オランダ語史料翻刻に関する研究会報告

　海外史料室においては、『日本関係海外史料　オランダ商館長日記』『大日本史料　十二編』欧文材料等のためにこれまで行なってきたオランダ語史料の翻刻作業の方法を振り返り、発展させていくために、原文編の校閲に協力を仰いできたオランダ歴史学研究所（ING）の編纂委員ワウテル・E・ミルデ氏を招聘し、2008年3月25日にオランダ語手書き史料の翻刻に関する研究会を開催した。以下は、研究会におけるミルデ氏の基調報告である。

GUIDELINES FOR THE TRANSCRIPTION OF 17TH CENTURY DUTCH MANUSCRIPTS

Wouter E. Milde

INTRODUCTION

The Dutch language is a Germanic language, but in order to reproduce it in writing, the Latin alphabet was used. However, several difficulties occurred, for example the Dutch language uses more vowels than Latin. On the other hand Dutch had no real use for consonants like c, x and q.

In order to make the Latin alphabet suitable to represent in writing the sounds of the Dutch language, several solutions were devised, mainly based on the joining and doubling of Latin letters.

Although to a lesser extent than the earlier Middle Dutch, Seventeenth Century Dutch was still regionally differentiated. There were differences in pronunciation in the various Dutch speaking regions. Since a generally accepted spelling system did not yet exist, these differences were also expressed in writing. In addition, there was a high degree of freedom in the use of letters that represented the same sound. Quite a few letters or combinations of letters were for example used for the sound k: k, c, ck, cq, q; and for the sound g: g, gh, ch gch etc. So contrary to Japanese and Chinese, where the written language, in spite of regional differences in pronunciation, was fixed, the Dutch written language was still marked by a great diversity.

Standardizing the Dutch written language has been a slow process. The important contributing factor in this unifying process was the influence of the language spoken in

Holland (the western part of the Low Countries, with cities like Amsterdam, Rotterdam and Delft) which was the leading region of the Republic. This factor was reinforced by the fact that the Statenbijbel, the bible for the protestant community in Holland, was translated into this language and that important writers like Hooft also used it. But it was not until the French period (the time of occupation by the French in the Napoleonic era) - after the demise of the VOC even - that the first official rules for spelling were laid down.

So spelling in the seventeenth century presented a highly irregular picture and this picture was made even more complex in VOC documents by the fact that a large proportion of VOC servants were of foreign descent. About half of the VOC personnel came from other countries, the majority of them from Germany.

As I just mentioned, only in the nineteenth century rules were laid down, in 1804 by Siegenbeek, in 1865 by De Vries en Te Winkel (editors of Woordenboek der Nederlansche Taal (WNT)), and in 1893 by Kollewijn respectively.

The three main rules were:
- To represent by letters each sound element that is heard in a word when it is pronounced by a cultured person. A kind of phonetic spelling therefore, but restricted by the following two rules.
- To introduce uniformity. So the numeral 5 should always be written vijf (and not vijv, fijf, veif etc.) and compounds should be recognizable: werelddeel not wereldeel, loopplank, not looplank.
- To incorporate traditional necessity: words that have acquired a permanency through long standing use, for example words with endings like -lijk and -isch.

PHONETICS/PHONOLOGY

While the Latin alphabet originally (in the second century B.C.) had only 21 letters - of which v and i acted both as a consonant and a vowel (j, u, w, y en z were missing) - and 4 diphthongs: ae, au, eu and oe, the Dutch language uses the following large number of vowels (in open and closed syllables) and diphthongs:

Vowels:

a: dak

aa: baat, vader

e: pet

ee: wees, wezen

i: pit, Egypte

i(e):   liter, dier

o: rok

oo: dood, boven

u: nut

uu: vuur, nu, duwen

eu: leuk, veulen

oe: zoet

e: de, monnik, heerlijk（silent vowel）

Diphthongs with i:

ei/ij: zeis, vijf

ui: ruit

aai: fraai

ooi: mooi

oei: foei

Diphthongs with u:

au/ou:rauw/goud

eeu: meeuw

ieu: nieuw

u: uw

（The u in combination with the q is not regarded as a vowel.）

On top of this, we have to take into account that the e is also used for lengthening a vowel instead of a double vowel (doet/doot), kwaed/kwaad).

To make matters worse, there are a lot of phonological peculiarities in the Dutch spoken language of the 17th century, which spill over in the written language. Below follows a list (not nearly a complete one). Looking at it one would almost think that all vowels in Dutch were freely interchangeable and that it could have been written with only consonants, like the Semitic languages.

Regarding vowels:

Before the r, a and e are often interchanged (hart/hert).

| The a is often used instead of | e | (bekend/bekand) |
| | ui | (wambuis/wambas). |

The e is often used instead of	aa	(twaalf/twelf)
	o	(worden/werden)
	i	(hitte/hette)
	ee	(vreemd/vremd).

| The i is often used instead of | e | (hert/hirt) |
| | u | (stuk/stik). |

Before cht the o is often used instead of a   (dacht/docht).

| The u is used instead of | o | (kontschap/kuntschap) |
| | i | (vrind/vrund). |

| The ee is used instead of | ei | (klein/kleen) |
| | aa | (paard/peerd). |

The eu is used instead of   oo   (voor/veur).

| The uu is used instead of | ui | (duisent/dusent); |
| and is sometimes inserted between r and w | | (taruwe, verruwe). |

The ie is used instead of   uu   (sturen/stieren).

The ou is used instead of   u   (huwelijck/houwelijck)

| The ei/ey is used instead of | e | (denken/deynken) |
| | ee | (menen/meynen). |

| The ij/y is used instead of | i | (binnen/bynnen) |
| | ie | (drie/drij). |

Regarding the consonants:

nw is sometimes written as   w   (lijnwaet/ lijwaet

ft is sometimes written instead of   cht   (gracht/graft, krachtig/kraftig)

The d between two vowels is sometimes dropped (neer/neder, mee/mede).

Things were complicated even more because of the (already mentioned) fact that many VOC servants did not come from Holland and had another mother tongue.

SPECIFIC FEATURES OF 17TH CENTURY TEXTS

In manuscripts a capital to mark the beginning of a sentence and a full stop to mark its end are often missing.

Seventeenth century writers in general do not make much use of the semicolon, instead they often use the colon.

Seventeenth century writers make much use of the comma, but not very consistently.

Quotations are not always placed between quotation marks, and if quotation marks are used they are often placed not at the beginning and end of the quotation, but at the beginning of the sentence(s).

There is great confusion about whether or not to write words together or separately. Words like het/'t, ze/se are often written together with the following or preceding word respectively.

A hyphen is often omitted (gins en wederreis instead of gins- (the hyphen indicating the omission of "reis") en wederreis).

GUIDELINES

Reproduction of text.

With regard to the reproduction of the text of written historical sources, in principle two clearly differentiated main approaches are possible: the critical and the non-critical approach.

In the first case the text is being revised to a certain degree, for example by standardizing the use of capitals, punctuation etc. However, it can even go so far as to reconstructing a no longer existent original text from available copies. The aim is bringing about an ideal text for a certain purpose.

With regard to the non-critical, also called diplomatic, approach the aim is to reproduce a text as exactly as possible (even in its concrete representation as document), word for word, and letter for letter. However, non-critical is of course not to be confused with uncritical, so adjustments are possible.

The answer to the question which approach should be selected for the reproduction of a text is determined by the objective that one has in mind.

By the way, apart from the aforementioned two main approaches there is a secondary mixed method as well.

These three approaches are derived from commonly accepted practice in editing texts. The critical-normalizing approach is mainly applied to the editing of material of historical scholarly importance, while the non-critical diplomatic approach is mainly used for philological purposes. The mixed approach is applied when there is a convergence of both interests.

It is obvious that the non-critical approach is not suitable for the editing of historical records like the dagregisters of the VOC, because it is in principle aimed at a typographic reproduction of a text with all its internal and external formal aspects. It is mainly used for linguistic/philological research.

In view of the nature of the texts and the background of the writers, the mixed approach also is not the ideal method for these texts.

The objective of the critical-normalizing approach is primarily to reproduce the textual aspects of the source for the sake of research that is solely directed at the contents of the text. It allows a rather high degree of normalizing of the internal-formal features of the source (of course observing the integrity of the text), while in principle leaving the external features as they are.

The critical-normalizing reproduction is usually applied to the editing of texts from after the Middle Ages.

THE CRITICAL-NORMALIZING METHOD

1 Transliteration

The single i and j in Dutch texts are printed i when having the function of a vowel, j when having the function of a consonant. (gejaecht instead of geiaecht)

The double i (ii or ij) in Dutch texts is printed either as ij or as y, depending on the following:

- in case of a short i-sound (between two consonants) the y is used, i.e. bynnen, not: bijnnen
- in case of a (present-day) long i-sound, the ij is used, i.e. ijder, ijmand, drij, brijven, not: yder, ymand, dry, bryven
- in case of a combination with a vowel the y is used, i.e. yeder, nyet, huys, oyt, not: ijeder, nijet, huijs, oijt
- in case of the (present-day) ei-sound the ij is used, i.e. wij, zijn, blijven, not: wy, zyn, blyven

The use of the letters u, v, uu (u doubled), vv (v doubled) and w is being standardized in accordance with the (assumed) modern usage, i.e. they are printed as u, and uu respectively when they have the function of a vowel, and as v and w respectively when their function is a consonant (boven, uut, not: bouen, wt).

Numbers for which Latin numerals are used are generally printed in Arabic numerals; the symbol for "half" is printed ½.

In extensive enumerations of quantities or numbers Roman numerals have to be replaced by Arabic numerals. However, in dates the Roman numerals are maintained.

The use of capitals and lower case / small letters is standardized according to current usage.

Capitals are used in proper names (of persons, places and things, such as Compagnie, Staten-Generaal, Synode) and in titles (Z.M., U.E.), but not in the names of months, quarters of the compass and functions/jobs (so: februari, noordoosten, gouverneur-generaal).

In case a prefix is written together with the following word (the silent e having been omitted), it is to be separated from that word by inserting an apostrophe, i.e. d'andere instead of dandere, t'allen instead of tallen etc. In other cases too where a silent e is omitted, an apostrophe is inserted, i.e. 'tselve instead of tselve, 'tgene instead of tgene. In general, the use of the apostrophe is being standardized, for example in the case of plurals (komma's instead of kommas).

The apostrophe can be used in Japanese names if otherwise there would be confusion about the underlying characters (Kan'ei).

2 Accents

The use of accents is standardized according to current usage (vóór instead of vòòr, dáár instead of dààr).

The use of a dieresis (two dots over a letter) is being standardized according to current usage, i.e. Indiën, resolutiën.

For the use of the apostrophe, see above.

3 Abbreviations

Abbreviations are tacitly / automatically written in full, as much as possible on the analogy of occurrences of the word written in full in the same text. If the editor is not absolutely certain of the correctness of an option, the abbreviation is being maintained, or the chosen option is provided with a note, if necessary indicating other possible options. Frequently occurring abbreviations that have more than one solution are printed abbreviated: voors. (voorseid/voorschreven)

The following exceptions apply:

- Abbreviations that are still in use, a°, etc.
- Abbreviations for coins, weights and measures. These abbreviations are written in small letters (t. for tael, r. for reaal)
- Titles and terms of address such as Z.M., Z.Ex., U.E., m.h^r, s^r.
- Abbreviations of words, the full version of which was no longer in use even at that time (adij)

4 Joining or separating words

Separate words that are written together in the text must, if possible, be printed separately, according to current usage (so "datze" becomes "dat ze", but mettet remains unchanged).

Parts of words that are separated in the text must, if possible, be printed written together, according to current usage (so "sich selven" becomes "sichselven" and "buyten 's lands" becomes "buytenslands").

5 Punctuation

The editor provides the text with punctuation according to current usage, always mindful of the fact that providing or changing punctuation can alter the meaning of the text.

6 Arranging and lay-out of the text

The editor can introduce an arrangement in sentences, (sub)sections, columns, articles, paragraphs, and chapters, as well as a typographic lay-out according to current usage.

7 Alterations/additions

In case the original text shows alterations and/or additions from the same time, then the altered and/or supplemented text will be printed. The character of the change will be mentioned in a note (for example: an addition in the margin or between the lines, a deletion etc.), if this is necessary for a good understanding of the text.

8 Marks

Marks that have no significance for a good understanding of the text as well as the use of different colours of ink are not reproduced in the text. The editor can mention it in a note.

9 Textual uncertainties

In case the editor is not sure how to read a certain part of the original text, he prints the version that in his opinion the most likely. In a note he specifies what is uncertain and if possible gives other possible versions.

10 Gaps / Lacunae

Parts of the original text that are lost or have become unreadable because of damage, fading or other causes:
- If the editor can reconstruct them with reasonable certainty, they are printed between square brackets ([ ])
- If reconstruction is impossible, the editor will either indicate how many characters in his opinion are lost by putting as many dots between square brackets, or if that is impossible he will put three asterisks between square brackets

In his introduction or in notes the editor explains the nature of the gaps / lacunae and justifies his reconstruction.

Lacunae that are not noticeable as such - for example because of the fact that something was omitted by accident in the original text - are not indicated in the text, but in notes, and as far as possible reconstructed.

11 Errors

Obvious errors in the original text are printed rectified. If necessary the original error can be reproduced in a note.

12 Abridging

If the editor does not print parts of the original text, this must be indicated by three dashes (- - -). The abridgement is being justified in the introduction if it concerns general categories and in notes if it concerns incidental cases. The editor can give a summary of what he has omitted, either in a footnote or in the text itself, but in that case it has to be distinguished clearly.

PUNCTUATION MARKS

A comma is used to indicate a short pause between two equivalent sentences or parts of sentences, except when these are joined by *en* or *of*. The use of commas is therefore closely connected with the grammatical structure and rhythm of the sentence.
If there is no pause between the parts of a compound sentence, no comma is put in. However, if clearly a pause is heard, then, even before *en* or *of* a comma is placed.
Before a restrictive subordinate clause no comma is placed. (De agent die daar loopt, is onze buurman. The officer who is walking there, is our neighbour.) On the other hand, a non-restrictive or amplifying subordinate clause must be placed between commas. (De agent, die immers het gezag vertegenwoordigt, moet worden gehoorzaamd. The officer, who represents authority, must be obeyed.)

Note the difference between:
"The soldiers who had fled, were punished." (Only the soldiers who had fled were punished, not the others.)
and
"The soldiers, who had fled, were punished." (The soldiers had fled and therefore were punished.)

Mijn bedoeling was eigenlijk, niet te blijven = Actually my plan was: not to stay
Mijn bedoeling was eigenlijk niet, te blijven = Actually it was not my plan to stay

The semicolon indicates a pause as at the end of a sentence. Two main sentences that belong together are usually separated by a semicolon, especially when (in Dutch) no conjunction is used. (Hij spant zich geweldig in; toch bereikt hij weinig.) (…; all the same …)
The use of the semicolon, however, is rather arbitrary.

The colon indicates that an explanation, an enumeration, an addition or a quotation is following.

Quotation marks are used when someone's words are quoted literally. In indirect speech quotation marks are not used. Sometimes an idiosyncrasy / a word peculiar to a certain group etc. is placed between quotation marks. (In soldiers language this is called "…".)

A question mark is only placed after a direct question, not after an indirect one.

An exclamation mark ends a sentence that is considered an exclamation; therefore it also ends a wish, an order and a warning.

A dash usually indicates a long pause, or something unexpected. (It was a peaceful night - suddenly a shot rang out.)

So far the rules I [try to] adhere to when editing Seventeenth Century Dutch texts. They are mainly based on the following works and guidelines compiled by the late Dr. M. E. van Opstall (Algemeen Rijksarchief).

Zeventiende-eeuwse Taal
Dr. A. Weijnen
Derde druk
N.V. W.J. Thieme & Cie
Zutphen　1960

Nederlandse Spraakkunst
Dr. C.G.N. de Vooys
Derde druk
J.B. Wolters
Groningen　1953

Nederlandse Spraakkunst
Dr. E. Rijpma en Dr. F.G. Schuringa
Achttiende druk
J.B. Wolters
Groningen　1961

Richtlijnen voor het uitgeven van historische bescheiden
Samengesteld in opdracht van het Nederlands Historisch Genootschap
en van de Rijkscommissie voor Vaderlandse Geschiedenis
Zesde herziene druk
's Gravenhage   1988

Bronontsluiting voor historisch onderzoek
Instituut voor Nederlandse Geschiedenis
's Gravenhage   1990

Beknopte geschiedenis van de Oud- en Middelnederlandse Letterkunde
Dr. J. van Mierlo S.J.
Zesde druk
1954
Uitgeversmij N.V. Standaard-Boekhandel
Antwerpen / Amsterdam

　海外史料室の出版物におけるオランダ語史料翻刻の基準は、基本的には上記のミルデ氏の報告に述べられるものに準拠しているが、別の方針を採っている部分も若干ある。その点にも触れつつ、以下に、当日の議論の要点を摘記する。

１）大文字と小文字の使い分け
①ミルデ氏の示された基準では、組織内での機能を示す語は小文字、称号は大文字とされているが、現在のところ『日本関係海外史料　オランダ商館長日記』では、史料内での識別のしやすさの観点から、gouverneur-generaal（総督）についてのみ、称号と同様に、大文字を用い、Gouverneur-Generaalと表記している。また、keyser 皇帝についても、徳川将軍を指す場合のみ大文字を用いている。
②単位の省略形についても、同様の理由で、T.（tael）　ƒ.（gulden）　r.（reaal）を用いてきたが、これについては単位によって大文字と小文字、イタリックとローマンが混在しているのはやはり統一したほうがよいとの意見を得た。今後検討したい。二人称の u、uw についても同様である。
２）省略形について
　省略形を元の形に戻す際、語尾を単数形にするか複数形にするか、綴字をどうするか等の問題がある。これについては非常に判断が難しいことが参加者の共通認識であった。ミルデ氏も best way in that text, あるいは　筆者自身が書いたかのように、と言うしかない、との見解だった。
３）単語の分割／一語化について
　概して17世紀のオランダ語は一語に書く傾向がある。依拠するとすれば現代の綴字辞典との由であったが、それも必ずしも絶対とは考えられていないようである。
４）文字の使い分け
　オランダ語史料翻刻における文字の使い分けでは、ij と y の使い分けが一番難しい。これは本文に記述されているように基本的には音節の開閉と母音の長短、前後が母音か子音かによって決定されるが、オランダ語以外の外国語や外国地名等を聴き取って綴ったもの等においては、当時の発音は復元困難であり、特に混乱が生じやすい。発音における隔たりの感覚、どちらのルールがより優先度が高いか等、オランダ語を母語としない編纂者にとっては容易に判定しがたい要素も多く、単語ごとの吟味は不可欠であろう。

　その他に、句読法、原文中の文法・語法上の誤りの扱いなど幅広い議論が行なわれた。
ミルデ氏の提示された方針は、オランダ歴史学研究所の編纂物等、オランダ語で、オランダ人の読者に、オランダ史の史料を読みやすく提供するための方法という発想が基本にある。その点で史料編纂所の出版物の欧文史料とは性格が異なる部分もあり、より原本の状況に手を加えた編纂が行なわれている印象を得た。一方、統一性、凡例における基準・例外の明示の重要性等については共通認識を確認することができた。
　今後も可能な機会にこのような意見交換を行ない、編纂に生かしていきたい。
本研究会は、科学研究費基盤研究（B）近世日蘭関係史料の相関的研究（代表者：松井）との連携により開催した。

<div align="right">（文責・松井洋子）</div>