## **Cleansing operations**

This is the comprehensive compendium of the cleansing operations. It includes detailed descriptions, treatment methods, and underlying motivations.

## Legend

- When find "a.e.", it means "anything else".
- The regular expression symbol ".\*?" indicates an undefined sequence of words, numbers, and/or graphic symbols.
- The brackets "(£ £)", stand for a comment made by the author to explain contents.

**Table 1: Translation to German or English** 

Meaning	Managing	Reason
Typo or data collection	Reduced to a single	It could distorts the
error	space	translation
Sentence that should	Erased	The sentence is fully
continue		translated with the only
		addition that it should
		continue
		It could distorts the
	sentence	translation
		It could distorts the
	sentence	translation
± ,		
•		
of a single word		
Space in which a lack or	Freed the entire	It could distorts the
=		translation
	Schichee	translation
translated is indicated		
	Typo or data collection error Sentence that should	Typo or data collection error  Sentence that should continue  Sentence not translated or which could not be translated "Destroyed" (various synonyms and specifications). The single use of seems to indicate the destruction of a single word  Space in which a lack or a part that cannot be  Reduced to a single  Erased  Erased the entire sentence  Sentence  Erased the entire sentence  Erased the entire sentence

Subject	Meaning	Managing	Reason
			2.000.002
[]			
keine Übersetzung vorhanden Keine Übersetzung möglich LEER GEFUNDEN	No translation available. No translation possible FOUND EMPTY		It could distorts the translation
 LEER GEFUNDEN			
(= )	Parenthesis indicating the explanation of the object reference	Erased brackets and content	They could distort the translation
(( ))	Parenthesis indicating ancient additions to the text (above the line or next to the column) <sup>1</sup>	Erased brackets, kept content	The translation was made thanks to these additions
[[ ]]	Antiche ricostruzioni sopra cancellazioni <sup>2</sup>	Erased brackets, kept content	The translation was made thanks to these reconstructions
\"arbustes à épines\" rôdeurs	Thorny shrubs Marauders	Translated into dornige Sträucher Translated into plünderer	French could distorts the translation
دن بي د	Quotation marks indicate complex proper nouns	Deleted quotation marks, kept content	The translation was made thanks to what is inside
(£ between two words £)	Identify two different translations, one before / and the second after	We decided to keep the first word and eliminate the second. The choice was arbitrary and made without observing which of the two was the better translation	It could distorts the translation
- LHG LHG -LHG -LHG - {LHG} LHG - LHG l.h.g. l.h,.g. l.h-g l.h.g . l.h.g - l.h.g - l.h.g	Leben, Heil, Gesundheit. Life, Prosperity, Health  Inverted quotation	Replaced with "Leben, Heil, Gesundheit" or "Life, Prosperity, Health"	To improve the translation  They could distort the
,,	- Inverted quotation	Defered quotation marks,	They could distort the

<sup>1</sup> Höpe2 *Ibid*. Höper J. et al., Berlin Text System 3.1 User Manual (V 3.0), 2018, p. 106.

Sub	ject	Meaning	Managing	Reason
		marks, in some cases	kept content	translation
>:	,	opened and not closed,	_	
		in others closed without		
		being opened		
Г	1	Brackets indicating	Erased brackets, kept	The translation was
		partially damaged	content	made thanks to these
		content or unclear		reconstructions
		words/consonants <sup>3</sup>		
\$[	]\$	Parenthesis indicating	Erased brackets and	They could distort the
	3.	the descriptions of the	content	translation
		depictions with		
		references to the		
		registers and areas		
[	1	Parenthesis indicating	Erased brackets, kept	The translation was
		reconstructions usually	content	made thanks to these
		of destructions marked		reconstructions
		by the signs "//" or "/"		
<	>	Parenthesis indicating	Erased brackets, kept	The translation was
		the additions to	content	made thanks to these
		reconstruct the meaning		additions for the
		theyr wanted to give to		reconstruction
		the sentence		
		Parenthesis indicating	Erased brackets, kept	They could distort the
C	C	the Serekh.	content	translation
		Parenthesis indicating		
		the Cartouche		
,	6	Superscripts	Deleted superscripts,	They could distort the
			kept content	translation
?	?	Parentheses in the	Erased brackets, kept	The translation was
		translation that indicate	content	made thanks to what is
		an uncertain reading		inside
		(sometimes also together		
		with square brackets) <sup>4</sup>		
(	)	Parenthesis indicating	Erased brackets, kept	They could distort the
		the Cartouche	content	translation
		In some cases found in	Erased	It could distorts the
		NN , in another case		translation
		found in  Saka  (these		
		were handled later)		
[§(£ +nur	nbers £)]	It probably indicates the	Erased brackets and	It could distorts the
		chapter number at its	content	translation
$[\S(\pounds + numb)]$	ers +letters	beginning		
£)	)]			
§(£ + num	nbers £)(£	It probably indicates the	Erased everything found	It could distorts the
+space, per	riod, end of	chapter number at its	inside the second	translation
		<del>-</del>		

*Ibid*.4 *Ibid*.

Subject	Meaning	Managing	Reason
sentence, colon, or a.e. £)	beginning	parenthesis of our comment, i.e. (£ +space, period, end of sentence,	
§(£ +space +numbers £) (£ +space, period, end of		colon, or a.e. £)	
sentence, colon, o a.e. £)			
§(£ +space +numbers, also divided by a - £) (£			
+space, period, end of			
sentence, colon, or a.e.			
£)			T. 11.11
- Variante.*?-	Translation variant	Erased the entire	It could distorts the
Variante.*? (£ with the writing		sentence	translation
Variante in any position			
of the sentence £)			
Variante.*?	Indicates a translation		In this case the variant is
(£ with Variante at the beginning of the	variant	kept the rest	the only sentence present, so we chose to
sentence and the end			keep it
of .*? at the end of the			кеери
sentence £)			
u.	und.	Replaced with their	It could improve the
V	von.	meaning:	translation
OÄ o ï	Oberägypten.	and	
O.Ä	Unterägypten	of	
OÄg OÄ.		Upper Egypt	
O.Äg		Lower Egypt	
0.Ä.			
OÄg.			
O.Äg.			
UÄ			
U.Ä			
UÄg			
UÄ.			
U.Äg U.Ä.			
UÄg.			
U.Äg.			
NN	Indicate a proper noun to	Erased	Translation done without
NN	insert		the noun
NN ( ii , ii'a)	T 1, 1 1 1	D 11 1 1	T. 11.11.
(wört.*?)	Literal explanation	Erased brackets and	It could distorts the
(wört.*?	Dravious varsion of the	content  Erosed breekets and	translation  It could distorts the
[ältere Fassung.*?]	Previous version of the	Erased brackets and	It could distorts the

Subject	Meaning	Managing	Reason
(älterer Text.*?)	translation	content	translation
(oder.*?)	Another version of the	Erased brackets and	It could distorts the
	translation	content	translation
[Beischrift.*?]:	Explanation of any other	Erased brackets and	It could distorts the
[Beischrift.*?] (£ a inizio	<u> </u>	content	translation
frase £)	surrounding the		
[.*?Beischrift.*?]	inscription		
6	Direct speech, residues	Erased	It could distorts the
66	from cleaning		translation
(£ both at the beginning	operations, or other		
and at the end of the	,		
sentence £)			
(£ spaces at the	-	Erased	It could distorts the
beginning and end of the			translation
sentence £)			
	-	Erased	It could distorts the
(£ at the beginning of the			translation
sentence £)			
$\{.*?\}\langle .*?\rangle$	Identify two different	Only the translation	Arbitrary choice, not
\(\lambda.*?\)\(\lambda.*?\)	translations, one in the	contained within the	because one is better
(£ also divided by spaces	first bracket and one in	curly bracket was	than the other. Because It
and other words, from a	the second	retained, the other was	could distort the
minimum of one word to		deleted	translation
a maximum of two			
words (and a variable			
quantity of spaces			
between them and the			
brackets) £) <sup>5</sup>			
{{ }}}	Ancient destructions <sup>6</sup>	Erased brackets, kept	Translation carried out
		content	thanks to these additions
$\langle\langle$	Haplography <sup>7</sup>	Erased brackets, kept	Translation carried out
		content	thanks to these additions
".*?/.*?"	Identify two different	We decided to keep the	It could distorts the
	translations within the	first word and eliminate	translation
	quotation marks, one	the second. The choice	
	before the / and the	was arbitrary and made	
	second after	without observing which	
		of the two was the better	
		translation	
\(\)	Parenthesis indicating	Erased brackets, kept	Translation carried out
	consonants/words	content	thanks to these additions
	mistakenly omitted by		
	the ancient scribe and		
	inserted or corrected by		

We also decided not to exceed the two words between the different brackets because we believed that they are no longer two translation variants if separated by more words. Höper J. et al., *op. cit.*, p. 106.

Ibid.

Subject	Meaning	Managing	Reason
3	the publisher (same of	3 3	
	the source) <sup>8</sup>		
{ }	Parenthesis indicating	Erased brackets, kept	Translation carried out
	consonants/words	content	thanks to these additions
	considered erroneous		
	and superfluous by the		
	publisher (same of the		
	source) <sup>9</sup>		
(	Probably a typing error	Replaced with '	It could distorts the
	to indicate a '	1	translation
<b>N</b>	Probably a typing error.	Replaced with '	It could distorts the
	It appears in one case	1	translation
	and could be in place of		
	a superscript		
&	Translated to "e"	Replaced with "und"	It could distorts the
		1	translation
=	Equivalent to the sign =	Replaced with =	It could distorts the
		1	translation
\$	Sign used to create a	Erased brackets, kept	It could distorts the
	bracket in which	content	translation
	transliteration is		
	maintained in the		
	translation, probably due		
	to the inability to		
	recognize the translated		
	object		
	(e.g. cloak \$mAs.t\$)		
#	Sign used in one case	Erased brackets, kept	It could distorts the
	only as the same of the	content	translation
	precedent \$		
(?)	Used in doubt translation	Erased	Translation carried out
			together with everything
			that generates doubt
"	Residues from cleaning	Erased	It could distorts the
(£ operation done again	operations		translation
£)			
(d.h*?)	(i.e*?)	Erased brackets and	It could distorts the
		content	translation
	Other possible	Erased brackets and	They could distort the
	translations; literal	content	translation
	meanings; used to reveal		
	the subject (both noun		
	and pronoun);		
	explanations of various		
	kinds		

*Ibid*.9 *Ibid*.

	Subject	Meaning	Managing	Reason
(	£ Double spaces again	Residues from cleaning	Reduced to a single	They could distort the
	£)	operations	space	translation

Table 1: Translation management.

**Table 2: Transliteration (wChar)** 

Subject	Meaning	Managing	Reason
(£ Double spaces £)	Typing or data collection	Reduced to a single	It could distorts the
	errors	space	translation
  	Space in which a lack or part of it that cannot be transliterated is indicated;	Erased the entire sentence	It could distorts the translation
_	one or more signs destroyed, but is not always in line with the hieroglyphic field in quantity or position		
-??-	Space in which a lack or part of it that cannot be transliterated is indicated	Erased the entire sentence	It could distorts the translation
(( ))	Parentheses indicating ancient additions to the text (above the line or next to the column)	Erased brackets, kept content	The transliteration was done thanks to these additions
[[ ]]	Ancient reconstructions over cancellations <sup>10</sup>	Erased brackets, kept content	The transliteration was done thanks to these reconstructions
? ?	Parentheses indicating a dubious reading (sometimes also together with square brackets) <sup>11</sup>	Erased brackets, kept content	The transliteration was done thanks to what is inside
~	Probably a graphic sign to separate objects of dubious translation which could have different transliteration	Erased	The transliteration was also made thanks to what this sign divides
.plpl .{pl} .{pl} .{pl} .pl} .pl	Used for plural	Erased	They could distort the translation
	Reconstruction of gaps restored by the publisher <sup>12</sup>	Erased brackets, kept content	The translation was made thanks to these reconstructions
-(Zahl)- Zahl	Number	Erased	It could distorts the translation
oder ḫr =s	Or $hr = s$	Erased	It could distorts the

<sup>10</sup> Ibid.

<sup>11</sup> *Ibid*.

<sup>12</sup> *Ibid*.

Subject	Meaning	Managing	Reason
			translation
ON GN	"Ortsname": Place name "Göttername": Diety name	Erased	It could distorts the translation
a	Probably a typing error as it is found in the names "Ptah" and "Wasjr"	Erased	It could distorts the translation
( )	Parentheses indicating words/consonants missing or omitted to abbreviate, and inserted by the translator (for example plurals, feminine, other letters), useful for correct translation	Erased brackets, kept content	The translation was made thanks to these additions
ГП	Parentheses indicating partially corrupted content or unclear words/consonants <sup>13</sup>	Erased brackets, kept content	It could distorts the translation
1Q	Indicates the destruction	Replaced with a / for	Adjustment to the use
2Q	of 1 or more squares in reference to hieroglyphics	each destroyed panel. Added a space in the middle in the case of two boxes	of / as the destroyed area of the hieroglyph
\(\( \mathbb{E} \) \word \mathbb{E}\)\\\\.*?\{(\mathbb{E} \word \mathbb{E})\}.*?\{(\mathbb{E} \word \mathbb{E})\}.\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	Most of the time <sup>14</sup> results in a construction of double translations within the two types of close brackets, namely () and {}	Erased everything before the end of the last parenthesis \( \) and all the parentheses inside the curly braces	It could distorts the translation
$ \begin{cases} .*? \\ \langle .*? \rangle \\ \langle .*? \rangle \\ \text{(£ e variations}^{15} \text{£)} \end{cases} $	Most of the time <sup>16</sup> results in a construction of double translations within the two types of close brackets, namely $\langle \rangle$ and $\{ \}$	Erased all the parenthesss and everything inside the $\langle \rangle$ brackets	It could distorts the translation
\(\)	Parentheses indicating words/consonants mistakenly omitted by	Erased brackets, kept content	Translation carried out thanks to these additions

<sup>13</sup> Ibid.

<sup>14</sup> Given the small number of such cases, we considered acceptable the error that could have been created when these brackets did not identify exactly as described. Unfortunately, this part of the raw data would require a more detailed review which could not be performed due to lack of time and funds.

<sup>15</sup> At the end of this table, we have inserted the variations taken from the source code. We did this because listing all of them would have taken too much time and space.

<sup>16</sup> Same as the note 14.

Subject	Meaning	Managing	Reason
	the ancient scribe and		
	restored or corrected by		
	the publisher <sup>17</sup>		
{ }	Parenthesss indicating	Erased brackets, kept	Translation carried out
	consonants/words	content	thanks to these additions
	considered erroneous		
	and superfluous by the		
	publisher (same of the		
	source) <sup>18</sup>		
:	It indicates the causative	Erased	They could distort the
	and is placed between		translation
	the "s" and the rest of		
1	the verb	T 1	771 1111
.du	Used for dual	Erased	They could distort the
,du	Vaniation of	C1	translation
≡	Variation of =	Substituded with =	Uniform the signs with
-Lücke-	Gon	Erased	the same meaning It could distorts the
Lücke	Gap	Liascu	translation
Lucke	Match two or more	Replaced with a space	For better understanding
	words of the	replaced with a space	of the single word by the
	transliteration		machine
+	Used with numbers (in	Erased	It could distorts the
	this way $[n+]$ or $[+n]$ , or		translation
	to indicate the union of		
	two terms in a single		
	lemma, which are		
	separated by a suffixed		
	pronoun (for example		
	rn+=f+nfr Sn.w)		
!	Used to indicate a	Erased	It could distorts the
	sentence that when		translation
	translated turns out to be		
	an exclamation	D 1	T. 111'
Ø	It probably indicates a	Erased	It could distorts the
	suffixed pronoun that		translation
	should be present but is		
	not (and it is not chosen		
	which one it might be), or it indicates an unclear		
	suffixed pronoun		
( )	Parentheses indicating a	Erased brackets, kept	They could distort the
	Cartouche	content	translation
:	Probably used to indicate	Erased	It could distorts the
	that the preceding sign is	214004	translation
	not present in the		- WILDIWITOII
	<u> </u>		

<sup>17</sup> Höper J. et al., *op. cit.*, p. 106. 18 *Ibid*.

Subject	Meaning	Managing	Reason
	transliteration		
Präp.	It probably indicates a	Erased	It could distorts the
	preposition		translation
( )	Parentheses cindicating	Erased brackets, kept	They could distort the
	names of fortresses	content	translation
ť'	The superscript is	Substituded superscript	It could distorts the
jmj-r'	probably a typing error	with '	translation
	to indicate '		
,	Probably a typing error	Substituded with '	It could distorts the
	to indicate '		translation
(£ Double spaces again	Residues from cleaning	Reduced to a single	They could distort the
£)	operations	space	translation

Table 2: Transliteration managing.

```
Variations of \{.*?\}\ \langle .*?\rangle\ e\ \langle .*?\rangle\ \{.*?\}
text = re.sub('(\langle \{.*? \rangle \} \setminus \{.*? \rangle \}) \setminus (\langle .*? \rangle \setminus (.*? \rangle )', r' \setminus 1', text)
text = re.sub('(\langle .*? \rangle \s\\langle .*? \rangle) \s(\langle .*? \rangle \s\langle .*? \rangle)', r'\2', text)
text = re.sub('(\langle \{.*? \rangle \} \setminus \langle (.*?) \rangle', r' \mid 1', text)
text = re.sub('\setminus (.*?)\setminus s(\setminus \{.*?\setminus \}\setminus \{.*?\setminus \})', r'\setminus 2', text)
text = re.sub('(\langle .*? \rangle) \setminus (\langle .*? \rangle \cdot , r' \rangle', r' \wedge 2', text)
text = re.sub('(\langle .*? \rangle \) \langle .*? \rangle) (\langle .*? \rangle)', r' 1', text)
text = re.sub('(\langle .*? \rangle s.*? \langle .*? \rangle) (\langle .*? \rangle)', r' \langle 1', text \rangle)
text = re.sub('(\langle .*? \rangle s.*? \langle .*? \rangle) (\langle .*? \rangle)', r' 1', text)
text = re.sub('\setminus \{(.*?)\setminus \} \lceil \land s \rceil \backslash \langle (.*?)\setminus \rangle', r'\setminus 1', text)
text = re.sub(' \backslash (.*?) \backslash [^ \s] \backslash \{(.*?) \backslash \}', r' \backslash 2', text)
text = re.sub('\setminus\{(.*?)\setminus\}\setminus s[^\setminus s]\setminus (.*?)\setminus ', r'\setminus 1', text)
text = re.sub(' \langle (.*?) \rangle \langle (.*?) \rangle', r' \langle 2', text \rangle
text = re.sub('(\{.*?\})[a-zA-Z0-9HhHhhh?'\underline{dt}s.,:=\underline{i}]+)\backslash\langle(.*?)\rangle', r'\backslash 1', text)
text = re.sub('(\langle .*? \rangle)(\langle a-zA-Z0-9 \underline{H}h\underline{H}\underline{h}\underline{h}^{3},\underline{dt}s,..=\underline{i}] + \langle .*? \rangle)', r' \langle 2', text\rangle
text = re.sub('(\{.*?}\)\s[a-zA-Z0-9HhHhhh]'d\underline{t}\check{s}, := \underline{i}] + \s) \setminus \langle (.*?) \setminus \rangle', r' \setminus 1', text)
text = re.sub('(\(\cdot\)^*?\))(\s[a-zA-Z0-9HhHhh}'dt\check{s},:=i]+\s(.*?\)', r'\2', text)
text = re.sub('(\(\langle .*? \rangle\)([a-zA-Z0-9\.H\.h\.H\.h\.H\.h\.h.'s'\underline{dt}\.s.,:=\underline{i}] + \(\langle .*? \rangle\)', \ r'\(2', \ text)
text = re.sub('(\\{.^*?\\}\s[a-zA-Z0-9HhHhhh]''\underline{dt}'', := \underline{i}] + s[a-zA-Z0-9HhHhhhh]''\underline{dt}'', := \underline{i}] + ) \langle (.^*?) \rangle', r' \wedge 1', text)
```

```
text = re.sub('(\langle .*? \rangle)(\langle s[a-zA-Z0-9HhHhh]')'dts., :=i] + \langle s[a-zA-Z0-9HhHhh]'(dts., :=i] + \langle s[a-zA-Z0-9Hh]'(s[a-zA-Z0-9Hh]'(s[a-zA-Z0-9Hh]'(s[a-zA-Z0-9Hh]'(s[a-zA-Z0-9Hh]'(s[a-zA-Z0-9Hh]'(s[a-zA-Z0-9Hh]'(s[a-zA-Z0-9Hh]'(s[a-zA-Z0-9Hh]'(s[a-zA-Z0-9Hh]'(s[a-zA-Z0-9Hh]'(s[a-zA-Z0-2]'(s[a-zA-Z0-9Hh]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[a-zA-Z0-2]'(s[
```

- $text = re.sub('(\langle \{.*? \rangle \} [a-zA-Z0-9HhHhhh]'\underline{dt}\check{s}.,:=i] + \langle [a-zA-Z0-9HhHhhhh]'\underline{dt}\check{s}.,:=i] + \langle [a-zA-Z0-9HhHhhhh]'\underline{dt}\check{s}.,:=i] + \langle [a-zA-Z0-9HhHhhh]'\underline{dt}\check{s}.,:=i] + \langle [a-zA-Z0-9HhHhhh]'\underline{dt}\check{s}.,:=i] + \langle [a-zA-Z0-9HhHhhh]'\underline{dt}\check{s}.,:=i] + \langle [a-zA-Z0-9HhHhhh]'\underline{dt}\check{s}.,:=i] + \langle [a-zA-Z0-9Hh]'\underline{dt}\check{s}.,:=i] + \langle [a-zA-Z0-9Hh]'\underline{dt},:=i] + \langle [a-zA-Z0-9Hh]'\underline{dt}$
- $text = re.sub('(\(\.^*?\))([a-zA-Z0-9\.Hh\.H)\.h^3'\underline{dt}\.s.,:=i] + \s[a-zA-Z0-9\.Hh\.H).h^3'\underline{dt}\.s.,:=i] + \s[a-zA-Z0-9.Hh.H).h^3'\underline{dt}\.s.,:=i] + \s[a-zA-Z0-9.Hh.H).h^3'\underline{dt}\.s.,:=i] + \s[a-zA-Z0-9.Hh.H).h^3'\underline{dt}.s.,:=i] + \s[a-zA-Z0-9.Hh.H].h^3'\underline{dt}.s.,:=i] + \s[a-zA-Z0-9.Hh.H].h^3'\underline{dt}.s$
- $text = re.sub('(\langle .*? \rangle)([a-zA-Z0-9HhHhhh}'dt\check{s}...=i]+\s[a-zA-Z0-9HhHhh}'dt\check{s}...=i]+\s[a-zA-Z0-9HhHhhh}'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-9Hh]'dt\check{s}...=i]+\s[a-zA-Z0-$
- $text = re.sub('(\{.*?}\)[a-zA-Z0-9HhHhhh]''\underline{dt}'', := \underline{i}] + \s[a-zA-Z0-9HhHhhh]''\underline{dt}'', := \underline{i}] + \s[a-zA-Z0-9HhH]''\underline{dt}'', := \s[a-zA-Z0-9HhH]''\underline{dt}''$
- $text = re.sub('\((.*?)\)\((.*?)\)[a-zA-Z0-9HhHhhh''\underline{dt}\S.,:=\underline{i}] + \s[a-zA-Z0-9HhHhhh''\underline{dt}\S.,:=\underline{i}] + \s[a-zA-Z0-9HhHhh''\underline{dt}\S.,:=\underline{i}] + \s[a-zA-Z0-9HhH''\underline{dt}\S.,:=\underline{i}] + \s[a-zA-Z0-9HhH'''\underline{dt}\S.,:=\underline{i}] + \s[a-zA-Z0-9HhH'''\underline{dt}\S.,:=\underline{i}] + \s[a-zA-Z0-2HhH'''\underline{dt}\S.,:=\underline{i}] + \s[a-zA-Z0-2HhH'''\underline{dt}\S.,:=\underline{i}] + \s[a-zA-Z0-2Hh'''\underline{dt}\S.,:=\underline{i}] + \s[a-zA-Z0-2Hh'''\underline{dt}\S.,:=\underline{i}] + \s[a-zA-Z0-2Hh''''\underline{dt}\S.,:=\underline{i}] + \s[a-zA-Z0-2Hh'''\underline{dt}\S.,:=\underline{i}] + \s[a-zA-Z0-2Hh''''$
- $text = re.sub('(\{.*?}\) s[a-zA-Z0-9HhHhhh]''\underline{dt}'', r'\) (.*?) \ ($

- $text = re.sub('(\{.*?\}\s[a-zA-Z0-9HhHhhh!'`\underline{dt}\check{s}.,:=\underline{i}] + \s[a-zA-Z0-9HhHhhh!'`\underline{dt}\check{s}.,:=\underline{i}] + \s[a-zA-Z0-9HhHhhhh!'`\underline{dt}\check{s}.,:=\underline{i}] + \s[a-zA-Z0-9HhHhhhh!'`\underline{dt}\check{s}.,:=\underline{i}] + \s[a-zA-Z0-9HhHhhhh!'`\underline{dt}\check{s}.,:=\underline{i}] + \s[a-zA-Z0-9HhHhhhh!'`\underline{dt}\check{s}.,:=\underline{i}] + \s[a-zA-Z0-9HhHhhhh!''\underline{dt}\check{s}.,:=\underline{i}] + \s[a-zA-Z0-9HhHhhhh!''\underline{dt}\check{s}.,:=\underline{i}] + \s[a-zA-Z0-9HhHhhhh!''\underline{dt}\check{s}.,:=\underline{i}] + \s[a-zA-Z0-9Hh].$
- $text = re.sub('\((.*?)\)\((.*)\)\((.*)$
- $text = re.sub('\setminus \{(.*?)\setminus \}\setminus \langle (.*?)\setminus \rangle', r'\setminus 1', text)$
- $text = re.sub('\setminus (.*?)\setminus -\setminus \{(.*?)\setminus \}', r'\setminus 2', text)$
- $text = re.sub('\setminus\{(.*?)\setminus\}\setminus -\langle(.*?)\setminus\}', r'\setminus 1', text)$
- $text = re.sub('\setminus \langle (.*?) \setminus \rangle \setminus \{(.*?) \setminus \}', r'\setminus 2', text)$
- $text = re.sub('\setminus \{(.*?)\setminus \}\setminus \langle (.*?)\setminus \rangle', r'\setminus 1', text)$
- $text = re.sub(' \setminus \langle (.*?) \setminus \rangle \setminus \{ (.*?) \setminus \}', r' \setminus 2', text)$

Table 3: Hieroglyphs coded in Gardiner code (Graphics)

Subject	Meaning	Managing	Reason
(£ Double spaces £)	Typing or data collection	Reduced to a single	It could distorts the
	error	space	translation
{m1}⟨S29⟩	Relating this to transliteration and translation of the datapoint, the hieroglyphic phrase containing it is too short. However no destruction of any kind is marked. Phrase ID: IBUBd7rPt2l7i0zahXbH yAelst0	Erased the entire sentence	It could distorts the translation
geschrieben	"Written".	Erased the entire	It could distorts the
SandhiForm	It was not clear what the second word referred to. Phrase ID: IBUBdzTACHPAGEaRu TT9RVGCmrk	sentence	translation since nothing of this kind should be found among the hieroglyphs
Detvon	The preposition "von" is	Erased the entire	It could distorts the
erhalten	probably meant.  "Receive".  Phrase ID:  IBUBdxFybGGQx0kahq  N2gGSVZNU	sentence	translation since nothing of this kind should be found among the hieroglyphs
// (£ as the only characters	Destruction	Erased the entire sentence	It could distorts the translation
£)			
"sic"	Term normally used in philology	Erased	It could distorts the translation
"var" "Var" "var." -var vae	Variant	Erased	They could distort the translation
+lvar+s	In Jsesh it is equivalent to the larger "var" lettering	Erased	It could distorts the translation
+linverted+s	In Jsesh it is equivalent to the larger "inverted" text	Erased	It could distorts the translation
"ein Vogel"	"A bird"	Replaced with / identifying a single destroyed sign	Standardize signs with similar meaning
"unleserliches Zeichen"	"Illegible sign"	Replaced with / identifying a single	Standardize signs with similar meaning

Subject	Meaning	Managing	Reason
V	<u> </u>	destroyed sign	
"lb" " lb"	Line break	Erased	They could distort the translation
"lb "b"			
"hierat"	Hieratic	Erased	They could distort the translation
"monogr" "monogram"	Monogram	Erased	They could distort the translation
"Spuren"	Traces	Erased	It could distorts the translation
"large"	It probably indicates the size of the sign	Erased	They could distort the translation
"hiero"	Hieroglyphic	Erased	It could distorts the translation
"mutil"	Mutilated	Erased	They could distort the translation
"composite"	Composite sign	Erased	They could distort the translation
"traces"	Traces of the sign	Erased	They could distort the translation
"senkrechte Zeichenspur"	"Vertical trace"	Erased	They could distort the translation
"senkrechtes Zeichen"	"Vertical sign"	Erased	They could distort the translation
**	Between one sign and another it identifies the overlap of the two	Erased	It could distorts the translation
*	In Jsesh separates two signs in juxtaposition <sup>19</sup>	Erased	It could distorts the translation
//	In Jsesh it is equivalent to the destroying of an entire square	Replaced with / identifying a single destroyed sign	Standardize and simplify signs with similar meaning
h/	In Jsesh it is equivalent to a horizontal half destruction of a square	Replaced with / identifying a single destroyed sign	Simplify signs used to indicate a destruction
v/	In Jsesh it is equivalent to a vertical half destruction of a square	Replaced with / identifying a single destroyed sign	Simplify signs used to indicate a destruction
#b-/#e	In Jsesh, if the damaged part covers more than one sign, or passage, it can be inserted between #b "shading beginning" and #e "shading end".  The / stands for a single	Replaced with / identifying a single destroyed sign	Simplify signs

<sup>19</sup> Van den Berg, Manuel de Codage, <<u>http://www.catchpenny.org/codage/</u>> (last viewed 28 July 2023).

St	ubje	ect	Meaning	Managing	Reason
	J		destroyed sign. In this	<del> </del>	
			case it represents only a		
			vertical destruction (like		
			"v/")		
	-:		In Jsesh, used in between	Erased	It could distorts the
	:		two signs, places one		translation
			above and one below the		
ГО		07	same squared space	T 11 1 1 1	771 1
[?		?]	In Jsesh they are	Erased brackets, kept	The translation was
			equivalent to brackets [	content	made thanks to these reconstructions
			<sup>1</sup> , i.e. partially corrupted content <sup>20</sup>		reconstructions
"		"?"	Parentheses of question	Erased brackets, kept	The translation was
	<b>¿</b> "		marks inside quotation	content	made thanks to these
			marks. Content of		additions
			dubious reading <sup>21</sup>		
i.č		<b>""?"</b> "	Parentheses of question	Erased brackets, kept	The translation was
			marks inside quotation	content	made thanks to these
			marks and single quotes.		additions
			Content of dubious		
ГГ		11	reading In Jsesh they are	Erased brackets, kept	The translation was
]]		]]	equivalent to a single	content	made thanks to these
			square parentheses.	Content	reconstructions
			Consonants/words		reconstructions
			missing from the original		
			text due to a gap and		
			reconstructed by the		
			publisher <sup>22</sup>		
[{*[{*		*}]*}]	In Jsesh they are	Erased brackets, kept	The translation was
[{-[{-		-}]-}]	equivalent to double	content	made thanks to these
			curly brackets, i.e.		reconstructions
F / ·		40.7	ancient destructions <sup>23</sup>	T 11 1 1	771
[{*		*}]	In Jsesh they are	Erased brackets, kept	The translation was
[{-		-}]	equivalent to a single	content	made thanks to these
			curly bracket, i.e.		reconstructions
			consonants/words considered erroneous		
			and superfluous by the publisher <sup>24</sup>		
[[*[[*	_	*]]*]]	In Jsesh they are	Erased brackets, kept	The translation was
-]]-]]		-]]-]]	equivalent to double	content	made thanks to these
			square brackets. Ancient		reconstructions

<sup>20</sup> Höper J. et al., op. cit., p. 106.

*Ibid*.

*Ibid*.

*Ibid*.

*Ibid*.

Subj	ect	Meaning	Managing	Reason
3		reconstructions over	8 8	
		cancellations <sup>25</sup>		
[(-[(-	-)]-)]	In Jsesh they are	Erased brackets, kept	The translation was
	7-7-	equivalent to double	content	made thanks to these
		parentheses. Old		additions
		additions to the text		
		(above the line or next to		
		the column) <sup>26</sup>		
(	)	In Jsesh these brackets	Erased brackets, kept	It could distorts the
		identify a subgroup <sup>27</sup> , or	content	translation
		indicates incorrectly		
		spelled		
		consonants/words or		
		abbreviations that have		
		been expanded by the		
Φ.	Φ.	publisher <sup>28</sup>	TT 1'C' 1	A 1 1 1
\$	\$	Parentheses indicating the use of transliteration	Unmodified	Already erased
		writing. Present in sentences that had		
		already been eliminated		
		because they contained		
		"Detvon", "erhalten",		
		"geschrieben" and		
		"SandhiForm"		
<1-0>-	-<0-2>	In Jsesh it is equivalent	Erased brackets, kept	It could distorts the
<1-	-2>	to a Cartouche, where 1	content	translation
-<1	[	stands for the beginning		
<2-	-1>	and 2 for the end. The 0		
<0-	-0>	stands for Cartouche		
		with no beginning or		
		end. The combined use,		
		i.e. the first one reported,		
		creates a Cartouche that		
		has only a beginning and		
		an end but without		
		connecting lines in the		
		middle. The -<1 is		
		probably a typing error	Engad brackets from	It could distorts the
<-	-> >	In Jsesh it is equivalent to a Cartouche	Erased brackets, kept	translation
##		In Jsesh it is equivalent	content Erased	It could distorts the
""	-	to making the two	Liascu	translation
		connected signs overlap		u ansianon
		connected signs overlap		

*Ibid*.

*Ibid*.
27 Van den Berg, *Manuel de Codage*, <a href="http://www.catchpenny.org/codage/">http://www.catchpenny.org/codage/</a>> (last viewed 28 July 2023).

<sup>28</sup> Höper J. et al., op. cit., p. 106.

Subject	Meaning	Managing	Reason
V	In Jsesh rotates some	Erased	It could distorts the
	signs (for example Aa11)		translation
	placing them vertically		
SS	Probably a typing error.	Replaced with S29	It could distorts the
	Phrase ID:		translation
	IBUBd6UJB382c0qdtPC		
	8EZHyaSA		
	In the object ID (lKey:		
	74950) within the TLA		
	shows the S29 sign		
nn	Equivalent to the	Replaced with M22-M22	It could distorts the
	combined sign of M22-	1	translation
	M22		
"lc"	It was not clear what he	Erased	It could distorts the
	was referring to		translation
prwn	Probably a mistake.	Replaced with O1. We	It could distorts the
	Phrase ID:	opted to avoid adding	translation
	IBUBd3vVMr7QFUd2n	E34 for safety	
	QHSiajyOoQ		
	The chiest ID (IV ev		
	The object ID (lKey: 93610) within the TLA		
	shows the O1 sign but		
	the E34 sign is absent		
rf	Probably a mistake. In	Replaced with D21-I9	It could distorts the
	transliteration it is	1	translation
	marked " $r = f$ ".		
	Phrase ID:		
	IBUBd8HQMlgcwkG3h		
	j0xc9pBWkk		
ZeA	Probably a typing error.	Replaced with Z2A	It could distorts the
	In transliteration it is a		translation
	plural.		
	Phrase ID:		
	IBUBd47Tu9KnIk4pgFp		
i	92h7qHOI	Panlaged with M17	It could distorts the
j	Probably a mistake. In transliteration it is the "j"	Replaced with M17	translation
	in the word "jw".		a anotanon
	Phrase ID:		
	IBUBd6tQoo05aEBko1		
	T8tdGWVlY		
y1	Probably a typing error.	Replaced with Y1	It could distorts the
	In the transliteration it is		translation
	part of "hpr.w" and is		
	found as determinative,		
	so it should be a capital		

Sul	bject	Meaning	Managing	Reason
		y.		
		Phrase ID:		
		IBUBd08T8AjvLE7ElW		
		Ez0Ugub64		
	z2	Probably a typing error.	Replaced with Z2	It could distorts the
		In transliteration it is a		translation
		plural.		
		Phrase ID:		
		IBUBdWM48uxFmkjW		
		v63AUS5oAvs		
	b1	Probably a typing error.	Replaced with B1	It could distorts the
		It is part of a group of		translation
		determinatives that		
		identify women.		
		Phrase ID:		
		IBUBd8m6lxq5ekjykeK		
	О.	Fj8oWwqg	D 1 1 1 1 F22	Tr 1.1.1' r r.1
1	pS	Probably a typing error.	Replaced with F22	It could distorts the
		It is part of a group of		translation
		signs whose		
		transliteration is "pht,t".		
		They probably wanted to write "pH".		
		Phrase ID:		
		IBUBd2sC7dUkyEX9q9		
		zlgGB9WhE		
		In Jsesh it indicates a	Erased	It could distorts the
	_	space, so it's probably a	Liuova	translation
		typing error		
""	?h"	Found within a series of	Erased	It could distorts the
		destroyed signs, which		translation
		the transliteration does		
		not take into account,		
		and without matching		
		either the object ID		
		(lKey) or the pos.		
		Phrase ID:		
		IBcCNJxk89Yz4U3vv78		
		dRlppnnc		
	!	In Jsesh it is equivalent	Erased	It could distorts the
		to the starting of a new		translation
		line		
	"	Typing error or cleaning	Erased	It could distorts the
		residue		translation
[&[&	&]&]	In Jsesh it is equivalent	Erased brackets, kept	The translation was
		to the parentheses $\langle \langle \rangle \rangle$ .	content	made thanks to these
	0.7	Haplography <sup>29</sup>	T 11 1 1	reconstructions
[&	&]	In Jsesh it is equivalent	Erased brackets, kept	The translation was

Subject	Meaning	Managing	Reason
J	to the parentheses $\langle  \rangle$ . Consonants/words mistakenly omitted by the ancient scribe, reconstructed or	content	made thanks to these reconstructions
	corrected by the author <sup>30</sup>		
&	In Jsesh it is equivalent to placing the signs within the same square space, not necessarily overlapping and often close together (for example in "z³-r'")	Erased	It could distorts the translation
- - (£ even more than one £)	Separation of each hieroglyphic sign placed next to another. If you find more than one it is due to cleaning residues or typing errors	Erased	This way the machine can learn on its own when a word begins and ends
\\R(£ number £)	In Jsesh it is equivalent to rotating the previous sign, by the same degrees of the numeric value, counterclockwise	Erased	It could distorts the translation
\R(£ number £)	In Jsesh it is equivalent to rotating the previous sign, by the same degrees of the numeric value, clockwise	Erased	It could distorts the translation
\(£ number £)	After a hieroglyph it identifies the dimensions of it according to the numerical value	Erased	It could distorts the translation
( )	Parentheses indicating the Cartouche	Erased brackets, kept content	It could distorts the translation
\	In Jsesh it is equivalent to mirroring the previous sign horizontally	Erased	It could distorts the translation
V205	A sign that in Jsesh it is equivalent to nothing	Kept	On the TLA website it is marked like this, while the transliteration is a _ sign (we marked here in case you decide managing otherwise)
(£ spaces at the beginning and end of the	Residues from cleaning operations	Erased	They could distort the translation

*Ibid*.

*Ibid*.

Subject	Meaning	Managing	Reason
sentence £)			
(£ Double spaces again	Residues from cleaning	Reduced to a single	They could distort the
£)	operations	space	translation

Table 3: Gardiner code managing.

Table 4: Part of speech (Word class)

Subjec t	Meaning	Managing	Reason
/	lKey mismatch	We have not considered the entire datapoint	It may conduit to a translation error. In some cases, it indicated the "" while in others it did not
			exist as an object

Table 4: Part of speech managing.

## References

HÖPER Jakob, KUPREYEV Maxim, SCHWEITZER Simon D., SPERVESLAGE Gunnar, HOEDT Sebastian, HYPSZER Joanna Maria, JONES Ellen, Berlin Text System 3.1 User Manual (V 3.0), 2018.

VAN DEN BERG Hans, "Manuel de Codage" A standard system for the computerencoding of Egyptian transliteration and hieroglyphic texts, < Manuel de Codage (catchpenny.org) >, (last viewed il 28/07/2023).