- PONTOPPIDAN.tbt

Pontoppidan, Erik. 1881 Einige Notizen über die Kreolensprache der dänisch-westindischen Inseln. In *Zeitschrift für Ethnologie* 13, pp. 130-138. Berlin.

Missionary data:

AUER74

- AUER74.tbt

Auerbach, Joh. Christoph. [Letter from Niesky in St. Thomas, 10 March 1774 to Friedrich Neisser, Barby, Germany.] 3 pp.

AUER84

- AUER84.tbt

Auerbach, Joh. Christoph. [Catechisatoon van die heilig Doop & Catechisation van die heilig Avendmaal, catechism manually added to the first and last pages of the Moravian psalm book of 1784 by Br. Loretz and Auerbach.] 4 pp.

Barby ABC

- BARBY.ABC_corrected.pdf

A B C-boekje voor die Neger-Kinders na St. Thomas, St. Croix en St. Jan. [Barby, 1800.]. Barby: 1800. 12 pp.

Beilage (3.3.4.)

334.C1.pdf

Beilage zum Diario von St. Thomas vom Monat August 1755 - a) Etliche Cariolische Lieder. 1755. 6 pp.

Böhner1 (3.2.1.)

321A.C1.pdf
321B.C1.pdf
321C.C1.pdf
321D_corrected.pdf
321E.COR.pdf
321Z.COR.pdf

The code 321 is an identification code of the source. The original manuscript is however so long that in the digitalisation process it has been decided to split the source into multiple files. The codes A-E and Z logically indicate the order of the files. This counts for all following sessions containing multiple files with similar codes.

[Böhner, Johann] *Die Geschichte users HEERN und Heilandes Jesu Christi, aus den Vier Evangelisten zusammengezogen, Un aus dem Deutsch in die Creolsprache übersetzt zum Gebrauch in dem Versammlung Haus der Neger Gemeine.* Mit einem Register zur anweisung der Stäglichten Evangelium. Before 1780. 15,0 x 19,0 cm, 406 pp.

Böhner2 (3.2.2)

322A1.C1.pdf
322A1.C1.tbt
322A2.C1.pdf
322A2.C1.tbt
322B.C1.pdf
322B.C1.tbt
322C.C1.pdf
322C.C1.tbt
322D.C1.pdf
322E.C1.pdf

[Böhner, Johann] *Die Handelingen of Geskiedenis van ons HEER en Heiland JEsus Christus ut die Vier Evangelisten na een tesamenhang gefoegt.* Before 1780. 16,1 x 20,1 cm, 4 + 528 pp.

Böhner3 (3.2.4.)

324_corrected.pdf
324_firstpage.pdf
324OP.C1.pdf

[Böhner, J.] *Die Briefen van die Apostel sender.* [and the revelation of John]. 15,6 x 20,1 cm, 259 pp.

Böhner4 (3.2.5.)

325A_corrected.pdf
325B_corrected.pdf
325C_corrected.pdf
325D_corrected.pdf
325E_corrected.pdf
325F_corrected.pdf

[Böhner, J.] *Ein abermaliger Versuch, Etwas aus den Büchern der heiligen Schrift Alten Testaments in die Creol (oder Neger) Sprache zu übersetzen, und in eine harmonische Zeitordnung gebracht, und in den*

dunkeln Stellen aus dem deutlichen Vortrag und Zusammenhang Des Grundtextes erlaeutert End 18th century. 3 volumes, 15,4 x 20,0 cm, 1038 pp.

Böhner5 (3.2.6.)

326A_corrected.pdf

Böhner, Johann. *Korte Begrieb van die Christlike Leer, nabin die evangelische Broedergemeenten daer geleeft van August Gottlieb Spangenberg*. 1780. 2 volumes, 15,5 x 19,3 cm, 650 pp.

Creole sermon1 (3.3.1.5.A)

3315A.C1.pdf

Creole sermon2 (3.3.1.5.B)

3315B.C1.pdf

Creole sermon3 (3.3.1.5.C)

3315C.C1.pdf

[Four Creole sermons from 1796/97]. In: [A number of diverse texts, partially dated]. 46 pp.

Genesis (3.3.1.3.)

3313.C1.pdf

[A free translation of Genesis 1,16 - 4,26]. Some commentaries added. In: [A number of diverse texts, partially dated]. 32 pp.

Geskiedenis, Böhner (3.2.3.1.)

3231.C1.pdf

Die Geskiedenis van ons Heere en Heiland Jesus Christus, opgeskreeven van die vier Evangelisten Matthens, Marcus, Lucas en Johannes; In: [A number of diverse texts, partially dated]. Late 18th or early 19th century. 15,5 x 19,8 cm., 107 pp.

Geskiedenis, unknown (3.2.3.2.A)

3232A.C1.pdf

Die Geskiednis van ons Heere en Heiland Jesus Christus, soo as die vier Evangelist sender ka skriev die op. In: [A number of diverse texts, partially dated]. Late 18th or early 19th century. 15,5 x 19,8 cm., 97 pp.

Grammatik_conversation

Grammatik_conversation.pdf

Grammatik der Creolischen Sprache in West-Indien. 1802.

taken from:

Hesseling, Dirk Christiaan. 1905. *Het Negerhollands der Deense Antillen*. Bijdrage tot de geschiedenis der Nederlandse taal, pp. 183-184. Leiden: A.W. Sijthoff.

HERRN65A

HERRN65A_corrected.pdf

Gebeden en Liederen voor die swart Broeder-Gemeenten na S. Thomas, S. Croix en S. Jan. 1765. unkn. pl., unkn. pub.

Isles (3.3.3.)

333.pdf

[Isles, Samy] *Diarium von St.Thomas vom Monat Septbr. 1754 bis Januar 1755 incl. Beilage zum 25sten Dec. gehörig: Lied welches am Kinder Bettage des 25. Dec [1754] abgesungen worden.* 1754-1755. 4 pp.

Kingo ABB

KINGO.ABB_corrected.pdf

Kingo, Johan Christopher Kørbitz Thomsen 1770. *Kreool A, B, buk. Door J.C. Kingo. St. Thomas na Amerika d. 7. Julii 1770. [Di ka druk na Mester Daniel Thibou na St. Croix, 1770.]* St. Croix: Daniel Thibou. 16 pp.

Loretz (3.3.6.)

EHELEHRE.txt

EHELEHRE.WP5_corrected.pdf

The first file contains annotations from the previous project.

[Loretz, Johann] *Bruder Loretz Entwurf zu einem kurzen Unterricht für die getauften und getrauten Eheleute.* Ca. 1783/1784. 6 pp.

LUND98

LUND98_corrected.pdf

[Lund, T(h)orkild]. 1798. *Leerboek na die Evangelis-kristelik Relisie tot Gebryk voor die Kooninglik Deen Mission-Kerk sender na Amerika.* [Kopenhagen, 1798. Ka prent van Boekprenter K. H. Seidelin]. Copenhagen: K.H. Seidelin.

Magens39

MAGENS39.pdf

Magens, Jochum Melchior. 1770. *Grammatica over det Creolske sprog, som burges paa de trende Danske Eilande, St. Croix, St. Thomas og St. Jans i America.* Sammenskrevet og opsat af en paa St. Thomas indföd Mand, pp. 39. Copenhagen: Gerhard Giese Salikrath.

Memorabilia (3.3.1.7.)

3317.C1.pdf

Memorabilia, in Creole. 1767. In: [A number of diverse texts, partially dated]. 8 pp.

OLDGESPR

OLDGESPR.tbt

OLDGESPR_2.tbt

The second file contains two more, different conversations than the first one.

Oldendorp, C.G.A. *Missionsgeschichte*, 3 vols.

recently published as:

Oldendorp, Christian Georg Andreas. 2002. *Historie der caribischen Inseln Sanct Thomas, Sanct Crux und Sanct Jan : insbesondere der dasigen Neger und der Mission der evangelischen Brüder unter denselben* : kommentierte Ausgabe des vollständigen Manuskriptes aus dem Archiv der Evangelischen Brüder-Unität Herrnhut. Hartmut Beck, Gudrun Meier (eds.). 3 volumes. Berlin : VWB, Verlag für Wissenschaft und Bildung.

The conversations in ‘OLDGESPR_2.tbt’ have been taken from:

Stein, Peter (ed.). 2010. Einige Gespräche, die mündlich gehalten sind, woraus zu ersehen, was manche Schwarze, sonderlich verehrliche, für Streitigkeiten vor ihren Baas bringen, und sie von ihm schlichten lassen. In Gudrun Meier, Peter Stein, Stephan Palmié, Horst Ulbricht (eds.), *Christian Georg Andreas Oldendorp. Historie der caribischen Inseln Sanct Thomas, Sanct Crux und Sanct Jan. Kommentarband*, pp. 252-258. Herrnhut: Herrnhuter Verlag.

Plicht (3.3.5.)

335.C1.pdf

335.C1.tbt

Plicht van Een helper Broeder en Süster. 4 pp.

PONTOPPI_Evangelium

PONTOPPI_Evangelium.tbt

Pontoppidan, Erik. 1881 Einige Notizen über die Kreolensprache der dänisch-westindischen Inseln. In *Zeitschrift für Ethnologie* 13, pp. 130-138. Berlin.

PRAET23

PRAET23_corrected.pdf

Creol Psalm-Buk of een Vergaeding van Oñwe en nywe Psalmen na Creol-Sprack. [Kopenhagen, 1823. Prented bie C. Graebe]. Copenhagen: C. Graebe, 1823. 140 pp.

Preface Böhner2

Preface_322A1.C1.tbt

The pdf-version of this text can be found under: Böhner2 (3.2.2), 322A1.C1.pdf.

[Böhner, Johann] *Die Handelingen of Geschiedenisen van ons HEER en Heiland JESus Christus ut die Vier Evangelisten na een tesamenhang gefoegt*. Before 1780. 16,1 x 20,1 cm, 4 + 528 pp.

PSBUK70

PSBUK70_corrected.pdf

Creool Psalm-buk voor die Deen Missioon na Westindien. 1770. 1770. unkn. aut.

Small Catechism M. Luther (3.3.1.1.)

3311.C1.pdf

*Die kleentje Catechismus van Docter Mart*in Luther**. In: [A number of diverse texts, partially dated]. 28 pp.

Vestindisk Glossarium (KINGOWL.MS)

KINGOWL.MS.pdf

Vestindisk Glossarium. manuscript roughly dates 1770.

published in:

Voort, Hein van der. 1996. Vestindisk Glossarium. In Peter Stein (ed.), *Christian Georg Andreas Oldendorp: Criolisches Wörterbuch, sowie das anonyme J.C. Kingo zugeschriebene Vestindisk Glossarium*, Lexicographica Series Maior 69, 165-187. Tübingen: Niemeyer Verlag.

Wold ABB

WOLD.ABB_corrected.pdf

Wold, Erich Röring. 1770. *Creool A B Buk voor die Deen Missioon na Westindien*. [*ka skrief door Erich Röring Wold, Catechet na St. Jans. 1770.*]; unkn. pl.: unkn. publ. 16 pp.

Zinzendorf Farewell Letter (3.1.1.)

Zinzendorf_Farewell_Letter_311.tbt

Zinzendorf, Nikolaus Ludwig von. 1742. Des Hrn. Grafen Zinzendorff Abschied-Schreiben an die Negers in St. Thomas, in Cariolischer Sprache. In Nikolaus Ludwig von Zinzendorf, *Büdingische Sammlung einiger in die Kirchenhistorie einschlagender, sonderlich neuerer Schrifften. Erster Band*, 453-457. Büdingen: Stöhr. Reprint: Hildesheim: Olms, 1965.

Slave letters:

Slave letter to the Danish King (3.1.1.)

Slave_Letter_to_the_Danish_King_311.tbt

Anonymous. 1739. Der erweckten Negros in St. Thomas Schreiben an Jhro Majest. den König in Dännemarck. In Nikolaus Ludwig von Zinzendorf. 1742. *Büdingische Sammlung einiger in die Kirchenhistorie einschlagender, sonderlich neuerer Schrifften. Erster Band*, 483-485. Büdingen: Stöhr. Reprint: Hildesheim: Olms, 1965.

Slave letter to the Danish Queen (3.1.1.)

Slave_Letter_to_the_Danish_Queen_311.tbt

Anonymous. 1739. Der Ältestin der Gemeinde der Negros in St. Thomas Schreiben an die Königin von Dännemarck. In Nikolaus Ludwig von Zinzendorf. 1742. *Büdingische Sammlung einiger in die Kirchenhistorie einschlagender, sonderlich neuerer Schrifften. Erster Band*, 485-487. Büdingen: Stöhr. Reprint: Hildesheim: Olms, 1965.

Slave letters collection

1738_Peter.pdf

1740_Lenathge.pdf
1752_Catarina.pdf
1752_Domingo_Gesoe.pdf
1753_Cornelius.pdf
1760_Nathanael.pdf
1762_Mari_Magdalene.pdf

Stein, Peter & Beck, Hartmut. In preparation. [Kommentierte Ausgabe der Sklavenbriefe aus St. Thomas 1737-1768].

2.4 Metadata

2.4.1 Metadata files

The metadata files can be viewed by clicking once on the session name. The information shows up in the screen to the right. The metadata start with the following two headers:

- *Session*: containing basic information on the source, such as the name of the session, the title of the source, and the date it was either published or finished.
- *Description*: a short description of the type of source and its content.

Then a number of collapsible headers follow:

- *Location*: an indication of the location the source was written, based on the continent, the country, and a region within this country. The field 'region' is used to indicate which one of the three US Virgin Islands the data come from.
- *Project*: information on the current project NEHOL.
- *Content*: this is one the most important sections of the metadata. The subfields it contains are discussed below.

1. Genre:

Genre is based on whether or not the text is a translation of a text in another language (Translation), or whether it contains spontaneous or original content (Text with original ... data). The latter genre is specified for whether it concerns written or spoken data (Text with original spoken data; Text with original written data). Despite the fact that they contain conversations meant to represent spoken data, the sources **1770 MAGENS**, **Grammatik_conversation** and **PONTOPPI** have been labelled 'Text with original written data', because we cannot verify that they are actual recorded conversations or (re)constructed by the author. The Wordlists are coded as a genre by themselves (Wordlist), because their content is of a different kind and serve a different purpose than actual texts. Sources containing multiple kinds of texts (i.e. those assigned a different genre label) are separated and included as separate sessions, each with their own metadata file. Two of the three wordlists in the database (**1926 DJDJ wordlist** and **1936 Nelson raw notes**) contain example sentences or, in the case of **1936 Nelson raw notes**, even small texts. These parts of the content should be labelled 'Text with original spoken data', but as they are only such a small portion of the whole content and perform a function within the wordlist, these two files are simply labelled 'Wordlist'. The file **Small Catechism M. Luther (3.3.1.1)** is not split into two files, but retained in its original form, and contains both the labels 'Text with original spoken data' and 'Translation', because it ends with a few prayers which are of course translated into Negerhollands.

2. Subgenre:

Subgenre specifies the nature of the text (in some cases specifying the kind of medium the text represents (e.g. letter, textbook, conversation), sometimes based on the type of content of the text (e.g. folktales, catechism, Gospel Harmony). The type of subgenre was chosen for each source in such a way that it would be most informative.

3. Languages.description:

The NEHOL corpus covers a broad range of different kinds of texts, from different time periods, and with different varieties of Negerhollands used: The language variety used in the religious material (e.g. the Gospel Harmony, letters written by missionaries) is

different in certain respects from the basilectal variety, which is used in e.g. **1926 DJDJ texts**. This section indicates what variety of Negerhollands is used in the text, and whether the author of the text or the provider of the data is a native speaker of this variety or not. If the variety is described as ‘Missionary variety of Negerhollands’, this means by definition that the author was not a native speaker of Negerhollands. The native language of the author (Danish, German, Dutch, or English) is given in the Actor section, Languages.

For wordlists, this section indicates what other language is used to translate the Negerhollands content, and what language is used for additional comments.

4. Description:

This section gives a much more detailed and specific indication of the content of the file, what source it is from, the author and all additional relevant information about the content of the file, the language used, and the author’s knowledge and/or command of the language, if available.

- *Actor*: gives information on different kinds of contributors to the source. These can be collectors (those who gather the data from others), consultants (those who contributed the data), authors (those who wrote the data), or translators (those who translated texts into Negerhollands). It is important to mention that the birthdate consists of the order year-month-day.

- *Written Resource*: contains information on the format of the file that contains the data.

2.4.2 References of the metadata

Hesseling, Dirk Christiaan. 1905. *Het Negerhollands der Deense Antillen. Bijdrage tot de geschiedenis der Nederlandse taal*. Leiden: A.W. Sijthoff.

Van Rossem, Cefas & Van der Voort, Hein (eds.). 1996. *Die Creol Tael. 250 years of Negerhollands texts*. Amsterdam: Amsterdam University Press.

Stein, Peter. 1986. Les Premiers créolistes: Les Frères Moraves à St. Thomas au XVIII^e siècle. In Hans den Besten (ed.), *Papers on Negerhollands, the Dutch Creole of the Virgin Islands. Amsterdam Creole Studies IX*, Publikaties van het Instituut voor Algemene Taalwetenschap 51, 3-18.

Stein, Peter. 2010a. Oldendorp und das Kreolische. In Gudrun Meier, Peter Stein, Stephan Palmié, Horst Ulbricht (eds.), *Christian Georg Andreas Oldendorp. Historie der caraibischen Inseln Sanct Thomas, Sanct Crux und Sanct Jan. Kommentarband*, pp. 207-247. Herrnhut: Herrnhuter Verlag.

Stein, Peter (ed.). 2010b. Einige Gespräche, die mündlich gehalten sind, woraus zu ersehen, was manche Schwarze, sonderlich verehrliche, für Streitigkeiten vor ihren Baas bringen, und sie von ihm schlichten lassen. In Gudrun Meier, Peter Stein, Stephan Palmié, Horst Ulbricht (eds.), *Christian Georg Andreas Oldendorp. Historie der caraibischen Inseln Sanct Thomas, Sanct Crux und Sanct Jan. Kommentarband*, pp. 252-258. Herrnhut: Herrnhuter Verlag.

2.4.3 Diplomatic editing codes

The diplomatic editing codes have been chosen and entered by Cefas van Rossem and Hein van der Voort in the NWO project in the first half of the 1990s. This section is therefore based on their written account of the choices made. The purpose of diplomatic editing is to be as complete as possible when presenting a text in providing information on how the text is presented in the original document or manuscript, which particularly includes information on corrections, revisions, and additions.

The diplomatic editing system chosen is an adaptation of Verkruysse (1973-1974). First, a list of functions is given, which can be matched with the list of the corresponding diplomatic editing codes following below that. The codes used before example sentences refer to the texts (before the dot) and the page number (after the dot) the example is taken from. All the examples are taken from Böhner's bible translations. E.g. 322.27 refers to text 3.2.2., to be found as Böhner2 (3.2.2.) in NEHOL, page 27 of this text.

1. Additions

1. a added on the line

<a>

322.27: *va*n die HEere; sender <a> loop weeraan na Ga

2. a added above the line

<↑a> (alt 024)

322.2: dat jender sal krieg <↑die> na die Handen

3. a added below the line

<↓a> (alt 025)

322.34: latstaan doop Em. <↓van em> Maar Johannes

4. a added in the left margin

<←a> (alt 027)

322.40: <←Leer-Baas> waar Joe woon? Em a see na sender:

5. a added in the right margin

<→a> (alt 026)

322.41: en a see na em: Ons ka vind die Messias <→die gesalvde.>

6. a added in the upper margin

<^a> (alt 030)

322.82:

82

op dat sender even sooveel krieg weeran
voor daarom hab jender Vyanden liev;
doe goets en leen, waar jender no verwagt

een goed voor die: dan sal jender Loon
 wees groot, en jender sal wees Kinders
 van die Allerrhohgste. Voordaarom jender
 sall wees volkom, glik jender Vader ben
 volkom ookal die ben nabinn die Hemel;
 en wees Barmhertig, glik jender Vader
 ben Barmhert ookal.
 <▲Matth. 6.>

7. a added in the lower margin

<▼a> (alt 031)
 322.96:

na Jesus <▼en a praat na Em die selvde Woorden, die em a ka lat see na Em
 door si Vrienden> [-*.] en Jesus a see na em:
 loop na hoes, die geskied joe soo as
 joe a ka gloof. En si Dienar *a* kom
 gesond na die selvde Uur. En as die
 ge\stiere a kom weeraan na hoes, jen=
 der a vind die sieke Dienar gesond.
 §.28.

@

8. added but illegible

<*. *>
 322.38: ka stier ons *. * wat joe see van joe selv?

9. a possibly added on the line

*<a> *
 321.17: vier en*<t>*achtig Jaar; die no a kom meer van die Tem=

10. a added at the end of the paragraph or the section

<§a> (alt 021)
 322.18:

en a wees <←[-na]> <§na die> Weldnes, tee em a sall staan
 na voor die Volk van Israel.

§.6.

Luc.

10b. a added at the end of the previous paragraph or the section

<§↑a> (alt 021 en alt 024)
 321.55:

der; d[at+]<ie><↑3> doe<↑2> jender<↑1> na<↑4> sender: Dat ben die Wet en die

Prophten. (sender Leer)

<§↑Ev. 8 Sonnt.p.Trinit.> Loop nabin door die naue Poort. Want die Poort ben wiet,

11. a added on the next page

<➤a> (alt 016)

325a.1:

a kom sonder f*a*soen en leeg, <➤▼Door die v*al*l van die Engel sender ben die verwusting en Duisterniss ontstaan. 2 Petr. 2, v.4. Judag v.6. - Na ons Christin die ben geopenbart, dat die no a wees van Begin; want Godt ben een Vader van die Licht, en nabin en ben geen Duisternis. Jac. 1, v.17. 1 Joh. 1, v.5.> en die a wees dui=

ter op die Afgrond. En die Geest Gods a sway op die Water. En Godt a see: die kom Licht, - en <→die a kom

2. Deletions

1. a deleted

[-a]

322a1.1 (Inleiding): wanneer jender sett nabin jender Hoes sonder, [-sonder] vor wees na Werk

2. deleted and illegible

[-*.*]

322a1.66: [-*.*] En Em a begin vor see na sender:

3. a deleted and b added on the line

[-a]

322a1.40: maar em, die a stier mi vor doop met Water, die sel[-de]<v>= de a see na mi:

4. a deleted and b added below or above the line

[-a]<↑b> of [-a]<↓b>

322a1.24: wat die Heer a see door die Prophet, [-die le]<↑a ka> see, die le s*ee:* Uit van Egypt mi a ka roep mi Soon.

3. Substitution

1. b written over a

[a+]

4. Uncertain readings

1. a is uncertain

a

322a1.2 (Inleiding): En dietoe*n* mi a ka krieg een Boeki

2. b after a is uncertain

a*b*

322a1.3: en die Leefen a wees die Menschen sender

*Lic*ht. 5 En die Licht skien nabin die

D*üste*rnis

3. uncertain whether there is something written

...(?)

322a1.22: a kom van Ost-Land na Jerusalem e*...*

see

4. the whole word is uncertain

woord

5. uncertain whether upper-case or lower-case is intended

A/a

6. uncertain whether a or b is intended

a/b

7. uncertain whether ab and cd written together or separately

ab\cd

5. Other symbols

1. abbreviations

italicise the solution. Example: a line over a letter may mean that that letter should be doubled, or it may mean that an –n or –en should be added.

2. illegible

one dot per letter, ± added when in doubt

3. variant b above a

b|a

4. comment on manuscript by editor

between dollar signs: \$...\$

5. variant b above a is partly identical to a

example: S<[↑]lok>|walg. It should be read as Slok|Swalg, but in the original there is a capital 'S', with 'walg' written behind it and 'lok' written above 'walg'.

6. left column, right column

no sign for the left column, the right column is indicated with an #

7. border of the page

@

3. Annotations

3.1 Introduction

3.1.1 The annotation layers

The tagset or annotations used in NEHOL consist of two layers: one gloss layer informing about the meaning or function of the word, and one tag layer specifying the part of speech. Two examples from ‘1926 DJDJ texts’, (1) and (2) below, best exemplify the different layers that the annotated files in NEHOL contain. The layer identification codes are those used in Toolbox, the free software programme used to annotate the NEHOL texts.

```
(1) \id DJDJtekst
    \ref .001
    \tx I.
    \page 11
```

```
(2) \ref .002
    \tx Di hā   ēn   juñ.
    \mb di   ha   ēn   juñ
    \ge 3SG   have IND boy
    \ps PRON V   DET   NN

    \page 11
```

Every text begins with an \id code giving the name of the file, as can be seen in (1). Every sentence has a unique reference marker \ref, starting from .001 in each file. The actual data is always preceded by the marker \tx. When an entry contains actual analysable data, as in (2), there are three annotation layers, which always come together. When there is a gloss layer, introduced by \ge (gloss English), there is also a tag layer, introduced by \ps (part of speech). There is also a layer of standardised forms to avoid spelling or form variation, introduced by \mb. An example of this can be seen in (2), where *hā* with <*ā*> is standardised to *ha* with <*a*> in the \mb-layer. The standardising always occurs within one source and its only goal is to facilitate search queries. When two or more forms co-occur in the source, in general the most frequent form has been chosen. In the missionary data, the spelling variance is so overwhelming that combined with the big size of the sources choosing a standardised form based on frequency, most resemblance to other similar forms, etc. was simply undoable. Although random, the only feasible solution has been to take as standardised form the form that occurs first.

The original main purpose of the \mb line is to be able to identify and separate morphemes. Since there is practically no morphological marking in the basilectal data, there is little or no need to identify morphemes beyond the word level in Negerhollands. The missionary data do contain morphologically marked forms, such as case marked forms and morphological plurals.

Page numbers (introduced by \page) indicate the location of the data in the original source. When an entry contains data in a language other than Negerhollands, this is indicated in a separate layer, introduced by \lang.

3.1.2 Justification of the classification used for annotating

Defining a set of categories and annotating data for these categories to the satisfaction of every potential user is quite an impossible task. Still, it is important for the usefulness of a corpus that the annotation is usable by as much potential users as possible. In order to allow for encoding

sufficient information, but at the same time warrant easy retrieval of this information, within the NEHOL project we have chosen to use an annotation system of two layers. In deciding what kind of information should go where, we have kept the primary goals of annotations in mind. These are bipartite: on the one hand, annotations make data in languages foreign to the reader accessible and understandable; on the other hand, annotations of word categories enable more complex and refined computer based search queries than just individual lexical items. In our annotations, we want to meet both aims, as they both serve the aim of the NEHOL database, i.e. making the Negerhollands data openly available and searchable. First, it is important to realize that the two purposes of annotation formulated above are not and should not be made incompatible. The information needed to make Negerhollands data understandable should in as much as possible be information usable to perform useful search queries.

The benefit of our choice for having two layers of annotation is best seen in lexical items. One layer serves the purpose of making these items understandable by providing a characterization of the meaning of that item with an appropriate and in meaning corresponding English lexical item if available. An annotation layer used for this practice is commonly labelled a gloss, which is the term adopted here. Lexical items can belong to various word classes, which are relevant to distinguish between. It is important to know whether a word is a noun, a verb, or belongs to another category. Such information is called part of speech, and a layer to provide such information is called a part of speech tag, or simply a tag.

This clearcut difference between gloss and tag becomes somewhat vague when describing functional items, due to the fact that the characterization of the meaning of a functional item is generally best provided through a characterization of its function. The problem is that word classes as we know them all refer to a function of the word in question in some way, in the sense that e.g. a noun (stereotypically) functions as an entity denoting word. A second problem is that one word can have more than one function and the classification of such items all depends on what function is considered to be most relevant to encode. Consider the case of a relative pronoun, is it more important to encode that it introduces a subordinate clause (a relative pronoun as a subordinating conjunction), or is it more important to encode that a relative pronoun functions as a pronominal? Note here that the word class category 'pronoun' is in fact a function denoting term, because it indicates that a pronoun functions as a substitute for a noun or a noun/determiner phrase.

To return to the first problem mentioned, the usefulness of the distinction between the gloss and the tag layer for functional items depends on the information we want to encode and how the categories used relate to other categories, i.e. is it a subcategory or a supercategory of other categories? There are categories, such as 'focus marker' that do not relate to any other categories in this way in Negerhollands. Given that 'focus marker' is a useful but also exhaustively specific characterization of the function of the item, it is consistent to encode this information in the gloss layer. Other categories, such as 'numeral', are more problematic. First, a numeral does relate to other categories we would like to express, such as quantifier - a numeral is a type of quantifier - and determiner - a quantifier is a type of determiner. When we want to use both the term quantifier and the term numeral, we run into problems with having only two layers, since neither category is exhaustively specific: There will have to be a gloss encoding the specific meaning of the item, such as 'three' or 'second' which will be obscured by a less specific gloss as numeral. Likewise, a gloss 'some', 'all', or 'many' is needed in addition to the less specific category 'quantifier'. It is evident that the two layer annotation system cannot maintain absolute consistency in degree of embeddedness between related categories.

It is important to note that the problem of not knowing what category to encode in the tag layer when a category is involved in a hierarchical relation with others cannot be resolved by reserving the tag layer only for word class denoting categories, because "standard" contemporary word classification systems are a mix of both traditional and modern classes with especially the traditional ones, such as 'noun', 'verb', 'adjective', 'pronoun', 'demonstrative', 'article', and

‘numeral’ to name a few, being associated with denoting word classes. It needs no further explanation that these classes are not very helpful in resolving the issues addressed here. It is of course not surprising that trying to fit the most commonly used grammatical categories into a two layer system cannot be done with absolute consistency for the simple reason that the most common classifications were not made to fit into such a structure.

Thus, forcibly a compromise needs to be reached between consistency in encoding the degree of embeddedness between related categories and between optimally informative annotations. We have chosen to give preference to the latter, resulting in the situation that quantifiers and numerals have their own tag, even though they are subcategories of determiners, a category that is also used as a tag for other items. We think that this more informative, albeit less consistent classification is preferable, because it has the advantage of making the NEHOL annotations more useful and better interpretable.

Another argument is the fact that the categories desired for annotation searches and the level of detail may vary from user to user. It is therefore better to provide a greater level of detail in the tags: It is possible to search for multiple categories at the same time if one is interested only in the supercategory not annotated, but it is obviously not possible to search for subcategories not annotated in the database.

Conversely, the desire to have the gloss function as a unique identifier that is context independent results in the situation that tags are not always a precise indication of the function or word class of each occurrence of an item. The underlying rationale is that the annotation should not or at least as little as possible be an analysis of the data. It is up to the user to analyze the data. Whenever there is a distinction between certain categories not reflected in the annotation, this is explicitly mentioned and exemplified in the following sections. First the open class categories are discussed in section 3.2, the closed class or functional categories are discussed in section 3.3. Additional categories are discussed in section 3.4. Subsections are named after the categories and list the tags and glosses used for items belonging to these categories.

3.2 Open class categories

3.2.1 Nouns

Tags: NN = common noun
 NP = proper noun
 NT = title noun

Negerhollands nouns are morphologically invariant for number, so there is no special distinction between count nouns and mass or non-count nouns, they are all tagged NN. Because nouns are lexical items with an indefinite number of meanings, the glosses are lexical as well.

The varieties other than the basilectal one do contain morphological plural and case marking and morphological inflection on adjectives. The corresponding tags and glosses to these morphemes are discussed in the subsection 3.3.6 on inflection.

In all varieties of Negerhollands, nouns can be pluralised by postponing the third person plural pronoun, which is glossed as such - 3PL – in order to avoid masking the formal identity between the third person plural and the nominal plural marker.

Abbreviations have a special coding at the end of the tag: .ABBR (see section 3.4.1).

The tag NP is assigned to proper nouns, which are names. These can be names of a specific individual, geographic location, country, etc., etc. Names of the days of the week and months are treated as common nouns.

The tag NT is given to title nouns, which are nouns of address. Common title nouns are *Bru* ‘Brother’ and *Baas* ‘Master’ (used to address the missionary ministers in the 18th century). The example below contains all three noun types. It also shows that compounds, such as *kabritafleis* in example with ref. nr. 1325, consisting of *kabrita* ‘goat’ and *fleis* ‘meat’, are generally treated as one noun.

```
\ref .1325
\tx Bru    Anānši a    ha    si       kabritafleis.
\mb Bru    Anānši a    ha    ši       kabritafleis
\ge Brother Anansi PST have 3.POSS goat.meat
\ps NT    NP    TAM V    DET    NN
“Brother Anansi had got his goat meat.” DJDJ 1926:27
```

3.2.2 Verbs

Basic tag: V = common verb

The tag V is attributed to verbs. We do not distinguish between auxiliaries and non-auxiliary verbs in the sense that an auxiliary is a verb syntactically accompanying another verb (Matthews 1997: 31). Specific functions which are often expressed in languages through auxiliaries, that is copulas and TAM-markers, including modals are treated as separate categories in their own right. This has the advantage of avoiding the discussion whether the elements that function as a copula or a TAM-marker are in fact verbal elements or not in Negerhollands. Read section 3.3.3 on copulas and section 3.3.13 on TMA markers, including modals.

Many creole languages share the feature that an adjective expressing a certain state or property behaves like a verb when used predicatively. This feature is discussed in the next subsection on adjectives.

3.2.3 Adjectives

Basic tag: PI = property item

It is a common feature in creole languages that predicatively used adjectives behave like a verb. We encounter predicative adjectives behaving like a verb in Negerhollands as well. The properties they share with verbs are i) they can combine with preverbal TMA marking, and ii) some even take objects, as in example ref. 2529. A comparison between example ref. 2529 and example ref. 1913 shows that the same item can occur both with verbal properties (ref 2529) and as an adjective (ref 1913). In order to be as consistent as possible in attributing the same item the same gloss, we have tagged predicative adjectives as ‘PI’, i.e. a property item, whether they obviously have verbal properties or not. Because the exact status of these items in Negerhollands is not clear, we have chosen to use the word class neutral term ‘property item’. By applying this tag, we do not want to imply in the least that Negerhollands does not have adjectives.

\ref.2529
\tx Bru Lion bañ am oka.
\mb bru lion bañ am oka
\ge brother lion afraid 3SG.AN too
\ps NT NN PI PRON ADV
“Brother Lion was afraid of him too.” DJDJ 1926:50

\ref.1913
\tx [...] di gut wa lō mā am bañ [...].
\mb [...] di gut wa lō mā am bañ [...].
\ge [...] DEF thing REL LO.IPFV make 3SG.AN afraid [...].
\ps [...] DET NN CS TAM V PRON PI [...].
“... the thing that was scaring him ...” DJDJ 1926:37

In order to avoid obscuring the formal identity between predicative and attributive adjectives and at the same time being able to avoid having to force a distinction between non-verbal and verbal predicative adjectives when such a distinction is not at all clear, we have decided to assign both attributive and predicative adjectives the tag ‘PI’. Similarly, there is no formal difference between adjectives and adverbially used adjectives, so both are assigned the tag ‘PI’.

Certain Negerhollands texts contain inflection on attributively used adjectives. This inflection is discussed in section 3.3.6.

3.2.4 Adverbs

Basic tag: ADV = adverb
Gloss: MOD = modal adverb

Many items that occasionally function as an adverb belong to other word classes as well. In order to be consistent in annotating items, only those items that are adverbial in nature are assigned the tag ‘ADV’. Adverbially used adjectives are tagged ‘PI’ (see section 3.2.3), and adverbially used prepositions are tagged ‘PREP’ (see section 3.3.1).

The item *wel* corresponds in meaning to the Dutch modal adverb *wel*, for which there is no English equivalent. Because its function may vary from context to context, it is hard to grasp it in a lexical representation, therefore it is glossed as ‘MOD’. Its use is exemplified in the typical example below, where it marks that the constituent or constituents that is/are new in the sentence given is/are a (possibly forced) alternative to a previously mentioned situation.

\tx Mi doop jender wel met Water tot bekeering;
\mb mi doop jender wel met water tot bekering
\ge 1SG baptise 2PL MOD with water (up)to conversion
\ps PRON V PRON ADV PREP NN PREP NN

“I will convert you by baptising you with water.” 322A1:34

Negerhollands has serial verb constructions where multiple verbs can be piled upon each other. This phenomenon has resulted in the situation that in Negerhollands, verbs as *ko* ‘come’ and *lo* ‘go’ can be used in a way that corresponds to the use of the English adverbials ‘here’ and ‘away’ respectively. An example of this can be seen in the example ref. 490 below, where the sentence

brin di difman ko is to be translated as ‘bring the thieves here’. In order to avoid forcing an interpretation on these items, *ko* and *lo* are simply tagged as verbs, ‘V’.

\ref.490

```
\tx Də polisman a skrēw: brin di difman ko.
\mb di polisman a skrēw brin di difman ko
\ge DEF policeman PST yell bring DEF thief come
\ps DET NN TAM V V DET NN V
```

“The policeman yelled: Bring th(os)e thieves here.” DJDJ 1926:16

3.2.4.1 Intensifier

There is a particular class of adverbs that describe the intensity of the meaning of a word, either in positive or negative sense. These items are called intensifiers. Because there are no intensifiers in the data that do not belong to any other class, we have no special annotation for intensifiers. Most intensifiers alternatively function as quantifiers: *atəvəl* ‘too’ versus ‘too much’, *muši* ‘very’ versus ‘much/many’, *bitsi* ‘a bit’. Below is an example of the item *frāi* ‘good’ used as an intensifier, tagged PI, plus an example of *muši* used as an intensifier, glossed as a quantifier.

\ref.1710

```
\tx Jin juñ mēnši mi frāi foflúk
\mb jini juñ mēnši mi frāi foflúk
\ge 2PL young girl be good sly
\ps PRON PI NN COP PI PI
```

```
\tx an ons hou man ha fo bi muši patiklā;
\mb en ons hou man ha fo bē muši patiklā
\ge and 1PL old man have FO be much/many careful
\ps CC PRON PI NN V CS COP QNT PI
```

‘You young girls are very sly and we old men have to be very careful.’ DJDJ 1926:34

The intensifier *gənú(g)* ‘enough’ – occurs as *genoeg* ‘enough’ in the missionary data – is tagged ADV.

```
\tx Die ben genoeg, dat
\mb die ben genoeg dat
\ge PRO.3 be enough that
\ps PRON COP ADV CS
‘It is enough, that ...’ 322A1:88
```

\ref.1979

```
\tx Weni am a ha guñgu gənú tou,
\mb weni am a ha guñgu gənú tou
\ge when 3SG.AN PST have big enough rope
```

\ps CS PRON TAM V PI ADV NN
 ‘When he had a big enough rope, ...’ DJDJ 1926:39

The item *so* has a broad range of functions. It often occurs in sentence initial position, where it is unclear whether it is an adverb or a coordinating conjunction. It also functions as an intensifier, as in the below example. To cover the broad range of uses, *so* is always annotated as an adverb, thus tagged ADV.

\ref .082
 \tx Am a wēs so lilik.
 \mb am a wēs so lelik
 \ge 3SG.AN PST be so bad
 \ps PRON TAM COP ADV PI
 ‘He was so evil.’ DJDJ 1926:12

3.2.5 Reduplication

Tag: .REDUP= reduplication

Reduplication is not a productive phenomenon in Negerhollands. Words of which the stem consists of a reduplicated part, such as *patpat* ‘duck’ are treated as consisting of one unanalyzable stem. The only exception is in the case of reduplicated adverbs or property items, where the reduplication has an intensifying function, as in the example below. In these cases, the reduplicated item has the element .REDUP added to the tag.

\ref .2967
 \tx [...] am a lo lo sudži sudži me am.
 \mb [...] am a lō lō sudži sudži mi am
 \ge [...] 3SG.AN PST go go slowly slowly with 3SG
 \ps [...] PRON TAM V V ADV ADV.REDUP PREP PRON
 ‘he walked very slowly with him.’ DJDJ 1926:59

3.2.6 Interjection

Tag: IJ = interjection

Interjections have been defined as “[f]orms that express ‘states of mind’ and do not enter into specific syntactic relations with other words” (Matthews 1997: 182). Most common are the interjections *jā* ‘yes’ and *nēn* ‘no’. Another example is in example with ref. 914 below.

\ref .914
 \tx weni ju mañkě am stop, ju fo sē: „dži”!
 \mb weni ju mañkě am stop ju fo sē dži
 \ge when 2SG want 3SG.AN stop 2SG FO say whoa
 \ps CS PRON TAM PRON V PRON CS V IJ
 “When you want him to stop, you have to say ‘whoa!’” DJDJ 1926:26

3.2.7 Ideophone

Basic tag: ID= ideophone

The following cross-linguistically valid definition of ideophones has been formulated: “Ideophones are marked words that depict sensory imagery” (Dingemanse 2011:25). The term ‘marked’ means that ideophones “stand out from other words” (Dingemanse 2011:25).

Onomatopoeic descriptions of sounds, including sounds meant to represent motion or any other “sensory imagery” have been labelled ideophone in Negerhollands. Two examples are given below. The second example shows that ideophones can also occur as verbs.

\ref.1272

\tx	Dan	Tekoma	a	sē:	awi	a	sē,	dōt	man	kan	mā	pûûûp?
\mb	dan	Tekoma	a	sē	awidi	a	sē	dōt	man	kan	mā	pûûûp
\ge	then	Ntikuma	PST	say	who	PST	say	dead	man	POSS	make	poo
\ps	ADV	NP	TAM	V	Q	TAM	V	PI	NN	TMA	V	ID

‘Then Ntikuma said: Who said that the dead can fart?’ DJDJ 1926:26

\ref.1452

\tx	Wani	am	a	lō	a	di	kas	hopo	di,
\mb	weni	am	a	lō	a	di	kās	hopo	di
\ge	when	3SG.AN	PST	go	NA	DEF	cupboard	open	3SG
\ps	CS	PRON	TAM	V	PREP	DET	NN	PI	PRON

na ha ēn steki kās, brōt, rezə;

na ha ēn steki kās brōt rezə

not have one piece cheese bread lard

NEG V N.CAR NN NN NN NN

‘When he walked to the cupboard and opened it, there was no cheese, bread, or lard at all.’

ēkēgut a lō fɟup.

ēkēgut a lō fɟup

everything PST go disappear

IPR TAM V ID

‘Everything was gone.’ DJDJ 1926:29

3.3 Closed class categories

3.3.1 Adpositions

Basic tag: PREP = preposition
 V.PREP = preposition as part of verb

Negerhollands has only prepositions, no postpositions. Prepositions used adverbially are also tagged as prepositions, because it is evident from the context alone when this is the case plus that the meaning (and therefore the gloss) is generally the same in both cases. An exception is *op* which is tagged as PREP and glossed ‘on’, when it occurs as a preposition, but *op* is tagged ADV and glossed ‘up’ when it occurs adverbially. This is because the meaning of *op* as a preposition and its meaning in interaction with other words cannot be recovered from these English glosses and would only be clear for people who have knowledge of Dutch. Compare the following two examples. If we *op* were glossed as ‘on’ in the second example, the meaning of the verb *lecht op* ‘lift up’ would not be clear from the glosses ‘lift’ and ‘on’ respectively.

```
\tx Die geskiedenis van ons HEer en Heiland op Aerden.  
\mb die geskiedenis van ons heer en heiland op aerde  
\ge DEF history of/from 1PL Lord and saviour on earth  
\ps DET NN PREP PRON NN CC NN PREP NN  
‘The history of our Lord and Saviour on earth.’ 322A1:16-17
```

```
\tx lecht op jender Oogen  
\mb lecht op jender oogen  
\ge lift up 2PL eyes  
\ps V ADV PRON NN.PL  
‘lift up your eyes’ 322A1:56
```

There are items that are formally similar to prepositions that only occur adverbially, such as *it* ‘out’ and *abiti* ‘outside’. These are always tagged ‘ADV’ in the basilectal data. The item *it* does combine with the preposition *fa* ‘of/from’ to form prepositional compounds, similar to English ‘out of’, as can be seen in the example below.

```
\ref .1245  
\tx [...] ons a ha fu kurí abít it fa di kérék.  
\mb [...] ons a ha fo kurí abiti it fan di kérék  
\ge [...] 1PL PST have FO run outside out of/from DEF church  
\ps [...] PRON TAM V CS V ADV ADV PREP DET NN  
‘We had to run outside out of the church.’ DJDJ 1926:26
```

Note that the related item *abini* is tagged as ‘PREP’, because it does occur as a preposition. The missionary data have the item *ut* ‘out.of’, which is a preposition and tagged as such.

Prepositions can also function as subordinating conjunctions. Also here, the same items are annotated the same. There are only two prepositions that are tagged as a subordinating conjunction: *fo* because its meanings and functions are so diverse (see section 3.3.2). Note that *fo* is tagged PREP and glossed ‘for/before’, when it functions as either a temporal or a benefactive preposition. The temporal and locational uses of the preposition *tē* are illustrated in the examples ref. 2789 and ref. 153 respectively. Exceptionally, the preposition *tē* is always tagged as a temporal

subordinating conjunction (‘CS’), for the simple reason that prepositional occurrences of *tē* are rather low frequent among the abundance of occurrences of *tē* as a subordinating conjunction.

\ref.2798

\tx Sinu a dig di pit ēn wēk fa Mānda tē Sātáda.
\mb sinu a dig di pit ēn wēk fan mānda tē Sātáda
\ge 3PL PST dig DEF well IND week of/from Monday until Saturday
\ps PRON TAM V DET NN DET NN PREP NN CS NN
‘They dug the well for a week from Monday to Saturday.’ DJDJ 1926:56

\ref.153

\tx Am a lō tē a di klip.
\mb am a lō tē a di klip
\ge 3SG PST go until NA DET cliff
\ps PRON TAM V CS PREP DET NN
‘He walked upto the cliff.’ DJDJ 1926:12

As in Dutch, prepositions can be a separable part of verbs. The separable prepositional part may follow the complement, when a complement is present. Thus, in such a case the verbal and the prepositional part are not adjacent. To distinguish the verbal part from a homomorphous verb with a different meaning, the meaning is glossed with the verbal part while the prepositional part is tagged ánd glossed V.PREP, because the meaning is already expressed in the verbal part. The first versus the second and the third example illustrate that verbs with a prepositional part are completely different lexical items than the formally identical verbs without prepositional part, and thus the two may need to be distinguished from each other by having separate glosses.

\tx joe sal bed Em,
\mb joe sal bed em
\ge 2SG FUT pray/ask 3SG
\ps PRON TAM V PRON
‘you will ask him’ 322A1:53

\tx Maar Ons weet, wat ons bed an;
\mb maar ons weet wat ons aanbed
\ge but 1PL know Q.INAN 1PL worship
\ps CC PRON V Q PRON V
‘But we know what we worship.’ 322A1:54

\tx em a bed Em aan.
\mb em a aanbed em aan
\ge 3SG PST worship 3SG V.PREP
\ps PRON TAM V PRON V.PREP
‘he worshipped him.’ 322A2:194

Note that in order to limit the workload this is only done for verbs with a prepositional part where the meaning of the whole is considerably different from the sum of the meaning of the bare verb plus the meaning of the preposition itself or where this meaning is not recoverable from the English glosses of the verb and the preposition. In the basilectal data, this is always the case and the tag V.PREP is not used. Below is an example from the missionary data, which shows that the (reflexive) verb *draai si om* ‘turn around’ can be constructed from the glosses of the individual items the verb *draai* ‘(re)turn’ and the preposition *om* ‘round’.

\tx Maar Jesus a draai si om,
 \mb maar Jesus a draai si om
 \ge but Jesus PST (re)turn 3SG.RX around
 \ps CC NP TAM V PRON PREP
 'But Jesus turned round' 322A1:40

Finally, items used in comparisons are all tagged as prepositions. There are two prepositions used in comparisons of equality:

1) *leik(i)*

\ref .894
 \tx grōt liki ši kop.
 \mb grōt leik ši kop
 \ge big like 3.POSS head
 \ps PI PREP PRON NN
 'big as his head.' DJDJ 1926:21

2) *as*

\ref .2527
 \tx Bru Kabritabok a wes džis so wis a am.
 \mb bru kabritabok a wēs džis so wis as am
 \ge brother billy.goat PST be just so wise as 3SG
 \ps NT NN TAM COP ADV ADV PI PREP PRON
 'Brother Billy Goat was just as clever as he.' DJDJ 1926:50

The preposition used in the comparison of inequality is *a*:

\ref .2534
 \tx mi kan slā di tamə́rín da betu a ju!
 \mb mi kan slā di tamarín dā bētə́r a ju
 \ge 1SG POSS hit DEF tambourine there better than 2SG
 \ps PRON TAM V DET NN ADV PI PREP PRON
 'I can hit that tambourine better than you!' DJDJ 1926:50

3.3.2 Conjunctions

3.3.2.1 Subordinating conjunctions

Basic tags: CS = subordinating conjunction
 Gloss: FOR = subordinating conjunction 'voor' in missionary data
 FO = subordinating conjunction 'fo' in basilectal data
 INF = infinitival complementizer
 REL = relative clause complementizer

The tag CS is assigned to subordinating conjunctions. The gloss FOR is assigned to the form 'for/voor' in the missionary data and FO to 'fo/fu' in the case of the DJDJ texts to capture all the various uses of this form. It can be used as:

i) the complementizer of a purposive clause

\ref. 2776

\tx am a bañ, Anānši mata am fo jet.
\mb am a bañ Anānši mata am fo jet
\ge 3SG PST afraid Anansi kill 3SG FO eat
\ps PRON TAM PI NP V PRON CS V
'He was afraid that Anansi would kill him to eat him.' DJDJ 1926:55

ii) complementizer of a compliment verb

\ref. 2261

\tx am a kan fo bli da stān.
\mb am a kan fo bli dā stan
\ge 3SG PST POSS FO stay there stand
\ps PRON TAM TAM CS V ADV V
'she was able to stay upright there.' DJDJ 1926:45

iii) a necessity modal.

\ref. 2272

\tx Am a sē, nēn, am nu fo mā grās.
\mb am a sē nēn am na fo mā grās
\ge 3SG.AN PST say no 3SG.AN not FO make fuss
\ps PRON tAM V IJ PRON NEG CS V NN
'He said, no, he must not make a fuss.' DJDJ 1926:45

Although the bare necessity modal *fo* is functionally a different item than the complex necessity modal *ba fo* (Van Sluijs 2011), the element *fo* receives the same annotation in both cases in order to avoid forcing an analysis on the data, as shown in the following example.

\ref. 2338

\tx Tekoma ha fo gi am.
\mb Tekoma ha fo gi am
\ge Ntikuma have FO give 3SG.AN
\ps NP V CS V PRON
'Ntikuma has to give [them to] him.' DJDJ 1926:46

Thus, in the basilectal data *fo* is glossed 'FO', when it is used in the above mentioned contexts. However, *fo* also occurs as a subordinating conjunction of reason or cause glossed 'because', and as a temporal subordinating conjunction glossed 'before'. When *fo* is a preposition it is tagged as such with the tag 'PREP'.

Fo is exceptional in receiving different tags depending on its function of a subordinating conjunction as opposed to that of a preposition. The item *tē* can also be used as a preposition, but it occurs so overwhelmingly frequently as a temporal subordinating conjunction that *tē* is always tagged CS. Other prepositions used as a subordinating conjunction (e.g. the temporal conjunction *astu* 'after') are always tagged PREP.

Another preposition also used as a subordinating conjunction is the item *leik(i)* (variably occurs as *leik*, *leiki*, *liki* etc) 'as/like'. An example of *leik(i)* as the subordinating conjunction in the same sense as the preposition is given below. Therefore, *leik(i)* is tagged as a PREP in those contexts.

\ref.2095

\tx So džis leiki am prāt də word, a so di kom it.
 \mb so džis leik am prāt di word a so di ko it
 \ge so just like 3SG.AN talk DET word FOC so 3SG come out
 \ps ADV ADV PREP PRON V DET NN COP ADV PRON V ADV
 ‘So just like she said, so it became true.’ DJDJ 1926:41

Leik(i) as a subordinating conjunction can also have the sense of ‘as if’.

\ref.100
 \tx Də mēnši a ki a ēn man
 \mb di mēnši a ki a ēn man
 \ge DEF girl PST see/look NA INDF man
 \ps DET NN TAM V PREP DET NN
 \tx wa glik leiki am a wēs ēn haləf kiniñ.
 \mb wa glik leik am a wēs ēn haləf kiniñ
 \ge REL look.like like 3SG.AN PST be INDF half king
 \ps CS V PREP PRON TAM COP DET PI NN
 ‘The girl looked at a man, who looked like he was almost a king.’ DJDJ 1926:12

The same counts for the preposition *a* used in a comparison of inequality, which can also occur as a subordinating conjunction:

\ref.1000
 \tx Am sa kri me jit as am mañkē.
 \mb am sa kri mē jet as am mañkē
 \ge 3SG.AN FUT get more food than 3SG.AN want
 \ps PRON TAM V QI NN PREP PRON TAM
 ‘She will get more food than she needs.’ DJDJ 1926:22

Leik(i) also occurs as a temporal subordinate conjunction, meaning ‘as soon as’. In this function, it is tagged CS and glossed as.soon.as.

\ref.2384
 \tx So leiki di klēn juñ bigin fo siñ
 \mb so leik di klēn juñ bigin fo siñ
 \ge so as.soon.as DEF small boy begin FO sing
 \ps ADV CS DET PI NN V CS V
 ‘So as soon as the little boy started to sing...’ DJDJ 1926:47

In the missionary data, there is one occurrence (see example below) of *toe* as an infinitival complementizer, as such glossed as ‘INF’. It seems to be in variation with *voor*, which can also function as an infinitival complementizer or a complementizer of a complement verb in the missionary data.

\tx Wagoed ons hab toe doen met joe,
 \mb wagoed ons hab toe doe met joe
 \ge what 1PL have INF do with 2SG
 \ps Q PRON V CS V PREP PRON
 ‘What do we have to do with you [...]?’ 322A1:103.

3.3.2.2 Relative pronouns

The relative pronoun *wa* is tagged CS and glossed REL as a subordinate conjunction of relative clauses in the basilectal data.

\ref .155

\tx	də	seləf	frou	wa	am	a	ki	a	ši	drom.
\mb	di	seləf	frou	wa	am	a	ki	a	ši	drom
\ge	the.same	woman	REL	3SG.AN	PST	see/look	NA	3.POSS	dream	
\ps	DET	NN	CS	PRON	TAM	V	PREP	PRON	NN	

‘the same woman that he had seen in his dream.’ DJDJ 1926:12

The relative pronoun *die* in the missionary data is tagged DET and glossed DEF, as a definite determiner, because it is formally identical to the definite article.

\tx	En	Em	a	kik	si	om	na	die,
\mb	en	em	a	kik	si	om	na	die
\ge	and	3SG	PST	see	3.POSS	around	NA	PRO.3
\ps	CC	PRON	TAM	V	PRON	PREP	PREP	PRON

\tx	die	a	ka	doe	die.
\mb	die	a	ka	doe	die
\ge	DEF	PST	COMPL	do	PRO.3
\ps	DET	TAM	TAM	V	PRON

‘And he turned round to look at the one, who had done it.’ 322A2:112.

3.3.2.3 Coordinating conjunction

Basic tags: CC = coordinating conjunction

The coordinating conjunctions are *en* ‘and’, *o* ‘or’, and *bot* ‘but’. Besides *en*, Negerhollands uses the preposition *mi* ‘with’ to conjoin two NPs. By annotating *mi* as PREP also when it is used to conjoin two NPs we do not obscure the nature of the different strategies that Negerhollands employs to conjoin NPs.

\ref .2867

\tx	am	me	Anānši	alē	lo	tre	di	batita,
\mb	am	mi	Anānši	alēn	lō	trē	di	batita
\ge	3SG	with	Anansi	alone	LO.FUT	pull(out)	DEF	type.of.tuber
\ps	PRON	PREP	NP	ADV	TAM	V	DET	NN

‘only he and Anansi will pull the tubers’ DJDJ 1926:57

3.3.3 Copula

Basic tag: COP = copula
 Gloss: FOC = focus marker
 be
 become

Negerhollands has a rather complex system of copulas. In the basilectal data, the following forms can be distinguished: *nēs*, *a*, *mi*, *bin*, *bi*, *bē*, and Ø. The zero copula has not been annotated. Just as West-African languages, creole languages may apply up to three or even four different strategies

in expressing copula relations that are expressed by only copular verb in most European languages (see Holm 1988:175). Negerhollands is a good example of this, as it possesses about seven items - depending on what forms to count as separate items and what forms as morphological variants, this number may be higher or lower – of which most are functionally specialised. The four functional distinctions are i) focus marking, ii) equation or indicating identity, iii) attributing a property, and iv) indicating a location. In the case of equation, the copular item links the subject with an NP that indicates the subject's identity. In the case of property attribution, the copular item links the subject with an adjective. The copular items links the subject with a locative item to indicate its location. The different functions of each item are shown in the table below, based on the findings in Stolz (1986:152) and Sabino (1988:204) combined. This shows that but for one, all items have multiple functions. This makes it hard to attribute the various items function specific glosses. For this reason, all items in the table below have been tagged COP and glossed 'be' as a function neutral term. The different copular functions can be seen in the table below.

The one exception is the item *a*, which has been glossed FOC, when it functions as a focus marker, because *a* is the only item that performs this function in basilectal Negerhollands. When *a* links an identity NP or an adjective, it is glossed 'be' just like the other copular items. Although this results in the situation that *a* is glossed differently when it is a focus marker than when it links an identity NP or an adjective, we do not want to imply that we are dealing with two separate items. In the missionary data, there is a separate focus marker *da* that does not function as a copula (see 3.3.5).

	Identity NP	Adjective	Location	Focus
<i>a</i>	++	+	-	+
<i>wēs</i>	+	+	+	-
<i>bi</i>	-	+	++	-
<i>bē</i>	+	-	+	-
<i>bin</i>	-	-	++	-
<i>mi</i>	±	++	+	-
<i>Ø</i>	+	+	+ / ±	-

++ = function of the majority of occurrences of this item

+ = a function of this item

± = a marginal function of this item

- = function not attested for this item

A particular distinctive feature of *wēs* is that only *wēs* is compatible with preverbal TMA markers (see 3.3.13), although marginally there is an exception to be found of other copular items having preverbal TMA markers.

Besides the so called copulas of state discussed before, Negerhollands also has copulas of change of state. These copulas mark a transition into a state and are assigned the gloss 'become'. The basilectal variant contains the item *kə*, while the missionary data have both the item *kə* and *wə*.

Basilectal Negerhollands does have a few lexical verbs that also function as a copula of change of state: *kəri* 'get/receive' that combines with adjectives, and *drāi* 'turn' that combines with NPs of (coming) identity (ref. 108) and adjectives (ref. 2706). Both the verb *drāi* and the verb *kəri* are not tagged as copulas, but as verbs (V), because the copular interpretation can easily be attributed to or be directly derived from their lexical meaning. In the case of *drāi*, example ref. 271 below shows this meaning to be present as well in contexts where a copula interpretation is not possible.

\ref .108

```
\tx Am      a      lō      drāi      matrōs,  
\mb am      a      lō      drāi      matrōs  
\ge 3SG.AN PST LO.FUT (re)turn sailor  
\ps PRON TAM TAM      V      NN  
'He was going to become a sailor.' DJDJ 1926:12
```

\ref .2706

```
\tx Bju,     hogo drāi      joia.  
\mb bju     hogo drāi      joia  
\ge brother eye (re)turn red  
\ps NT      NN V      PI  
'Brother, my eyes are turning red.' DJDJ 1926:54
```

\ref .271

```
\tx am      a      drāi      am      ēn      guñgu stēn.  
\mb am      a      drāi      am      ēn      guñgu stēn  
\ge 3SG.AN PST (re)turn 3SG.AN INDF big stone  
\ps PRON TAM V      PRON DET PI      NN  
'He turned him into a big stone.' DJDJ 1926:14
```

The missionary variety of Negerhollands uses the copulas as passive auxiliaries too. In these constructions we have chosen to regard the use of the passive auxiliary as an extension of use of the auxiliary as a copula, which from a diachronic perspective is how the construction developed in Dutch as well (Cornelis & Verhagen 1995). Therefore, the passive auxiliaries are tagged and glossed the same as the formally identical copulas: *wees* and *bin* have been tagged COP and glossed 'be', while *wort* and *kom* have been tagged COP and glossed 'become'. The following example shows both uses of the item *kom*.

```
\tx Em      sal      kom      groot      en      sal      kom  
\mb em      sal      kom      groot      en      sal      kom  
\ge 3SG      FUT become big/grand and FUT become  
\ps PRON TAM COP PI      CC TAM COP
```

```
\tx Geroepen      een      @  
\mb ge- roep -en      een      @ {page.break} {PU}  
\ge PTCP- call -PL INDF page.break  
\ps INFL- V -INFLDET PU
```

```
\tx een      Soon van      die      Hooghste;  
\mb een      soon van      die      hogh -ste  
\ge INDF son of/from DEF high -SUP  
\ps DET NN PREP DET PI -INFL  
'He will become grand and will be called a son of the Highest.' 322A1:11-12
```

3.3.4 Determiners

Basic tag:	DET	= determiner
Gloss:	INDF	= indefinite
	DEF	= definite
	DEM	= demonstrative
	other	

such

The class of determiners is a functionally motivated one that combines grammatical elements that limit the “potential referent of a noun phrase” (Matthews 1997:95). This encompasses items such as articles, attributive demonstratives, and quantifiers, which in turn is a supercategory encompassing cardinal numerals. Quantifiers and numerals are tagged differently than other determiners, because they form a subcategory that is useful to distinguish (see sections ???).

Formally, there is no distinction between definite articles and attributive demonstratives in basilectal Negerhollands. Only one form occurs: *di*, which also occurs with reduced vowel as *də*. Therefore, the determiner *di* is glossed DEF, for having definite reference. A distinction between proximal and distal referents can be expressed by additionally using the locational adverbs *bi(so)* ‘here’ and *da(so)* ‘there’. That these combinations cannot be said to form conventionalised periphrastic demonstratives follows from the facts that i) they do not seem obligatory; and ii) there is no fixed position for the locational adverbs, when they occur, to combine with the determiner *di*.

The missionary data do contain one item other than *die* that functions as an attributive demonstrative. This is *deese*, which has proximal reference. Therefore, it is glossed DEM.PROX. *Deese* also occurs as a demonstrative pronoun, but it is not separately annotated for this function: therefore the form *deese* is always assigned the gloss DEM.PROX and the tag DET.

An item that may arguably function as an attributive demonstrative with distal reference is *die* which is identical in form to the definite article *die* and thus always glossed DEF. *Die* also occurs as a relative pronoun in the missionary data. Because we assume the latter use derived from the determiner use, *die* in the missionary data is glossed ‘DEF’ and tagged ‘DET’ also when it occurs as a relative pronoun.

The gloss INDF is assigned to the item *en*, which functions as the indefinite article. It is identical in form to the cardinal numeral ‘one’. It is often ambiguous whether *en* is an indefinite article or a cardinal numeral. Therefore, we have chosen to annotate *en* as a cardinal numeral only, when the reading of indefinite article is not available or clearly not appropriate, as in the example below, which talks about the Fates, the three goddesses that share one eye.

\ref.181

\tx	Di	dri	fa	zinə	mi	blin,	sini	ki	dē	ēn	hōgō.
\mb	di	dri	fan	sinu	mi	blin	sinu	ki	dē	ēn	hogo
\ge	DEF	three	of/from	3PL	be	blind	3PL	see/look	through	one	eye
\ps	DET	N.CAR	PREP	PRON	COP	PI	PRON	V	PREP	N.CAR	NN

‘The three of them were blind, they looked through one eye.’ DJDJ 1926:12

The multifunctionality of items of one word class to occur in the function of another word class is, as we have already seen, ubiquitous. This is particularly true for items that function as determiners (including quantifiers), as the majority of them can occur independently as well, a type of occurrence that I call pronominal. Thus in principle, when items that occur as determiners or quantifiers occur pronominally, they do not receive a different annotation. The exception is pronominal *di*, a high frequent core pronoun as such annotated as a pronoun, which is discussed in section (3.3.10).

Other determinative elements are *andu* ‘other’ and *sulke* ‘such’, of which the latter only occurs in the missionary data.

The missionary data contain occurrences of a fusional determiner, which incorporates a genitive case marking. This form is tagged DET and glossed DEF.GEN (see also section 3.3.6 on inflection).

\tx die Christus des HEere.
 \mb die Christus des heer
 \ge DEF Christus DEF.GEN Lord
 \ps DET NP DET NN
 ‘... the Christ of the Lord.’ 322A1:26

3.3.5 Quantifiers

Basic tag: QNT = quantifier
 Gloss: QNT.NEG = negative quantifier

Quantifiers are items that indicate quantity. As such, numerals are quantifiers as well, but we distinguish them from numerals in the annotations, because numerals form a functionally coherent class of items. Numerals are discussed in section 3.3.6.

The quantifiers other than numerals give a “relative or indefinite indication of quantity” (Matthews 1997:305). Common examples of quantifiers in the basilectal data are: *alma/alga/alda* ‘all’, *ēkē* ‘every’, *ēntēn* ‘no’ (glossed QNT.NEG), *muši* ‘much/many’, *ēnēstə* ‘any’, and *hēlə* ‘whole’.

Despite its gloss, the quantifier *sowē(l)* ‘so.much’ can occur both with count and noncount nouns.

\tx Sowē dak mi sowē jā mi bi hi
 \mb sowēl dag mi sowēl jā mi bē hi
 \ge so.much day with so.much year 1SG be here
 \ps QNT NN PREP QNT NN PRON COP ADV
 ‘So many days and so many years that I have been here...’ DJDJ 1926:17

The quantifiers *som* ‘some’, *ēnestə* ‘any’, *ēkē* ‘every’ and *ēntēn* ‘no’ combine with the noun *gut* ‘thing’ to form the indefinite pronominals ‘something’, ‘anything’, and ‘nothing’ respectively. *Ēntēn* can also be used as the indefinite pronominal ‘nothing’ by itself. As the examples below show, the compound indefinite pronominals have only been annotated as separate items when they occur as one word (see ref. 1121). The indefinite pronouns are discussed in section 3.3.11.

\ref .665

\tx as am wēt enestə gut fa di nistuk.
 \mb as am wēt ēnəstə gut fan di nestuk
 \ge when/if 3SG.AN know any thing of/from DEF handkerchief/shawl
 \ps CS PRON V QNT NN PREP DET NN
 ‘[The king called the girl] if she knew anything about the handkerchief.’ DJDJ 1926:18

\ref .1121

\tx Mi sa du soṅgut fo ju.
 \mb mi sa du somgut fo ju
 \ge 1SG FUT do IPR for/before 2SG
 \ps PRON TAM V PRON PREP PRON
 ‘I will do something for you.’ DJDJ 1926:24

3.3.6 Numerals

Basic tags: N.CAR = cardinal numeral
 N.ORD = ordinal numeral

Negerhollands distinguishes between cardinal and ordinal numerals. Ordinal numerals are always tagged as such in their entirety, and not as being a complex of a cardinal numeral and a suffix *-de* or *-ste*, even though this is the way the ordinal numerals are formed in the superstrate language Dutch.

In DJDJ 1926:67-68, the cardinal numerals upto and including one hundred are given (ref. 3304-3307 in the database). The cardinals upto and including sixteen are:

ēn, twē, dri, fi, feif, ses, sewun, ak, negə̃n, tin, èləf, twələf, dèrtin, vertin, feiftin, sestin,

The ordinal numerals upto and including twelve are given below:

\ref.3309

ēstə, twēdə, dridə, fide, feifdə, sesdə, sewəndə, akdə, negəndə, tində, èləfdə, twələfda

Note however that in the below example (ref. 1351), we find an occurrence of the form of the cardinal numeral used as an ordinal numeral.

\ref.1351

\tx	Di	fi	māl	wani	am	a	slā	si	tetsi	werán,
\mb	di	fi	māl	weni	am	a	slā	ši	tetsi	werā´
\ge	DET	four	time	when	3SG.AN	PST	hit	3.POSS	toe	again
\ps	DET	N.CAR	NN	CS	PRON	TAM	V	PRON	NN	ADV

‘The fourth time he hit his toe again, ...’ DJDJ 1926:28

Below we find the determiner *andə* used in the sense of ‘second’.

\ref.1975

\tx	Di	ēstə	ēn	ha	di	seskopdibəl,
\mb	di	ērstə	ēn	ha	di	seskopdibəl
\ge	DET	first	INDF	have	DET	six-headed.devil
\ps	DET	N.ORD	DET	V	DET	NN

\tx	di	andə	ēn	ha	di	fikopdibəl,
\mb	di	andə	ēn	ha	di	fikopdibəl
\ge	DEF	other	INDF	have	DEF	four-headed.devil
\ps	DET	DET	DET	V	DET	NN

\tx	di	dridə	ēn	ha	di	twēkopdibəl.
\mb	di	dridə	ēn	ha	di	twēkopdibəl
\ge	DEF	third	INDF	have	DET	two-headed.devil
\ps	DET	N.ORD	DET	V	DET	NN

‘The first one had got the six-headed devil, the second one had got the four-headed devil, the third one had got the two-headed devil.’ DJDJ 1926:39

The item *lāstə* ‘last’ is tagged as an adjective with the additional element .SUP in the tag to indicate that it is a superlative form (see section 3.3.8). It is not glossed as an ordinal numeral.

\ref.2889

\tx Di lāstu dak sinu a lo dig,
 \mb di lāstə dag sinu a lō dig
 \ge DET last day 3PL PST LO.IPFV dig
 \ps DET PLSUP NN PRON TAM TAM V
 ‘The last day they were digging, ...’ DJDJ 1926:58

See section 3.3.4 on the ambiguity of the item *ēn* between being an indefinite article and a cardinal numeral.

3.3.7 Focus marking

Basic tag: DISC = discourse marker
 Gloss: FOC = focus marker

The tag DISC and gloss FOC is assigned to the focus marker *da* that we encounter in the missionary data. Below is a representative example, that shows that it typically occurs in questions.

\tx 19. Da no Moses a ka gie jender die Wet?
 \mb 19. da no Moses a ka gie jender die wet
 \ge 19. FOC not Moses PST PFV give 2PL DET law
 \ps 19. DISC NEG NP TAM TAM V PRON DET NN
 ‘Didn’t Moses give you the/that law?’ 322A2:179

The focus marker *a* in the basilectal data can also be used as a copula, and is therefore tagged as a copula (see section 3.3.3).

3.3.8 Inflection

Basic tag: INFL = inflection
 Gloss: ABL = ablative
 ACC = accusative
 GEN = genitive
 INFL = inflection
 PL = plural
 PTCP = participle
 SUP = superlative

The Netherhollands missionary data contain some grammatical elements, which have very likely been transferred from Dutch or German. This includes morphological plural marking on nouns (glossed as PL); inflection on attributive adjectives (glossed as INFL); genitive case marking on article and noun (glossed as GEN); Latin case marking on Biblical names, which occurs either as ablative (glossed as ABL) or genitive case marking; and the use of a (past) participle form of the verb, combined with auxiliaries (tagged as copulas, see 3.3.3). We glossed these items as intended by the author/translator even though they probably were not a part of the NH as spoken by the NH mothertongue speakers of African descent. Another type of inflection is the marking of the superlative form of property items (glossed as SUP), which occurs in all varieties of NH.

Sometimes the inflectional elements are marked as separate items and tagged INFL, but in some cases the inflectional elements have been annotated as fused with the lexical item they are affixed to. Consider the following examples that compare past participle forms where the inflection has been annotated as a separate functional item and those tagged as V.PTCP.

\tx Ver doem niet; soo kom jender ook niet verdoemt.
 \mb verdoem niet soo kom jender ook niet verdoem -t
 \ge damn not so become 2PL also not damn -PTCP
 \ps V NEG ADV COP PRON ADV NEG V -INFL
 ‘Do not damn; then you will not be damned either.’ 322A1:88

\tx Oordeelt niet; soo jender ook no word ge oordeelt.
 \mb oordeel niet soo jender ook no word geoordeeld
 \ge judge not so 2PL also not become condemned
 \ps V NEG ADV PRON ADV NEG COP V.PTCP
 ‘Do not judge; then you will not be condemned.’ 322A1:88

3.3.9 Negation

Basic tag: NEG = negation

The tag NEG is assigned only to those items that express purely negation. On the clause level, these are *na*, *no*, *nu*, *ni*, *nə*, *ne* ‘not’. In example ref. 1775, the element *na* is the negation, tagged as NEG.

\ref.1775
 \tx So am na a stop tē am a slā am dōt.
 \mb so am na a stop tē am a slā am dōt
 \ge so 3SG not PST stop until 3SG PST hit 3SG dead
 \ps ADV PRON NEG TAM V CS PRON TAM V PRON PI
 ‘So he didn’t stop, until he killed him.’ DJDJ 1926:35

In the missionary data we find *no* ‘not’.

\tx Die Stadt, [...] no kan wees verborg.
 \mb die stad [...] no kan wees verborg
 \ge DET town [...] not can be hidden
 \ps DET NN [...] NEG TAM COP V.PTCP
 ‘The city [, that lies on a mountain] cannot be hidden. 322A1:76

Both in the missionary and in the basilectal data we find a separate item negating NPs only: *niet* and *nit* ‘not’ respectively.

\ref.1727
 \tx Nit ēn jamus kā sprout.
 \mb nit ēn jamus kā sprout
 \ge not one yam COMPL sprout
 \ps NEG N.CAR NN TAM V
 ‘Not a single yam had sprouted.’ DJDJ 1926:34

The negative quantifiers *ēntēn* ‘no’ (basilectal data) and *geen* ‘no’ (missionary data) have been tagged QNT.NEG (see section 3.3.5).

3.3.10 Pronouns

Basic tag: PRON = pronoun
 Glosses: 1SG = first person singular
 2SG = second person singular

3SG	= third person singular (inanimate or expletive)
3SG.AN	= third person singular animate
3.POSS	= third person possessive
1PL	= first person plural
2PL	= second person plural
3PL	= third person plural
DEM	= demonstrative pronoun
DEM.PROX	= demonstrative pronoun with proximal reference
DEM.DIST	= demonstrative pronoun with distal reference
IPR	= indefinite pronoun
RX	= reflexive pronoun
RC	= reciprocal pronoun
PRO.3	

3.3.10.1 Personal and possessive pronouns

In Negerhollands, pronouns are invariable. There is only one form for both subjects and non-subjects. This same form is also used attributively to function as a possessive pronoun, with the exception of the third person possessive, which has its own distinct form *ší* or *si(e)*, glossed 3.POSS. Although this form particularly has third person singular referents, and the third person plural pronoun can be used as a possessive pronoun, *ší/si(e)* does seem sometimes to have third person plural referents. Therefore, it is assigned the number neutral gloss 3.POSS.

There is an animacy difference in the third person singular pronouns: the pronoun *am* refers to animate referents, regardless of sex (see example ref. 920, with its human referent *Hans* (a boys' proper name) in example ref. 919, and example ref. 914, with its animate referent *dí kabái* 'the horse' in example ref. 912).

\ref.919

\tx	Hans	a	tumbl	bini	dí	baba.
\mb	Hans	a	tumbl	abini	dí	baba
\ge	Hans	PST	throw/tumble	inside/into	DEF	mud
\ps	NP	TAM V	PREP	DET	NN	

'Jack tumbled into the mud.' DJDJ 1926:21

\ref.920

\tx	Ham	a	sit	a	gron	lō	skrēw.
\mb	am	a	sit	a	gron	lō	skrēw
\ge	3SG	PST	sit	NA	ground	LO.IPFV	yell
\ps	PRON	TAM V	PREP	NN	TAM	V	

'He sit on the ground crying.' DJDJ 1926:21

\ref.912

\tx	Də	man	a	du	am	bō	dí	kabái.
\mb	dí	man	a	du	am	abo	dí	kabái
\ge	DEF	man	PST	do	3SG.AN	on	DEF	horse
\ps	DET	NN	TAM V	PRON	PREP	DET	NN	

'The man put him on the horse.' DJDJ 1926:21

\ref.914

\tx	weni	ju	mañké	am	stop,	ju	fo	sē:	„dži”!
-----	------	----	-------	----	-------	----	----	-----	--------

\mb weni ju mañké am stop ju fo sē dži
 \ge when 2SG want 3SG.AN stop 2SG FO say whoa
 \ps CS PRON TAM PRON V PRON CS V IJ
 ‘When you want it to stop, you have to say: “Whoa!” DJDJ 1926:21

The pronoun *di* refers to inanimate referents (see ref. 1706, with its inanimate referent *də houtu bak* ‘the wooden bowl’ in ref. 1705).

\ref.1705

\tx So am a sē də mēñši
 \mb so am a sē di mēñši
 \ge so 3SG.AN PST say DEF girl
 \ps ADV PRON TAM V DET NN

\tx du nē də houtu bak me di frokós.
 \mb du nē di houtu bak mi di frokós
 \ge do down DEF wooden plate/bowl with DEF breakfast
 \ps V ADV DET PI NN PREP DET NN

‘So he said to the girl: Put down the wooden bowl with the breakfast.’ DJDJ 1926:34

\ref.1706

\tx Dan di mēñši a du di nē.
 \mb dan di mēñši a du di nē
 \ge then DET girl PST do 3SG down
 \ps ADV DET NN TAM V PRON ADV
 ‘Then the girl put it down.’ DJDJ 1926:34

Di is also used as expletive subject (ref. 1694), in which case it has no reference at all.

\ref.1694

\tx a so mā di ha roto fandā.
 \mb a so mā di ha roto fandā
 \ge FOC so make 3SG have rat today
 \ps COP ADV V PRON V NN ADV
 ‘That’s why there are rats nowadays.’ DJDJ 1926:34

In the missionary data, the item *die* is more multifunctional than the basilectal item *die*, in that the former cannot only be used to refer to inanimates and as an expletive pronoun, it can also function as a pronominal demonstrative referring to both animates and inanimates. In order to do justice to this broad range of uses, pronominal *die* is glossed as PRO.3 in the missionary data.

3.3.10.2 Demonstrative pronouns

Basilectal Negerhollands contains two demonstrative pronominal items:

\ref.1688

\tx Puši a sē: dā mi gənú.
 \mb puši a sē dat mi gənú
 \ge cat PST say DEM.DIST be enough

\ps NN TAM V PRON COP ADV
 ‘Cat said: “That’s enough.”’ DJDJ 1926:34

\ref .1324
 \tx Mi lō sē ju diso nu.
 \mb mi lō sē ju diso nu
 \ge 1SG LO say 2SG DEM now
 \ps PRON TAM V PRON PRON ADV
 ‘I will tell you this (one) now.’ DJDJ 1926:27

The first item *da(t)* certainly has distal reference, but for the second item it is hard to tell whether it has only proximal reference is neutral in this respect or may also have distal reference. Even though the data do not contradict a proximal interpretation of *diso*, we have chosen to assign *diso* the distance neutral gloss DEM.

Note that there is one attributive occurrence the distal demonstrative *da*. Because of the rareness of such occurrences, attributive demonstrative *da* does not seem to be a central part of the Negerhollands attributive demonstrative system - if basilectal Negerhollands has ever had a stable attributive demonstrative system at all, that is. Therefore, this occurrence is annotated as a demonstrative pronoun, see ref. 2006.

\ref .2006
 \tx džis a da tit am a bigín fo du houtu bo Fergí
 \mb džis a dat tit am a bigín fo du houtu abo fergi
 \ge just NA DEM.DIST time 3SG.AN PST begin FO do wood on pig
 \ps ADV PREP PRON NN PRON TAM V CS V NN PREP NN
 ‘Just at that moment he began to put wood on Pig.’ DJDJ 1926:39

In the missionary data, we have the items *dat* and *deese*, but we are more sure of their reference. Therefore *dat* has been glossed as DEM.DIST, a pronominal demonstrative with distal reference, while *deese* has been glossed DEM.PROX, a pronominal demonstrative with proximal reference.

\tx Kik, dees word gesett tot een vall
 \mb kik dees word ge- set tot een vall
 \ge see DEM.PROX become PTCP- sit/set (up)to INDF fall
 \ps V PRON COP INFL- V PREP DET NN
 ‘Look, this is set as a trap...’ 322A1:27

3.3.10.3 Indefinite pronouns

Indefinite pronouns are glossed IPR. Examples are *ēntēngut* and *ēntēn* ‘nothing’, *ēkēgut* ‘everything’, *somgut* ‘something’, *ēngut* ‘anything’, *ekeren* ‘all/everyone’. As discussed in section 3.3.5, quantifiers can also occur as indefinite pronouns. In such cases, the items are annotated as quantifiers, as in the example below, where *muši* is used as an indefinite pronoun, but annotated as a quantifier.

\ref .2718
 \tx Sinu nu kan kri muši fan di jet.
 \mb sinu na kan kri muši fan di jet
 \ge 3PL not POSS get much/many of/from DEF food
 \ps PRON NEG TAM V QNT PREP DET NN
 ‘They cannot get much of the food.’ DJDJ 1926:54

Also discussed in section 3.3.5 is that certain quantifiers combine with the noun *gut* ‘thing’ to form indefinite pronouns. When the quantifier and *gut* do not occur as one word, they are tagged for what they are: a quantifier plus the noun *gut* (see example ref. 145). When they do occur as one word, they are glossed as an indefinite pronoun (see example ref. 1027).

\ref.145

\tx	Am	wēt	Prisjas	nə	ha	ēntēn	gut,	am	bi	pubu.
\mb	am	wēt	Prisjas	na	ha	ēntēn	gut	am	bē	pobu
\ge	3SG.AN	know	Perseus	not	have	QNT.NEG	thing	3SG.AN	be	poor
\ps	PRON	V	NP	NEG	V	QNT	NN	PRON	COP	PI

‘He knows that Perseus does not have anything, he is poor.’ DJDJ 1926:12

\ref.1027

\tx	ju	no	wēt	entengut	fa	sini.
\mb	ju	na	wēt	ēntēngut	fan	sinu
\ge	2SG	not	know	IPR	of/from	3PL
\ps	PRON	NEG	V	PRON	PREP	PRON

‘you know nothing of them.’ DJDJ 1926:23

Below is an example of the item *ēn* occurring as an indefinite pronoun (ref. 052). As can be seen, it is annotated as an indefinite determiner.

\ref.051

\tx	Ju	pupā	hā	fo	gi	siwun	juñ	mi	sēwun	minši.
\mb	ju	pupā	ha	fo	gi	sewun	juñ	mi	sewun	mēnši
\ge	2SG	father	have	FO	give	seven	boy	with	seven	girl
\ps	PRON	NN	V	CS	V	N.CAR	NN	PREP	N.CAR	NN

‘Your father has to give seven boys and seven girls.’ DJDJ 1926:11

\ref.052

\tx	Den	am	a	sē:	mi	sa	wēs	ēn.
\mb	den	am	a	sē	mi	sa	wēs	ēn
\ge	then	3SG.AN	PST	say	1SG	FUT	be	INDF
\ps	ADV	PRON	TAM	V	PRON	TAM	COP	DET

‘Then he said: “I will be one [of them].”’ DJDJ 1926:11

3.3.10.4 Reflexive pronouns

Reflexive pronouns are glossed for their person and number, plus the gloss .RX. When the reflexive element occurs as a separate element from the personal pronoun, the reflexive is only glossed RX.

\ref.1861

\tx	mi	sa	bidrāg	misél	leik	ēn	juñ	man.
\mb	mi	sa	bidrāg	miself	leik	ēn	juñ	man
\ge	1SG	FUT	behave	1SG.RX	like	INDF	young	man
\ps	PRON	TAM	V	PRON	PREP	DET	PI	NN

‘I will behave myself as a young man.’ DJDJ 1926:36

\ref .786

\tx ēn difman liki am self
\mb ēn difman leik am self
\ge INDF thief like 3SG RX
\ps DET NN PREP PRON PRON
'a thief like himself' DJDJ 1926:20

3.3.10.5 Reciprocals

The basilectal data do not contain a reciprocal element. The missionary data have *malkander* 'each other'. *Manikandə* does occur in the basilectal data, but only in the sense of 'together'.

\tx en soo alltwee met malkander bliev behouden.
\mb en soo alltwee met malkander bliev behouden
\ge and so both with RC stay preserved
\ps CC ADV QNT' PREP PRON V V.PTCP
'And so both are preserved with each other.' 322A2:110

3.3.11 Proform

Basic tag: PROF = proform

The tag PROF is used for items that refer to whole clauses, but that are not pronouns. It applies to the item *so* when used to refer back to clauses, typically the content of quoted speech in the narrative, as in the following example.

\ref .1271
\tx Wani si nom a ho so,
\mb weni ši nom a hō so
\ge when 3.POSS uncle PST' hear so
\ps CS PRON NP TAM V PROF
'When his uncle heard that, ...' DJDJ 1926:26

3.3.12 Question words

Basic tag: Q = question word
Glosses: Q.AN = animate question word
 Q.INAN = inanimate question word
 Q.LOC = locative question word
 Q.MNR = question word of manner
 Q.QNT' = quantificational question word
 Q.CAUS = question word of cause or reason
 Q.DET' = determinative question word

The category of question words does not easily fit in with most of the other categories, as a question word can in principle belong to any other category. Our solution to this problem is that we include an element in the gloss indicating the nature of the question word. A list of question words and how they are glossed is given below:

Item	Gloss	English translation
(a)widi	Q.AN	Who?
wa	Q.INAN	What?
(a)(w)api	Q.LOC	Where?

wamā	Q.CAUS	Why?
huso	Q.MNR	How?
huwēl	Q.QNT	How much/how many?

The question word *wa* ‘what’ (see ref. 372) also occurs as a determining question word, as in example ref. 296.

\ref.372

\tx Wa ju mēstər nam?
\mb wa ju mēstər nām
\ge Q.INAN 2SG master be.called
\ps Q PRON NN V
‘What is your master’s name?’ DJDJ 1926:15

\ref.296

\tx Prišjas ha sē: wa sōt fa fristə ju bē?
\mb Prišjas a sē wa sōt fan fristə ju bē
\ge Perseus PST say Q.INAN kind of/from lover/suitor/courtship 2SG be
\ps NP TAM V Q NN PREP NN PRON COP
‘Perseus said: “What kind of lover are you?”’ DJDJ 1926:14

Note that in some cases, the question words’ function is sometimes akin to a relative conjunction, as in the following example (ref. 2197).

\ref.2197

\tx Bot də hou frou fo hou am dā a di pit
\mb bot di hou frou fo hou am dā a di pit
\ge but DEF old woman FO keep/guard 3SG.AN there NA DEF well
\ps CC DET PI NN CS V PRON ADV PREP DET NN

\tx obu di tit wapi am kan kri fo mata di kapitein fan di bōt.
\mb obu di tit wa api am kan kri fo mata di kapitein fan di bōt
\ge over DEF time Q.LOC 3SG.AN can getFO kill DEF captain of/from DEF boat
\ps PREP DET NN Q PRON TAM V CS V DET NN PREP DET NN
‘But the old woman had to keep him there at the well for the time where he could get to kill the captain of the boat.’ DJDJ 1926:43

The missionary data contain the following question word which refers both animates and inanimates/abstract entities: *welk(s)* ‘which.one/what/who’. It is glossed Q.DET.

\tx Welks ben meer leucht vor see
\mb welks ben meer leucht vor see
\ge Q.DET be more light FOR say
\ps Q COP COMP PI CS V
‘What is easier(?) to say...’ 322A2:107

3.3.13 TMA marking

Basic tag: TAM = tense, aspect, mood/modality
Glosses: FUT = future
PST = past
COMPL = completive

IPFV	= imperfective
HAB	= habitual
LO.IPFV	= imperfective <i>lo</i>
LO.FUT	= future <i>lo</i>
LO	= <i>lo</i>
POSS	= possibility/ability/permission modal
NEC	= necessity/obligation modal
PERM	= permission modal
want	
like	

In Negerhollands, like most creole languages, verbs can be marked for tense, aspect, mood, and modality by preverbal elements. Because it is not uncommon that these elements express notions that cannot be characterised by tense or aspect or mood or modality only, but combine for example tense and aspect or mood and modality, any of these preverbal elements is tagged as TAM. Thus, we treat the items expressing notions that belong to the domain of tense, aspect, mood and modality as one class.

Preverbal markers that express past time reference of the verb are glossed as PST, while the gloss FUT is used for markers that express future time reference. It should be mentioned that the item *sa* glossed as FUT may also have modal connotations - which needs further investigation - or may be a marker of mood or of mood and a specific kind of modality if combined with the perfective aspectual marker *ka*.

\ref .848

\tx	Ju	sa	kā	drā	di	a	ju	han.
\mb	ju	sa	kā	drā	di	a	ju	han
\ge	2SG	FUT	COMPL	carry/bring	3SG	NA	2SG	hand
\ps	PRON	TAM	TAM	V	PRON	PREP	PRON	NN

'You should have carried it in your hand.' DJDJ 1926:20

The gloss COMPL is used for the preverbal item *ka*. It expresses completive aspect, which focusses on the state or condition of things resulting from the completion of the situation (Graves 1977:140; Bybee et al. 1994:54). *Ka* may also express resultative aspect, which focusses on the resulting situation of an event, so the gloss completive is not meant to exclude any other (closely related) aspectual uses.

In the basilectal data, we encounter the preverbal element *lo*, which has a range of uses. It is thus not always straightforward how to interpret this element. It can express progressive, habitual, and inchoative aspect, and it can also express future time reference. Besides these TMA-functions, it can also occur as a full verb whether involved in a serial verb construction or not. Where we think it is safe to assume one particular reading, we have annotated *lo* that way: i) when it has progressive or habitual aspect it is glossed LO.IPFV; ii) when it has future time reference it is glossed LO.FUT. When *lo* is a full verb it is glossed 'go' and tagged V. Note that also as a full verb in a serial verb construction, *lo* may express some TMA-related notion. When it is not clear how to interpret *lo*, it has been assigned the neutral gloss LO.

The missionary marker *le* is glossed IPFV, thus, as an element signalling imperfective aspect.

The habituality marker *kan* (identical in form to the modal *kan*) is glossed HAB.

Basilectal Negerhollands contains the following modal verbs:

Item	Gloss	Tag	Modality
<i>kan</i>	POSS	TAM	possibility, ability, permission
<i>ha fo</i>	have + FO	V + CS	necessity, obligation
<i>fo</i>	FO	CS	obligation, necessity
<i>mu(t)</i>	NEC	TAM	necessity, obligation
<i>manké</i>	want	TAM	volition
<i>wel</i>	like/want	TAM	volition

The volitional modals *manké* ‘want’ and *wel* ‘like’ have been assigned lexical glosses, because they are more frequent as lexical verbs than as modal verbs. When modal, both function as a volitional modal verb which translates as ‘want’.

Missionary Negerhollands has the following modal verbs:

Item	Gloss	Tag	Modality
<i>kan</i>	POSS	TAM	possibility, ability, permission
<i>moet</i>	NEC	TAM	necessity, obligation
<i>daerf</i>	PERM	TAM	ambiguous: permission or ‘dare’
<i>will</i>	want	TAM	volition
<i>mankeer</i>	want	TAM	volition

It is often not clear what reading is intended for the modal *daerf*. The lexical source of the item seems to be the German modal of permission *darf*, hence the gloss ‘PERM’. However, the possibility cannot be excluded beforehand that its use in missionary Negerhollands be inspired by the Dutch verb form *durf* ‘dare’.

3.4 Other categories

3.4.1 Abbreviations

Basic tag: ABBR = abbreviation

The tag ABBR is added to the tag of the abbreviated word, which is annotated as the full item for the rest (see example below).

```
\tx  §.          33. Matth.    9, 9=17.  
\mb  §{§}{PU} 33 Matth    9, 9=17.  
\ge  §         33 Matthew  9, 9=17.  
\ps  PU        33 NP.ABBR 9, 9=17.  
322A2:107
```

3.4.2 Foreign words

Basic tag: FW = foreign word

The tag FW is reserved for functional or closed class items – excluding prepositions - such as the use of the English pronoun *he* instead of the Negerhollands *am*. This because free borrowing is expected to be much less common and much more restricted with closed class items than with lexical items. An exception to this principle are English interjections and short pieces of quoted speech in English: These have been annotated as foreign word (see for example ref. 913).

```
\ref .913  
\tx  ju      fo   sē   am:    „hurry up”;  
\mb  ju      fo   sē   am     hurry up  
\ge  2SG     FO   say   3SG.AN hurry.up  
\ps  PRON    CS   V     PRON   FW  
‘You have to say to him: “hurry up”.’ DJDJ 1926:21
```

We have tried to use the tag FW as little as possible in the NEHOL database, because of the mixed origin of the vocabulary of Creole languages. This makes it hard to say which words are not and which ones are foreign, i.e. not a part of the language in question, e.g. as a result of code switching. Some of the English words in the text occur relatively often or at least multiple times. Therefore, English (and thus foreign) words - or group of words that make up only a part of speech - are not considered to be foreign words but borrowings. As such, they are considered part of the language and thus glossed and tagged as such. An example in favour of this strategy, is the occurrence of *bear it*, where *bear* is a borrowing from English, but the word *bear* is not completely identifiable with the English word *bear* because of the adverbial part *it* ‘out’. Only the whole construction *bear it* can be translated as ‘bear’. It would be strange to consider *bear* in this case as not being a NH word, because the whole construction it occurs in (‘bear it’) exists in NH only. Another example is the word *rocking-stul* which consists of the borrowed *rocking* combined with the NH word *stul* ‘chair’ to form the loan translation of English ‘rocking chair’. Another example is the English word *mischance* which is used and pronounced the English way, but in English it is a noun, while in NH it is used as a verb or perhaps even a verbal property item, meaning ‘abort’. It goes without saying that this does not imply that there are no items in the dataset for which it would be appropriate to see them as ‘foreign words’.

When whole sentences are English, this is encoded in the layer ‘\lang’ with the annotation: ‘english’. Thus, the tag FW is not needed in these cases either.

3.4.3 Formulaic expressions

Basic tag: FRM = formulaic expression

To be used for symbols and words that fulfill a certain function in the text, but do not have a clear meaning. Examples are: mathematical characters and the abbreviation symbol § for section, as in the example below.

```
\tx § I. Luc. 1, 5=25.
\mb §{§} {FRM} I Lucas 1, 5=25.
\ge § 1 Luke 1, 5=25.
\ps FRM N.CAR NP 1, 5=25.
322A1:8
```

3.4.4 Punctuation

Basic tag: PU = punctuation

Common punctuation, such as ‘, ’, ‘, ’, ‘, ’, and ‘?’ is not annotated: they can easily be searched for. Only the non-standard symbols that are part of the diplomatic editing system have been annotated (see section 2.3.3 for their meaning).

Item	Gloss	Tag	Example
@	page.break	PU	A)
		PU	B)

A)

```
\tx En @
\mb en @{page.break} {PU}
\ge and page.break
\ps CC PU
322A1:15
```

B)

```
\tx die HEer {sonder vout|onberispelik.}
\mb die heer sonder vout - | - onberispelik
\ge DEF Lord without mistake/flip - | - irreproachable
\ps DET NN PREP NN - PU - PI
“... the Lord without mistake/irreproachable ...” 322A1:8
```

3.4.5 Unclassified words

Basic tag: UNC = unclassified words

Typical unclassified items are noise words and pause fillers. Also annotated as unclassified in the basilectal data are magic formulas and song lines consisting of nonsense words in folk stories. An example containing such a line is given below:

```
\ref .1293
\tx Dan skilpat a anturt: ziñkaiô dahô amali wêlô marifomí marigêl!
\mb dan skilpat a anturt ziñkaiô dahô amali wêlô marifomí marigêl
\ge then turtle PST answer ziñkaiô dahô amali wêlô marifomí marigêl
\ps ADV NN TAM V UNC UNC UNC UNC UNC UNC
```

“Then turtle answered: “ziŋkaiŋ dahŋ amali wēlŋ marifomí marigēl”.’ DJDJ 1926:26

4. References

- Den Besten, Hans & Van Rossem, Cefas. to appear. Diplomatische editie van de Negerhollandse woordenlijsten van Frank G. Nelson. In *Tijdschrift voor Nederlandse Taal- en Letterkunde*. (digital)
- Matthews, P. H. 1997. *The Concise Oxford Dictionary of Linguistics*. Oxford: Oxford University Press.
- Sabino, Robin. 1988. The copula in vernacular Negerhollands. *Journal of Pidgin and Creole Languages* 3, 2, 199-212.
- Stein, Peter. 1986. The documents concerning the Negro-Dutch Language of the Danish Virgin Islands, St. Thomas, St. Croix, and St. John – Negerhollands -, in the Unitäts-Archiv (Archives of the Moravian Brethren) at Herrnhut. A commented Bibliography. In Hans den Besten (ed.), *Papers on Negerhollands, the Dutch Creole of the Virgin Islands*. *Amsterdam Creole Studies* IX, Publikaties van het Instituut voor Algemene Taalwetenschap 51, 19-31.
- Stolz, Thomas. 1986. *Gibt es das kreolische Sprachwandellmodell? Vergleichende Grammatik des Negerholländischen*. Frankfurt am Main/Bern/New York: Lang.
- Van Rossem, Cefas. to appear. De Negerhollandse woordenlijsten van Frank G. Nelson. In *Tijdschrift voor Nederlandse Taal- en Letterkunde*.
- Van Rossem, Cefas & Van der Voort, Hein (eds.). 1996. *Die Creol Tael. 250 years of Negerhollands texts*. Amsterdam: Amsterdam University Press.
- Van Sluijs, Robbert (2011). The modal system of 20th century Negerhollands. M.A. thesis, Linguistics Dept, Radboud University Nijmegen.
http://www.ru.nl/publish/pages/645263/robbert_van_sluijs_-_masters_thesis_-_definitief_sep_2011.pdf
- van Sluijs, Robbert. 2013. Negerhollands. In: Susanne Michaelis & Philippe Maurer & Martin Haspelmath & Magnus Huber (eds.), *Survey of Pidgin and Creole Languages, Volume I: English-based and Dutch-based languages*, Oxford: Oxford University Press.
- Verkruysse, P.J. 1973-1974. Over diplomatisch editeren van handschriften en het gebruik daarbij van diacritische tekens. In *Spektator* 3, 325-346.