

Human Evaluation of model ALL

After the execution of experiments and analyzing the results, we tested the effectiveness of the model HIERO-TRANSFORMER¹. We trained it to translate 9 data pairs listed in Table 1.

Source	Target
Hieroglyphics (ea)	German (de)
Hieroglyphics (ea)	English (en)
Hieroglyphics (ea)	Transliteration (tnt)
Hieroglyphics (ea)	Lemma ID (lKey)
Hieroglyphics (ea)	<i>Wordclass</i> (wordClass)
Transliteration (tnt)	German (de)
Transliteration (tnt)	English (en)
Transliteration (tnt)	Lemma ID (lKey)
Transliteration (tnt)	<i>Wordclass</i> (wordClass)

Table 1: List of data pairs in which the model ALL trained.

The evaluation of HIERO-TRANSFORMER took place in three distinct phases.

1. First, we subjected sentences of progressively increasing grammatical complexity, taken from Gardiner's grammar².
2. Second, we submitted a few passages from four literary works (The Eloquent Peasant³, the Tale of Sinuhe⁴, the Tale of the Shipwrecked Sailor⁵ and the Admonitions of Ipuwer⁶).
3. Third, we stressed the model with a long passage from the Story of Sinuhe.

For each path we divided the evaluation in two waves based on the inputs we entered. In the first wave, the input was presented to the model as Hieroglyphics (Gardiner code), in the second wave, it was presented as transliteration. Results of both waves have been collected in tables to compare better the differences and similarities between the model prediction and the control sentence. We underline that Lemma IDs and POS tags could be found within the TLA⁷ database using the search tool. Thanks to this it is also possible to obtain other information on Lemmata, such as attestations in the *corpus*, reference bibliography and related Lemmata.

-
- 1 For the Automatic Evaluation of the model and the comparison with other models, you can find the complete results in the paper at: <<https://aclanthology.org/2024.ml4al-1.9/>>.
 - 2 Gardiner A. H., *Egyptian Grammar, being an introduction to the study of hieroglyphs*, Oxford, 1957.
 - 3 [The Eloquent Peasant] Original documents: pBerlin 3023 and pAmherst 1, pBerlin 3025 and pAmherst 2, pButler (BM 10274), pRamesseum A (Berlin 10499, recto). Published in Gardiner and Vogelsang (1908). Italian translation and commentary in Bresciani (1969). English translation and commentary in Allen (2015). German translation and commentary in Vogelsang (1913).
 - 4 [Tale of Sinuhe] Original documents: pBerlin 3022 and pAmherst n-q, pRamesseum A (Berlin 10499, verso), pUCL 32106C, pUCL 32106C, papyrus from the Natural History Museum of Buenos Aires. Text in hieroglyphic transcription in Blackman (1972). Italian translation and commentary in Bresciani (1969). English translation and commentary in Allen (2015). German translation and commentary in Grapow (1952).
 - 5 [Tale of the Shipwrecked Sailor] Original document: pHermitage 1115. Published in [Golenischeff]. Text in hieroglyphic transcription in Blackman (1972). Italian translation and commentary in Bresciani (1969). English translation and commentary in Allen (2015).
 - 6 [Admonitions of Ipuwer] Original document: pLeiden I, 344 recto. Published in hieroglyphic transcription, translated and commented in English in Gardiner (1969). Italian translation and commentary in Bresciani (1969).
 - 7 Thesaurus Linguae Aegyptiae <<https://thesaurus-linguae-aegyptiae.de>> (accessed: 2 august 2024).

When inserting the Gardiner Code, the data was cleaned following the same cleansing operations scheme used to clean the training data⁸. In particular, we erased any element that did not correspond to the sole alphanumeric codes representing the hieroglyphs, such as "-" or "*".

For the transliteration, we followed a method similar to the one implemented for the Gardiner code: we erased superfluous characters and adopted the same conventions for inserting transliteration used by TLA⁹. This convention differs in some formal elements from the most commonly adopted systems. Adopting it permitted us to have no differences between the texts we transliterated and the ones we submitted to the model.

Before subjecting to the model, each sentence was studied by the author: the hieroglyphic texts were transliterated, translated and grammatically analysed. However, for the transliteration, we applied the same convention used by the TLA.

It results that, from Hierogliphys and transliteration, HIERO-TRANSFORMER is capable of generating:

- German translation;
- English translation;
- Transliteration;
- Lemma IDs;
- Part-of-speech tags (POS tags).

When possible, we compared the model predictions with the work of other publications or the author of this work. We collected all results produced by the model in tables to improve reading and make easier comparisons with the correct answers.

⁸ You can find the list of all these operations in the file “[cleansing_operations.pdf](#)”.

⁹ Here <<https://github.com/mattia-decao/hiero-transformer>>, you can find a list of tips while providing to HIERO-TRANSFORMER hieroglyphic or transliteration inputs.

1 – First phase: Exercise from Gardiner, *Egyptian Grammar*

Here, we subjected to HIERO-TRANSFORMER various exercises taken from Gardiner's grammar. These examples were of progressively increasing grammatical complexity.

For the first example¹⁰ we analyzed:



R'w m p.t

Re is in the sky

It is a simple non-verbal sentence with an adverbial predicate. The subject *R'w* is followed by the adverbial predicate *m p.t*.

The sequence of hieroglyphics in Gardiner code is as follows:

D21 D36 N5 G17 Q3 X1 N1

By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:

source → target	Model prediction	Check
ea → de	Re am Himmel	/
ea → en	Re in the heavens	Re is in the sky
ea → tnt	R m p,t	R'w m p.t
ea → lKey	96700 400007 58710	400015 400007 58710
ea → wordClass	gods_name preposition substantive_fem	gods_name preposition substantive_fem

Table 2: Phase 1, example 1. Prediction from hieroglyphics in Gardiner code and comparison.

With Gardiner code as input (Table 2), you can see that the model predicted an almost entirely correct outcome. There are some differences in the translation, in which it reports “heavens” instead of “sky”, and a couple of errors: the first in the transliteration, R instead of *R'w*; the second in the lemma ID (lKey), 96700 instead of 400015.

The translations “sky” and “heavens” are both possible, so it may not be considered an error, but just a variation of the translation.

¹⁰ Gardiner, *Egyptian Grammar* cit., p. 36.

The imprecision in the transliteration is not very incisive since the model recognizes the capital letter, and therefore understands that it is a proper name. Furthermore, in both translations and POS tags, the model can understand that it is the name of a deity, specifically Re. However, there is a difference regarding the error on the lemma ID. Here, the model, instead of the ID that should represent Re^{w} , provided us with the ID that represents rdj , which is the irregular verb with the meaning of "to put, to give". Two factors may lead to this error: firstly there could be a consistent presence of this verb concerning Re, and consequently, the model could find it difficult to separate the two meanings; secondly, it is an extremely short sentence, and therefore the model would have difficulty finding a context, although it is a quite simple phrase.

By inserting the transliteration *Re m p.t* as input, we obtained the following results:

source → target	Model prediction	Check
ea → de	Re im Himmel	/
ea → en	Re in the sky	Re is in the sky
ea → IKey	400015 400007 58710	400015 400007 58710
ea → wordClass	gods_name preposition substantive	gods_name preposition substantive_fem

Table 3: Phase 1, example 1. Prediction from transliteration and comparison.

Here in Table 3, we can observe an improvement compared to the previous result. The English translation “heavens” has become “sky”, and there is only one error in the POS tags (wordClass).

Once again, “sky” and “heavens” represent a possible translation difference and therefore can be considered variations.

The error in the POS tags is related to *p.t* that the model does not write. As we will see later, such a lack occurs more frequently as the length of the input provided to the model increases. In the case of this sentence, however, we found that the model differentiates the use of punctuation for endings. In fact, after a quick test in which we inserted transliteration *p,t* instead of *p.t*, the model recognized it as "substantive", even if it did not specify its gender.

The second example¹¹ we analyzed was:



wbn Re m h.t

Re rises in the horizon

11 Gardiner, *Egyptian Grammar* cit., exercise II.1, p. 37.

It is a simple verbal sentence. The verb *wbn* is in first position, in *sdm=f* form, followed by the subject *R'w* and ends with the preposition *m* and the noun *ḥ.t*.

The sequence of hieroglyphics in Gardiner code is as follows:

G43 D58 N35 N8 D21 D36 N5 A40 G17 N27 X1 Z1



By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:


source → target	Model prediction	Check
ea → de	Re im Horizont	/
ea → en	Re in the horizon	Re rises in the horizon
ea → tnt	wbn R m,t	wbn R'w m ḥ.t
ea → IKey	45030 400015 400007 227	45050 400015 400007 227
ea → wordClass	verb_3-lit gods_name preposition substantive_fem	verb_3-lit gods_name preposition substantive_fem

Table 4: Phase 1, example 2. Prediction from hieroglyphics in Gardiner code and comparison.

With Gardiner code as input (Table 4), you can see that the model makes one error in the translation, one in the lemma IDs and a few in the transliteration.

In the English translation, we can observe that the verb is not recognized, probably because during training there may have been more adverbial sentences relating to Re than verbal ones.

In the transliteration,  *R'w* is reported as *R*, while it seems to not recognize  *m ḥ.t* considering that it writes *m,t*.

Among lemma IDs, 45030 is incorrect. In fact, this is equivalent to  *wb'yt*, that is “Aufwärterin; female butler; servant”¹², while it should have reported 45050.

In any case, most of the model results are correct, especially regarding POS tags.

By inserting the transliteration *wbn R'w m ḥ.t* as input, we obtained the following results:

source → target	Model prediction	Check
ea → de	Re wird am Horizont aufgehen	/
ea → en	Re appears in the horizon	Re rises in the horizon
ea → IKey	45050 400015 400007 182390	45050 400015 400007 227
ea → wordClass	verb_3-lit gods_name preposition	verb_3-lit gods_name preposition substantive_fem

¹² TLA ID 45030, <<https://thesaurus-linguae-aegyptiae.de/lemma/45030>> (accessed: 18 october 2023). WB I, 292.7-8. The latter refers to the text *Wörterbuch der ägyptischen Sprache*, where WB refers to the name of the text itself, and the first number identifies the volume. Then, there is the page number, often followed by a dot and a number that identifies the term inside the page. For any dictionary entry, usually, there is more than one term, as it is possible to translate it in multiple ways depending on the context.

The lemma ID 182390 is equivalent to $\mathbb{A}\mathbb{A}\mathbb{A}\mathbb{A} d^3 d^3 .t$, that is “Behörde; Kollegium; (Toten-)Gericht; assembly; council; magistrates (at the judgment of the dead); assemblies; advice; tribunal”¹³. We observed that this error is also reflected in the POS tags, where the writing of the lemma is absent. Similar to the difference we found between $p.t$ and p,t in the first sentence of this phase, we further check the insertion of $^3h.t$. Again, the model differentiates the use of punctuation, generating a more accurate prediction using the comma rather than the dot. In fact, by submitting the sentence $wn R'w m ^3h,t$, the IDs are correct; “substantive” appears in the POS tags, imprecisely reported without the feminine gender; while the translation “May Re appear in the horizon” can be considered a variation of the correct meaning.

You are my servant

The sequence of hieroglyphics in Gardiner code is as follows:

14 Gardiner, *Egyptian Grammar* cit., exercise III.8, p. 44.

source → target	Model prediction	Check
ea → de	Meine Dienerin	/
ea → en	My servant	You are my servant
ea → tnt	jw = <u>t</u> m bk,t=j	jw = <u>t</u> m b ³ k.t=j
ea → lKey	851515 174900 53830 10030	851515 10120 500292 53870 10030
ea → wordClass	particle personal_pronoun substantive_fem personal_pronoun	particle personal_pronoun preposition substantive_fem personal_pronoun

Table 6: Phase 1, example 3. Prediction from hieroglyphics in Gardiner code and comparison.

With Gardiner code as input (Table 6), you can see inaccuracies and small errors in every field. These are all linked by the fact that the model failed to understand $\Rightarrow \text{𓂏}$ as two separate terms: in the translation, it does not recognize the first part of the sentence, and in fact, it is omitted; in the transliteration, $=\underline{t}m$ is erroneously reported instead of $=\underline{t} m$; in the lemma IDs, is reported 174900 ($\Rightarrow \text{𓂏} \underline{t}w$), i.e. the dependent pronoun of second person masculine singular, instead of 10120 500292 ($\Rightarrow \text{𓂏} =\underline{t} m$), i.e. suffix pronoun of second person feminine singular plus preposition; finally, in the POS tags, we note the absence of "preposition" which should indicate $\text{𓂏} m$.

Furthermore, the model makes one inaccuracy reporting the lemma ID 53830 ($\text{𓂏} \text{𓂏} b^3k$), instead of 53870 ($\text{𓂏} \text{𓂏} b^3k.t$), thus only getting the gender wrong.

By inserting the transliteration $wbn jw =\underline{t} m b^3k.t=j$ as input, we obtained the following results:

source → target	Model prediction	Check
ea → de	Du bist mein Diener	/
ea → en	You are my servant	You are my servant
ea → lKey	21881 10120 500292 53820 10030	851515 10120 500292 53870 10030
ea → wordClass	particle personal_pronoun preposition verb personal_pronoun	particle personal_pronoun preposition substantive_fem personal_pronoun

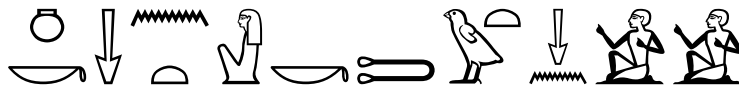
Table 7: Phase 1, example 3. Prediction from transliteration and comparison.

Here in Table 7, we can observe that the errors previously generated are almost entirely resolved, but we observe that the model mistake a POS tag: while referring to $\text{𓂏} \text{𓂏} b^3k.t$, it reports "verb" instead of "substantive_fem". Among the lemma IDs it writes a variation of a lemma and imprecisely reports another one.

For what concerns lemma IDs, the model proposes 21881 instead of 851515, but it is the same term, i.e. jw . Instead, he makes an inaccuracy with 53820 ($\text{𓂏} \text{𓂏} b^3k.w$), that is “Arbeit; Arbeitsleistung; tasks; labor ; travail; redevance”¹⁵, while it should report 53870 ($\text{𓂏} \text{𓂏} b^3k.t$).

¹⁵ TLA ID 53820, <<https://thesaurus-linguae-aegyptiae.de/lemma/53820>> (accessed: 18 october 2023). WB I, 427.13-428.5.

The fourth example¹⁶ we analyzed was:



jnk sn.t =k twt sn =j

I am your sister and you are my brother

This sentence comprises two nominal clauses, both of A B type. The subject of the first clause is the independent pronoun *jnk* (first person singular), while the nominal predicate is *sn.t=k*; the second clause is similar to the first, in fact, we find the subject *twt*, an archaic form of the dependent pronoun of the second person masculine singular¹⁷, followed by the nominal predicate *sn =j*.

The sequence of hieroglyphics in Gardiner code is as follows:

W24 V31 T22 N35 X1 B1 V31 V13 G43 X1 T22 N35 A1 A1

By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:

source → target	Model prediction	Check
ea → de	Ich war deine Schwester, meine Schwester	/
ea → en	I was your sister, my brother	I am your sister and you are my brother
ea → tnt	<i>jnk sn,t =k tw,t sn =j</i>	<i>jnk sn.t =k twt sn =j</i>
ea → lKey	27940 136260 10110 174900 136230 10030	27940 136260 10110 175050 136230 10030
ea → wordClass	personal_pronoun substantive_fem personal_pronoun verb substantive_masc personal_pronoun	personal_pronoun substantive_fem personal_pronoun personal_pronoun substantive_masc personal_pronoun

Table 8: Phase 1, example 4. Prediction from hieroglyphics in Gardiner code and comparison.

With Gardiner code as input (Table 8), you can see inaccuracies in all fields, all due to the misunderstanding of $\Rightarrow \text{𓏏} \text{𓏏} \text{𓏏}$ *twt*: in the English translation, the model understands that these are two sentences, in fact, it divides them with a comma, but the second one lacks the subject and the copula; the transliteration reports *tw,t* instead of *twt*; in the IDs of the lemmas we find 174900 ($\Rightarrow \text{𓏏} \text{𓏏}$ *tw*), i.e. a dependent pronoun of the second person masculine singular, instead of 175050 ($\Rightarrow \text{𓏏} \text{𓏏} \text{𓏏}$ *twt*).

¹⁶ Gardiner, *Egyptian Grammar* cit., exercise V.5, p. 57.

¹⁷ Gardiner, *Egyptian Grammar* cit., p. 53.

tw); and finally, in POS tags is reported “verb” instead of “personal_pronoun”. Possibly, since $\Rightarrow \text{𓂏} \text{ } tw$ is an archaic form, the model may have received less input than the pronoun $\Rightarrow \text{𓂏} \text{ } tw$, so the model considers it more plausible to translate it this way.

By inserting the transliteration *jnk sn.t =k twt sn =j* as input, we obtained the following results:

source → target	Model prediction	Check
ea → de	Ich bin deine Schwester, du bist mein Bruder	/
ea → en	I am your brother; you are my brother	I am your sister and you are my brother
ea → lKey	27940 136260 10110 175050 136230 10030	27940 136260 10110 175050 136230 10030
ea → wordClass	personal_pronoun substantive_fem personal_pronoun personal_pronoun substantive_masc personal_pr	personal_pronoun substantive_fem personal_pronoun personal_pronoun substantive_masc personal_pronoun

Table 9: Phase 1, example 4. Prediction from transliteration and comparison.

Here in Table 9, we can observe that the mistake with *tw* is no longer committed, because the model understands the difference between the transliteration *tw* and *tw*, but we can observe other inaccuracies.

In the English translation, we can note that the model cannot distinguish the noun’s gender since it reports "brother" instead of "sister".

In the POS tags, the model does not finish the writing of “personal_pronoun”, getting stuck at “personal_pr”.

All other results generated by the machine are correct.

The fifth example¹⁸ we analyzed was:



n nḏnḏ =j ḥn' ḏw qd

I do not ask advice to (lit. with) one of evil character

This is a negative verbal sentence. The negation *n*, followed by the *sḏm* =*f* form of the verb, introduces a negation which usually acquires a past value. Then follows the subject in suffixed pronoun and the preposition ḥn'. The last two terms, namely the adjective *ḏw* and the noun *qd*, is a *nfr ḥr* construction.

The sequence of hieroglyphics in Gardiner code is as follows:

¹⁸ Gardiner, *Egyptian Grammar* cit., exercise VIII.1, p. 84.

By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:

source → target	Model prediction	Check
ea → de	Ich war nicht zerschnitten, ich war entblößte	/
ea → en	I did not defame the disgraceful	I do not ask advice to one of evil character
ea → tnt	n nn =j hn,w qd =j	n ndnd =j hn' dw qd
ea → IKey	78890 854526 10030 854526 10030 162430 10030	850806 91600 10030 106200 182840 162430
ea → wordClass	particle verb_2-lit personal_pronoun preposition substantive_masc personal_pronoun	particle verb_4-lit personal_pronoun preposition adjective substantive_masc

Table 10: Phase 1, example 5. Prediction from hieroglyphics in Gardiner code and comparison.

With Gardiner code as input (Table 10), you can see various errors and inaccuracies in all fields.

In the English translation, although the model understands the past tense, it reports “defame” and “disgraceful”, demonstrating that it has not understood the fundamental points of the sentence.

In the transliteration, we find *nn* instead of *ndnd*, *hn,w* instead of *hn'* and *dw*, and finally *qd =j* instead of *qd*.

In the lemma IDs, the model reports 78890 instead of 850806, but it is the same lemma (particle of negation 𓂏); or the others, are correct only 10030, i.e. the first person singular pronoun =j, and 162430 (𓂏𓂏𓂏 *qd*).

In the POS tags we find that the model does not recognize *ndnd* as “verb_4-lit”, and reports “verb_2-lit”; finally it reports “substantive_masc personal_pronoun” instead of “adjective substantive_masc”.

The presence of so many errors is indicative of an incapacity to understand the sentence. Some errors, however, are understandable: when the model encountered 𓂏𓂏𓂏, it may have thought that 𓂏 was itself a determinative while staying between one determinative and another did not make the model change its translation; conversely the presence of 𓂏𓂏𓂏 could lead the model to translate it as a noun that supports a suffixed pronoun.

By inserting the transliteration *n ndnd =j hn' dw qd* as input, we obtained the following results:

By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:

source → target	Model prediction	Check
ea → de	Ich kenne dich, ich kenne deinen Namen	/
ea → en	You know me, I know your name	I know you, I know your name
ea → tnt	r.kwj tw r.kwj rn =k	rh.kw tw rh.kw rn =k
ea → lKey	95620 44000 174900 95620 44000 94700 10110	95620 174900 95620 94700 10110
ea → wordClass	verb_2-lit personal_pronoun personal_pronoun verb_2-lit personal_pronoun personal_pronoun substanti	verb_2-lit personal_pronoun verb_2-lit substantive_masc personal_pronoun

Table 12: Phase 1, example 6. Prediction from hieroglyphics in Gardiner code and comparison.

With Gardiner code as input (Table 12), you can see that the model recognizes most of the sentence, but makes inaccuracies in all fields.

In the English translation, we observe that it reverses the meaning of the first sentence, so instead of “I know you”, it translates “You know me”. You can better understand the reasons for this error by looking at the lemma IDs. Here, there is the ID 44000 (𓂏𓂏 wj), which represents the dependent pronoun of the first person singular, corresponding to the ending that identifies the stative. Thus, we can infer that the model does not intend the first clause to be stative, but rather a verbal clause with dependent pronouns used as subjects (which is grammatically incorrect).

In the POS tags, the error is similar to the one just analyzed. We find “personal_pronoun” twice after the recognizing of the first verb, which would be attributable to recognizing 𓂏𓂏 wj instead of 𓂏𓂏 kw. In the second clause, in addition to making mistakes in the same way, the model swaps the two POS tags referring to 𓂏𓂏, as it reports “personal_pronoun substanti” (without completing “substantive_masc”) instead of the correct “substantive_masc personal_pronoun”.

By inserting the transliteration *rh.kw tw rh.kw rn =k* as input, we obtained the following results:

source → target	Model prediction	Check
ea → de	Ich kenne dich und ich kenne deinen Namen	/
ea → en	I know you, I know your name	I know you, I know your name
ea → lKey	95620 174900 95620 94700 10110	95620 174900 95620 94700 10110
ea → wordClass	verb_2-lit personal_pronoun verb_2-lit substantive_masc personal_pronoun	verb_2-lit personal_pronoun verb_2-lit substantive_masc personal_pronoun

Table 13: Phase 1, example 6. Prediction from transliteration and comparison.

Here in Table 13, we can observe that the model does not make errors in any field of the prediction.

This result is extremely satisfactory because, even if in the previous examples we observed that the results are better with transliteration input instead of the Gardiner code, usually few inaccuracies persist or new ones are made. In this case, however, not only does the model correct its errors, but it manages to provide an exact result.

The seventh example²⁰ we analyzed was:



jw =j hr hrp kʔ.t nb.t

I lead every work

This is a pseudo-verbal sentence. It is introduced by the particle *jw*, which supports the subject *=j*, while the pseudo-verbal construction is represented by *hr hrp*. At the end, we find the noun *kʔ.t* modified by the adjective *nb.t*.

The sequence of hieroglyphics in Gardiner code is as follows:

M17 Z7 A1 D2 Z1 Aa1 D21 Q3 A24 D28 X1 A9 V30 X1

By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:

source → target	Model prediction	Check
ea → de	Ich beherrsche jede Arbeit	/
ea → en	I was in every work	I lead every work
ea → tnt	jw =j hr rp k,t nb.t	jw =j hr hrp kʔ.t nb.t
ea → lKey	851515 10030 107529 120150 163010 81660	851515 10030 107529 120150 163010 81660
ea → wordClass	particle personal_pronoun preposition verb_3-lit substantive_fem adjective	particle personal_pronoun preposition verb_3-lit substantive_fem adjective

Table 14: Phase 1, example 7. Prediction from hieroglyphics in Gardiner code and comparison.

²⁰ Gardiner, *Egyptian Grammar* cit., exercise XXIII.2, p. 256.

With Gardiner code as input (Table 14), you can see a couple of small imperfections attributable to an error in understanding a term.

In the English translation, we observe that the model does not seem to consider 𐎧𐎥𐎷𐎫 *hrp*, while it translates 𐎧 *hr* in its variant “in”. For the same term, the lack of translation of the verb could be linked to the failure to write the consonant *h* in the transliteration. Furthermore, in the transliteration field, *k,t* is found instead of *kʔ.t*.

Regardless of the error in the English translation and the imperfections in the transliteration, we observe that lemma IDs and POS tags are correct. Here, the model understood the sentence, while it had little difficulty translating and transliterating correctly.

By inserting the transliteration *jw =j hr hrp kʔ.t nb.t* as input, we obtained the following results:

source → target	Model prediction	Check
ea → de	Ich leite jede Arbeit	/
ea → en	I lead every work	I lead every work
ea → lKey	851515 10030 107529 120150 163010 81660	851515 10030 107529 120150 163010 81660
ea → wordClass	particle personal_pronoun preposition verb_3-lit substantive_fem adjective	particle personal_pronoun preposition verb_3-lit substantive_fem adjective

Table 15: Phase 1, example 7. Prediction from transliteration and comparison.

Here, in Table 15, we can observe that, as in the previous example, the model also repairs the errors generated when receiving the Gardiner code input but generates a completely correct response. This result is in line with the previous example and outlines the hypothesis that the model is not sensitive to variations in grammatical complexity (given that we found more errors in simple sentences compared to more complex ones), while it is sensitive to the inputs received during training. The more similar and numerous the single terms were, the more the model understood small variations in meaning. On the contrary, as the rarity of the lemmas increased, the model’s inability to recognize them also increased.

2 – Second phase: Literary works

In the second testing phase, we subjected to HIERO-TRANSFORMER various Middle Egyptian literary works. Similarly to the previous phase, we chose the passages that contained different grammatical elements. Here, the increasing of complexity was due to the length of the passages, instead of the grammar.

For the first example, taken from *The Admonitions of Ipuwer*²¹, we analyzed:



$m = tn \ sm^{\prime}.tw \ z \ r \ gs \ sn = f$

Look, a man is slain next to his brother

We found a difference between the transcriptions of Gardiner's and Enmarch's publications, in particular, Gardiner transcribed the $\downarrow sn$ hieroglyph with his variation \downarrow . In our work we preferred to report Enmarch transcription work because it is more recent (which was collected within the TLA database).

It is a verbal sentence. The particle $m.tn$ introduces the verb sm^{\prime} , in the $s\dot{d}m.tw = f$ form. The suffix tw is used to make the passive. The subject is z and is followed by the adverbial clause introduced by $r \ gs$.

The sequence of hieroglyphics in Gardiner code is as follows:

G17 D36 X1 N35 Z2 S29 U1 G1 G17 Z7 T30 D40 X1 Z7 O34 A1 Z1 D21 Aa13 Z1 T22 A1 I9

By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:

21 In particular: pLeiden I, 344, 9.3. Gardiner, *The Admonitions of an Egyptian Sage*, p. 67. Enmarch, *Sentence ID IBUBdxnGXSW0507nkxucVzKGEJs*, <https://thesaurus-linguae-aegyptiae.de/sentence/IBUBdxnGXSW0507nkxucVzKGEJs>, in: *Thesaurus Linguae Aegyptiae* (accessed on 10 October 2023).

source → target	Model prediction	Check
ea → de	Siehe, ein Mann wird an die Seite seines Bruders getötet	/
ea → en	Look, a man should be killed at his brother	Look, a man is slain at his brother's side ²²
ea → tnt	m =tn sm.tw z r gs sn =f	m =tn sm ³ .tw z r gs sn =f
ea → lKey	64440 172370 134370 125010 92390 136230 10050	64440 172370 134370 125010 92390 136230 10050
ea → wordClass	particle_nonenclitic personal_pronoun verb_3-lit substantive_masc preposition substantive_mas	particle_nonenclitic personal_pronoun verb_3-lit substantive_masc preposition substantive_masc personal_pronoun

Table 16: Phase 2, example 1. Prediction from hieroglyphics in Gardiner code and comparison.

With Gardiner code as input (Table 16), you can see that the differences between the results and the correct answer vary little. Towards English, there are few differences, towards transliteration it misses the ³ of sm³, and finally it is unable to complete the writing of the POS tags.

As for the English translation, we find “should be killed” instead of “is slain”, furthermore the fact that the action takes place next to the brother is missing, and in fact it is translated “at his brother” instead of “his brother's side”. The first difference is due to the fact that the model does not recognize the ending ^{ae} tw as passive, and mistakenly adds the modal verb “should”; the second reflects the absence of translation for the term r gs “beside; next to”.

The lack of a ³ in the transliteration should not be considered very incisive, given that the corresponding ID of lemma 134370 is correct.

Finally, we can observe an incompleteness in the POS tags: “substantive_mas” is not wrote completely and “personal_pronoun” is missing. This type of error, as we will also see in the following examples, occurs more often in medium and long sentences where the input inserted is greater. However, in cases like this, it is also found in shorter sentences. It seems as if the prediction stopped just before completing the analysis.

By inserting the transliteration m =tn sm³.tw z r gs sn =f as input, we obtained the following results:

source → target	Model prediction	Check
ea → de	Siehe, ein Mann wird neben seinem Bruder getötet	/

²² Enmarch, *Sentence ID IBUBdxnGXSW0507nqxucVzKGEJs*, <<https://thesaurus-linguae-aegyptiae.de/sentence/IBUBdxnGXSW0507nqxucVzKGEJs>>, in: *Thesaurus Linguae Aegyptiae* (accessed on 10 October 2023).

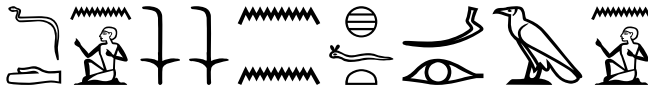
ea → en	Look, a man is killed next to his brother	Look, a man is slain at his brother's side ²³
ea → lKey	64440 172370 134370 125010 92390 136230 10050	64440 172370 134370 125010 92390 136230 10050
ea → wordClass	particle_nonenclitic personal_pronoun verb_3-lit substantive_masc preposition substantive_mas	particle_nonenclitic personal_pronoun verb_3-lit substantive_masc preposition substantive_masc personal_pronoun

Table 17: Phase 2, example 1. Prediction from transliteration and comparison.

Here in Table 17, we can observe that almost all previous errors have disappeared. The English translation is correct as are the lemma IDs; the only problem is found once again in the POS tags, in which the writing does not end. In fact, as the last part, the model reports "substantive_mas" without completing it and leaving out "personal_pronoun". This error is similar to the one just observed in the Gardiner code results of this example.

Overall, the results provided by the model are correct, we can consider it very positive.

For the second example, taken from *The Admonitions of Ipuwer*²⁴, we analyzed:



dd.n =j nn hft m³.n =j

I said this in accordance with (what) I saw

It is a verbal sentence in the *sdm.n =f* form. The preposition *hft* introduces the adverbial clause which contains another verb in the *sdm.n =f* form.

The sequence of hieroglyphics in Gardiner code is as follows:

I10 D46 N35 A1 M22 M22 N35 N35 Aa1 I9 X1 U2 D4 G1 N35 A1

By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:

-
- 23 Enmarch, *Sentence ID IBUBdxnGXSW0507nxxucVzKGEJs*, <<https://thesaurus-linguae-aegyptiae.de/sentence/IBUBdxnGXSW0507nxxucVzKGEJs>>, in: *Thesaurus Linguae Aegyptiae* (accessed on 10 October 2023).
- 24 In particular: pLeiden I, 344, 9.3. Gardiner, *The Admonitions of an Egyptian Sage*, p. 67. Enmarch, *Sentence ID IBUBdxnGXSW0507nxxucVzKGEJs*, <<https://thesaurus-linguae-aegyptiae.de/sentence/IBUBdxnGXSW0507nxxucVzKGEJs>>, in: *Thesaurus Linguae Aegyptiae* (accessed on 10 October 2023).

source → target	Model prediction	Check
ea → de	Ich habe dies gesagt, weil ich gesehen habe	/
ea → en	I said this to the one I had seen	I have said this in accordance with what I have seen ²⁵
ea → tnt	d.n =j nn ft m.n =j	dd.n =j nn hft m³.n =j
ea → lKey	185810 10030 84570 116770 66270 10030	185810 10030 84570 116770 66270 10030
ea → wordClass	verb_2-lit personal_pronoun dimostrative_pronoun preposition verb_2-gem personal_pronoun	verb_2-lit personal_pronoun dimostrative_pronoun preposition verb_2-gem personal_pronoun

Table 18: Phase 2, example 2. Prediction from hieroglyphics in Gardiner code and comparison.

With Gardiner code as input (Table 18), you can see that, although the model comes close to the correct translation, in many places it generates errors that change the meaning of the translation. Moreover, it leaves out several consonants in the transliteration.

In the English translation, the differences between “said” and “have said”, and between “had seen” and “have seen”, can be considered variations since the form *sdm.n =f* does not point out a tense, but a completed action (perfective aspect). However, the difference between “to the one I had seen” and “in accordance with” modifies the meaning of the sentence.

In the transliteration, we note that the model does not write the *d* of *dd.n*, the *h* of *hft* and the *ʾ* of *m³.n*. Anyway, the translation is affected only in correspondence with the term *hft*, whose error modifies its meaning, while it does not affect either the lemma IDs (lKey) nor the POS tags (wordClass). In fact, in IDs and POS tags we can observe that the model provides the correct prediction.

By inserting the transliteration *dd.n =j nn hft m³.n =j* as input, we obtained the following results:

source → target	Model prediction	Check
ea → de	Ich habe dies gesagt, gemäß dem was ich gesehen habe.	/
ea → en	I said this according to what I had seen.	I have said this in accordance with what I have seen ²⁶
ea → lKey	185810 10030 84570 116770 66270 10030	185810 10030 84570 116770 66270 10030
ea → wordClass	verb_2-lit personal_pronoun demonstrative_pronoun preposition verb_2-gem personal_pronoun	verb_2-lit personal_pronoun dimostrative_pronoun preposition verb_2-gem personal_pronoun

Table 19: Phase 2, example 2. Prediction from transliteration and comparison.

²⁵ Gardiner, *The Admonitions of an Egyptian Sage* cit., p. 99.

²⁶ *Ibid.*

Here in Table 19, we can observe that the model translated everything extremely well. The only difference can be found in the English translation, which can be considered a translation variation.

For the third example, taken from *The Eloquent Peasant*²⁷, we analyzed:



jn jw =k r z n nhḥ

Will you be a man of eternity?

This is an adverbial sentence. The particle *jn* introduces the question followed by the particle *jw*, which supports the suffix pronoun with function of subject. The preposition *r* introduces the adverbial clause, which here function as a clause of future identity (what one is destined or led to do).

The sequence of hieroglyphics in Gardiner code is as follows:

M17 N35 M17 Z7 V31 D21 O34 A1 Z1 N35 N35 G21 V28 V28 N5

By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:

source → target	Model prediction	Check
ea → de	Bist du ein Mann der Ewigkeit?	Wirst du nicht ein mann der ewigkeit sein? ²⁸
ea → en	Would you be a man of eternity?	are you to be eternal? ²⁹
ea → tnt	<i>jn jw =k r z n nhḥ</i>	<i>jn jw =k r z n nhḥ</i>
ea → IKey	26680 851515 10110 91907 125010 79800 86570	26680 851515 10110 91907 125010 79800 86570
ea → wordClass	particle particle personal_pronoun preposition substantive_masc nisbe_adjective_preposition substantive_	particle particle personal_pronoun preposition substantive_masc nisbe_adjective_preposition substantive_masc

Table 20: Phase 2, example 3. Prediction from hieroglyphics in Gardiner code and comparison.

²⁷ In particular: pBerlin 3023, 95 (Vogelsang), 126 (Allen). Vogelsang, *Kommentar zu den Klagen des Bauern* cit., p. 88. Allen, *Middle Egyptian Literature* cit., p. 260.

²⁸ Vogelsang, *Kommentar zu den Klagen des Bauern* cit., p. 89.

²⁹ Allen, *Middle Egyptian Literature* cit., p. 261.

With Gardiner code as input (Table 20), you can see that the model translates in a very good way. However, there is a small difference in the English translation and an incompleteness in the POS tags.

As for the English translation, we can consider it a variation. The word "would" is used to introduce a polite invitation or request.

In the POS tags, however, we can observe an incompleteness: the model is unable to write "substantive_masc" in its entirety.

By inserting the transliteration *jn jw =k r z n nh* as input, we obtained the following results:

source → target	Model prediction	Check
ea → de	Bist du ein Mann der Ewigkeit?	Wirst du nicht ein mann der ewigkeit sein? ³⁰
ea → en	Would you be a man of eternity?	are you to be eternal? ³¹
ea → lKey	26680 851515 10110 91907 127770 78875 86570	26680 851515 10110 91907 125010 79800 86570
ea → wordClass	particle particle personal_pronoun preposition undefined nisbe_adjective_preposition substantive_masc	particle particle personal_pronoun preposition substantive_masc nisbe_adjective_preposition substantive_masc

Table 21: Phase 2, example 3. Prediction from transliteration and comparison.

Here in Table 21, we can observe that the translation into both languages is not different from the previous results, but there are changes both in the lemma IDs and the POS tags. This time we can note the completion of the POS tags.

Lemma ID 127770 (*𓂏𓂛 sj*) is incorrect, in fact this corresponds to the dependent pronoun of the third person feminine singular. This error is probably due to the fact that the phonetic distinction between *s* and *z* in Middle Egyptian tends to be reduced and the two spellings to overlapped³². Consequently, we thought that, during training, the model alternately learned these two transliterations as if they were the same. By submitting one of these two graphemes, therefore, the model could exchange them. We find another difference in the lemma ID 78875. It is always *𓂏𓂛 n*, but its meaning is “innerhalb von (temporal); in (Zeitraum); [temporal]; within; in (temp.)”³³ and differs from 79800, i.e. “von [Genitiv]; belonging to; [genitives]; de”³⁴.

In the POS tags there is now an "undefined" instead of "substantive_masc" corresponding to the ID that was incorrectly recognized, i.e. 127770 (*𓂏𓂛 sj*).

30 Vogelsang, *Kommentar zu den Klagen des Bauern* cit., p. 89.

31 Allen, *Middle Egyptian Literature* cit., p. 261.

32 Allen, *Middle Egyptian* cit., p. 19.

33 TLA ID 78875, <<https://thesaurus-linguae-aegyptiae.de/lemma/78875>>, in: *Thesaurus Linguae Aegyptiae* (accessed on 10 October 2023).

34 TLA ID 79800, <<https://thesaurus-linguae-aegyptiae.de/lemma/79800>>, in: *Thesaurus Linguae Aegyptiae* (accessed on 10 October 2023).

For the fourth example, taken from The Eloquent Peasant³⁵, we analyzed:



jn nn r =f dj =k swʹ =n hr wʹt

So, won't you let us pass on the road? (lit. Won't you make sure sure that shall we pass on the road?)

This is a negative interrogative verb sentence formed by the *rdj sdm =f* construction. The particle *jn* is used to introduce the question, while the particle *nn* is used to mark the negation of the *sdm =f* form verbs in future tense. The *rdj sdm =f* construction is very common and is used to form causative, although the presence of the *s* in *swʹ* already makes it so. The sentence ends with the preposition *hr*, followed by the noun *wʹt*.

The sentence is translated in the future tense because the use of *nn* as the negation of a verb in the *sdm =f* form usually gives the sentence a future meaning.

The sequence of hieroglyphics in Gardiner code is as follows:

M17 N35 D35 N35 D21 I9 D36 V31 S29 V4 G1 Z9 N31 N35 Z2 D2 Z1 N31 Z1 X1

By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:

source → target	Model prediction	Check
ea → de	Wirst du nicht zulassen, daß wir auf diesem Weg weggegangen sind?	/
ea → en	Would you not allow us to go on this road?	So, won't you let us pass on the path? ³⁶
ea → tnt	<i>jn nn r =f i =k sw =n hr w,t</i>	<i>jn nn r =f dj =k swʹ =n hr wʹt</i>
ea → lKey	26680 84550 28170 10050 550028 10110 129740 10070 400090 42490	26680 84550 28170 550028 10110 129740 10070 400090 42490
ea → wordClass	particle particle particle_enclitic verb_irr personal_pronoun verb_4-inf personal_pronoun preposition	particle particle particle_enclitic verb_irr personal_pronoun verb_4-inf personal_pronoun preposition substantive_fem

Table 22: Phase 2, example 4. Prediction from hieroglyphics in Gardiner code and comparison.

³⁵ In particular: pBerlin 3023, 8-9 (Vogelsang), 39-40 (Allen). Vogelsang, *Kommentar zu den Klagen des Bauern* cit., p. 45. Allen, *Middle Egyptian Literature* cit., p. 239.

³⁶ Allen, *Middle Egyptian Literature* cit., p. 241.

With Gardiner code as input (Table 22), you can see that the model agree with the correct answer, but we find some differences and some errors. For English translation the model set the sentence without considering the enclitic particle $\overline{\text{r}} = f$ and without correctly setting the future tense. It made some errors in the transliteration and left out a POS tag (wordClass).

The errors in English translation are probably made because the model did not have the opportunity to train on enough sentences in which the $nn sdm = f$ form was present to be able to translate it into the future tense. The same can be hypothesized for the absence of translation of the enclitic particle.

In the transliteration there are three differences compared to the human answer: the first is \dot{j} instead of dj , an inaccuracy probably made because the model understood $\overline{\text{r}}dj$ as $rd\dot{j}$ instead of dj , and it only reported its weak radical \dot{j} ; the second is $sw\dot{i}$ instead of sw^3 , probably because the transliterated term $sw^3\dot{i}^{37}$ is present in the TLA database, where the data for training was extracted; third and final difference is w^3t instead of w^3t . All three were considered not very significant given that the lemma IDs are correct, i.e. 550028 for the first, 129740 for the second and 42490 for the third. The last important consideration to make concerns the insertion of the ID 10050 ($\overline{\text{r}} = f$) by the model. This, in fact, is technically incorrect, since it is part of the enclitic particle $\overline{\text{r}} = f$, but the data received from the model, and extracted from the TLA database, often reports this particle together with the personal pronoun. As a result, the model learned in this way. However, you can note that this problem does not arise in the POS tags.

The last error occurred precisely in the POS tags, in which you can see the absence of the term "substantive_fem" which refers to $\overline{\text{r}} = f w^3t$.

By inserting the transliteration $jn nn r = f dj = k sw^3 = n hr w^3t$ as input, we obtained the following results:


source → target	Model prediction	Check
ea → de	Bist du nicht derjenige, der uns den Weg überquert?	/
ea → en	Do you not allow us to pass on the way?	So, won't you let us pass on the path? ³⁸
ea → lKey	26680 84550 28170 10050 51510 10110 129740 10070 400090 42490	26680 84550 28170 550028 10110 129740 10070 400090 42490
ea → wordClass	particle demonstrative_pronoun particle_enclitic personal_pronoun verb_3-inf personal_pronoun verb	particle particle particle_enclitic verb_irr personal_pronoun verb_4-inf personal_pronoun preposition substantive_fem


Table 23: Phase 2, example 4. Prediction from transliteration and comparison.

37 TLA ID 129740, <<https://thesaurus-linguae-egyptiae.de/lemma/129740>>, in: *Thesaurus Linguae Aegyptiae* (accessed on 10 October 2023).

38 Allen, *Middle Egyptian Literature* cit., p. 241.

Here in Table 23, we can observe that the model makes some errors in lemma IDs and POS tags.

In the lemma IDs, 51510 ( *wḏi*) is found instead of 550028, because the model did not include the verb *dj*. This error is also found in the POS tags, in which the same lemma is reported "verb_3-inf".

In the POS tags, we can see that the negation *nn* ("particle") is exchanged with "demonstrative_pronoun", an imprecision given by the training of the model itself. It is probable that it encountered the demonstrative pronoun more times than the negation, and therefore, without the context of the hieroglyphics, the odds of defining it that way are increased. We then observe that, again, ID 10050 ( =f) is reported.

This time, the error of the personal pronoun that we have just observed among in the lemma IDs is found in the POS tags too: the "personal_pronoun" appears when nothing should be found. Furthermore, "verb_3-inf" is reported instead of "verb_4-inf", and "verb" instead of "preposition", while the writing of "substantive_fem" is completely omitted.

For the fifth example, taken from *The Tale of the Shipwrecked Sailor*³⁹ we analyzed:



h³.kw r wḏ wr

I went down towards the sea (lit. big green)

This is a sentence in which *.kw* ending identifies the verb form as stative.

The sequence of hieroglyphics in Gardiner code is as follows:

D54 V31 G43 A1 D21 M14 G36 D21 N36

By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:

source → target	Model prediction	Check
ea → de	Zu einem Weihrauch kam ich	/
ea → en	I came to the sea	I went down to the sea ⁴⁰
ea → tnt	jwḏ.kw r w wr	h³.kw r wḏ wr
ea → lKey	21930 91901 43820	97350 91901 43820
ea → wordClass	verb_irr preposition substantive	verb_3-inf preposition substantive_masc

³⁹ In particular: pHermitage 1115, 25. Golenischeff, *op. cit.*, p. 11. Blackman, *op. cit.*, p. 42.

⁴⁰ Allen, *Middle Egyptian Literature* cit., p. 15.

Table 24: Phase 2, example 5. Prediction from hieroglyphics in Gardiner code and comparison.

With Gardiner code as input (Table 24), you can see that the model generated fairly accurate results, while some errors can be considered logically acceptable. For the same terms, English translation, transliteration, lemma ID and POS tags differ slightly from the correct answer, but this could be considered as a translation variation. However, in the transliteration and in the POS tags, it produced some inaccuracies.

In the transliteration, we can observe that the model interpreted 𓂏 as the verb $jw\text{𓂏}$ instead of $h\text{𓂏}$, and this is an understandable error. All the results produced follow the same line: the English translation “I came”; the lemma ID 21930; and finally the POS tag “verb_irr”. Although the inaccuracies, the model provided all the data to what it actually translated, so it should be considered a translation variation rather than an actual error.

Regarding the other inaccuracies, in the transliteration we find $w\text{ }wr$ instead of $w\text{𓂏} wr$. The model, however, recognizes the term since it translates it correctly, and also writes correctly the lemma ID.

In the POS tags, the “substantive” instead of “substantive_masc” is inaccurate.

By inserting the transliteration $h\text{𓂏}.kw\text{ }r\text{ }w\text{𓂏} wr$ as input, we obtained the following results:

source → target	Model prediction	Check
ea → de	Ich bin zum Großen Grünen hinabgestiege	/
ea → en	I went down to the Great Gate	I went down to the sea ⁴¹
ea → lKey	97350 91901 43820	97350 91901 43820
ea → wordClass	verb_3-inf preposition substantive	verb_3-inf preposition substantive_masc

Table 25: Phase 2, example 5. Prediction from transliteration and comparison.

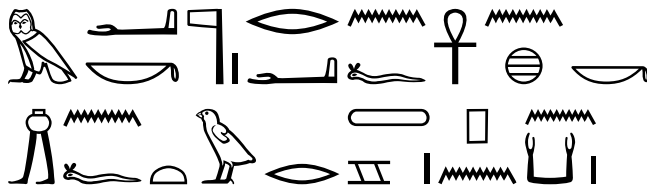
Here in Table 25, we can observe correctly recognized the verb $h\text{𓂏}$, reporting its ID and the corresponding POS tag. However, it also incorrectly translated $w\text{𓂏} wr$ into English, as well as it presented the same imprecision in the definition of its POS tag.

In the English translation, the model reported “Great Gate” instead of “sea” or “Great Green”, and although it is an error, it is interesting that it understood the adjective wr referring to that noun.

The lemma IDs here are correct, but we find the same inaccuracy encountered previously in the POS tags: the model is unable to define the masculine gender of $w\text{𓂏} wr$, reporting it simply as “substantive”.

41 Allen, *Middle Egyptian Literature* cit., p. 15.



For the sixth example, taken from *The Tale of the Shipwrecked Sailor*⁴² we analyzed:



m =k ntr rdj.n =f 'nh =k

jnṯ =f tw r jw pn n k'

Look, the god made sure that you live,
taking you to (lit. towards) this island of Ka

Between the publications of Blackman and Golenischeff there is a difference in the transcription of the seventh sign, that is : Blackman reports it this way, while Golenischeff reports . We decided to report Blackman's transcription because it is more recent.

This is an emphatic sentence that can be split into two units. The first unist is a verbal clause with the *sdm.n =f* form and is characterized by having the subject placed before the verb in order to focus attention (topicalization); after being introduced by the particle *m =k*, the subject is represented by the noun *ntr*, while the predicate is *rdj.n =f 'nh =k*, i.e. a *rdj sdm =f* construction introduced by *rdj* in the *sdm.n =f* form. The second unit, due to the context, is an implicit adverbial subordinate clause (circumstantial) that indicates a concomitant action. It is introduced by the verb *jnṯ* in the *sdm =f* form, and the subject *=f* point to the same third person subject of the previous clause. The preposition *r* is then used to introduce an adverbial clause.

We subjected this example to the model in two waves. The first, in which we maintained the subdivisions of the two units, and the second, in which we inserted the entire text.

The sequence of hieroglyphics in Gardiner code is as follows:

G17 D36 V31 R8 Z1 D21 D37 N35 I9 S34 N35 Aa1 V31

W25 N35 I9 X1 G43 D21 N18 N23 Z1 Q3 N35 N35 D28 Z1

By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:

⁴² In particular: pHermitage 1115, 113-115. Blackman, *op. cit.*, p. 44. Golenischeff, *op. cit.*, p. 17.

source → target	Uni No.	Model prediction – Separated units	Model prediction – Full text	Check
ea → de	1	Siehe, Gott gibt ihm dein Leben	Siehe, Gott hat ihm dein Leben gegeben,	/
	2	Er bringt dich zu dieser Insel für den Ka	nachdem er dich zu dieser Insel des Ka gebracht hat	
ea → en	1	Behold, God has given you life	Behold, God has made you live, so that he can bring you to this island of Ka	Look, the god, he has let you live by fetching you to this island of ka ⁴³
	2	He brought you to this island of Ka		
ea → tnt	1	m =k nṯr rī n =f n =k	m =k nṯr rī.n =f n =k jnī	m =k nṯr rdj.n =f 'nh =k jnī =f tw r jw pn n k ³
	2	jnī =f tw r jw pn n k	=f tw r jw pn	
ea → lKey	1	64440 10110 90260 851711 10050 38530 10110	64440 10110 90260 851711 10050 38530 10110 26870 10050 174900 91901 21	64440 10110 90260 851711 10050 38530 10110 26870 10050 174900 91901 21940 59920 79800 162870
	2	26870 10050 174900 91901 21940 59920 79800 162870		
ea → wordClass	1	particle_nonenclitic personal_pronoun substantive_masc verb_irr personal_pronoun substantive	particle_nonenclitic personal_pronoun substantive_masc verb_irr personal_pronoun substantive	particle_nonenclitic personal_pronoun substantive_masc verb_irr personal_pronoun verb_3-lit personal_pronoun verb_3-inf personal_pronoun personal_pronoun preposition substantive_masc demonstrative_pronoun nisbe_adjective_preposition substantive_masc
	2	verb_3-inf personal_pronoun personal_pronoun preposition substantive_masc demonstrative_pronoun nisbe		

Table 26: Phase 2, example 6. Prediction from hieroglyphics in Gardiner code and comparison.

With Gardiner code as input (Table 26), you can see many differences between the predictions and the response produced manually.

In the separated unit stage, $\text{𓂏} \text{𓂏}$ was incorrectly transliterated *n*, plus it was translated as a noun. This is in line with the corresponding POS tag, but in conflict with the lemma ID (lKey) 38530 which is a verb. Instead of *rdj.n =f* the model transliterated *rī n =f*, and the rest of the results are reported correctly, although 𓂏 of the second unit is only reported as *k*. In the POS tag section (wordClass), we see that in the first unit the model does not report "personal_pronoun" referring to the second person singular suffix pronoun $\text{𓂏} =k$, while in the second sentence the writing stops at "nisbe", omitting both its specification "nisbe_adjective_preposition" and the tag

43 Allen, *Middle Egyptian Literature*, p. 32.

“substantive_masc”, i.e. the tag referring to k^3 . Although there are several errors, the lemma IDs are generated perfectly and the translation is quite accurate.

In the full text stage, the model suffers from the effect of the input length. We can notice that: in the transliteration, although $rj.n = f$ is corrected by adding the dot, $n k^3$ is absent; the lemma IDs stop at 21, while it should be 21940 59920 79800 162870; finally, in the POS tags, in addition to incorrectly inserting “substantive” for $\text{𓆎} \text{𓆑}$, all the tags of the second sentence are missing.

However, for both stages we observe that the English translation is precise and provides a different version than the one produced manually.

By inserting the transliteration $m = k \text{ ntr rdj.n} = f \text{ 'nh} = k \text{ e jni} = f \text{ tw r jw pn n k}^3$ as input (first maintaining the unit subdivisions and then inserting the entire text), we obtained the following results:

source → target	Uni No.	Model prediction – Separated units	Model prediction – Full text	Check
ea → de	1	Siehe, Gott hat veranlaßt, daß du lebst	Siehe, Gott hat veranlaßt, daß du lebst, und er bringt dich zu dieser Insel für den Ka	/
	2	Er bringt dich zu dieser Insel für den Ka		
ea → en	1	Behold, God has left you to live	Behold, God has caused you to live by bringing you to this island of Ka	Look, the god, he has let you live by fetching you to this island of ka ⁴⁴
	2	He brings you to this island of Ka		
ea → lKey	1	64440 10110 90260 550028 10050 38530 10110	64440 10110 90260 550028 10050 38530 10110 26870 10050 174900 91901 21	64440 10110 90260 851711 10050 38530 10110 26870 10050 174900 91901 21940 59920 79800 162870
	2	26870 10050 174900 91901 21940 59920 400055 162870		
ea → wordClass	1	particle_nonenclitic personal_pronoun substantive_masc verb_2-lit personal_pronoun verb_	particle_nonenclitic personal_pronoun substantive_masc verb_2-lit personal_pronoun verb_	particle_nonenclitic personal_pronoun substantive_masc verb_irr personal_pronoun verb_3-lit personal_pronoun verb_3-inf personal_pronoun personal_pronoun preposition substantive_masc demonstrative_pronoun preposition nisbe_adjective_preposi tion substantive_masc
	2	verb_3-inf personal_pronoun personal_pronoun preposition substantive_masc demonstrative_pronoun preposition		

Table 27: Phase 2, example 6. Prediction from transliteration and comparison.

Here in Table 27, we can observe many differences between the predictions and the correct answer.

In the separated unit stage, the English translation has “left” instead of “let” or “caused”. In the lemma IDs, the difference between ID 550028 and ID 851711 of the first sentence, referring to *rdj*, can be considered as a translation variation, slightly enough that it can be ignored; on the other side, in the second unit, there is 400055, i.e. the preposition *n* used as a dative, instead of 79800, i.e. always *n* but used as a genitive. Finally, in the POS tags of the first sentence, the model incorrectly recognize *rdj* as “verb_2-lit” instead of “verb_irr”, and the writing is truncated to “verb_” leaving out “verb_3-lit personal_pronoun”; while in the second sentence, in line with the error made in the lemma IDs, the model reports "preposition" instead of "nisbe_adjective_preposition", but "substantive_masc" referring to *kʿ* is completely missing. Aside from small differences and errors, the English translation is very precise, as is the list of the lemma IDs.

In the full text stage, once again the model suffered from the input length. The initial part of the lemma IDs and POS tags are identical to the separated unit stage, but they cut off suddenly. Lemma IDs stop at 21 leaving out 21940 59920 79800 162870, while POS tags stop at “verb_” leaving out everything else. Apart from this, however, the correctness of the English translation is extremely satisfactory.

For the seventh example, taken from the *Story of Sinuhe*⁴⁵, we analyzed:



jr.t =j šm.t m ḥnt.yt

Mi feci strada (lit. Io feci l’andare) verso sud

I made my way (lit. I made the going) southward

It is a verbal sentence in which the verb *jr.t* is a narrative infinitive. The subject is *=j* and the object is *šm.t*, another verb in infinitive form. The sentence ends with the adverb *m ḥnt.yt*.

The sequence of hieroglyphics in Gardiner code is as follows:

D4 X1 A1 N40 G17 X1 D54 G17 W17 N35 X1 M17 M17 X1 P1

By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:

⁴⁵ In particular: pBerlin 3022, 5-6. Blackman, *op. cit.*, p. 8.

source → target	Model prediction	Check
ea → de	Mein Marsch nach Süden	Ich machte das Gehen südwärts ⁴⁶
ea → en	My journey to the south	My making off upstream ⁴⁷
ea → tnt	jṛ.t =j šṃ.t m nt,yt	jr.t =j šm.t m hnt.yt
ea → IKey	28550 10030 154340 850635	28550 10030 154340 850635
ea → wordClass	verb_3-inf personal_pronoun verb_3-inf adverb	verb_3-inf personal_pronoun verb_3-inf adverb

Table 28: Phase 2, example 7. Prediction from hieroglyphics in Gardiner code and comparison.

With Gardiner code as input (Table 28), you can see that the model produce almost exact predictions, the translation presents a variation, however correct, while in the transliteration, it reports a small inaccuracy and a couple of differences.

In the English translation we see “My journey” instead of “My making off”: the lack of recognition of the verb in the infinitive form could be considered as a less literal way of translating the sentence, therefore we considered it a translation variation rather than a mistake; furthermore, the model reports “to the south” instead of “upstream”, but both translations are possible for *m hnt.yt*⁴⁸.

Both verbs of the transliteration are presented with the weak radical *j*, and in fact we find *jṛ.t =j šṃ.t*, but this can be considered an inexactness rather than a mistake; also, the *h* is missing in *hnt.yt*. The lack of *h* is similar of that one we had already observed⁴⁹. It could therefore be hypothesized that model has some problems in recognizing the hieroglyph *ⲙ*, encoded with the Gardiner code Aa1.

Regardless of small differences and errors, it seems that the model understood the entire sentence well, as can also be seen from the correctness of lemma IDs and POS tags.

By inserting the transliteration *jr.t =j šm.t m hnt.yt* as input, we obtained the following results:

source → target	Model prediction	Check
ea → de	Mein Marsch nach Süden	Ich machte das Gehen südwärts ⁵⁰
ea → en	My journey to the south	My making off upstream ⁵¹
ea → IKey	28550 10030 154340 850635	28550 10030 154340 850635
ea → wordClass	verb_3-inf personal_pronoun verb_3-	verb_3-inf personal_pronoun verb_3-inf adverb

46 Grapow, *op. cit.*, p. 14.

47 Allen, *Middle Egyptian Literature* cit., p. 65.

48 TLA ID 850635, <<https://thesaurus-linguae-aegyptiae.de/lemma/850635>> (accessed on 23 October 2023). WB III, 309.15-18.

49 Second Literary example, p. 17.

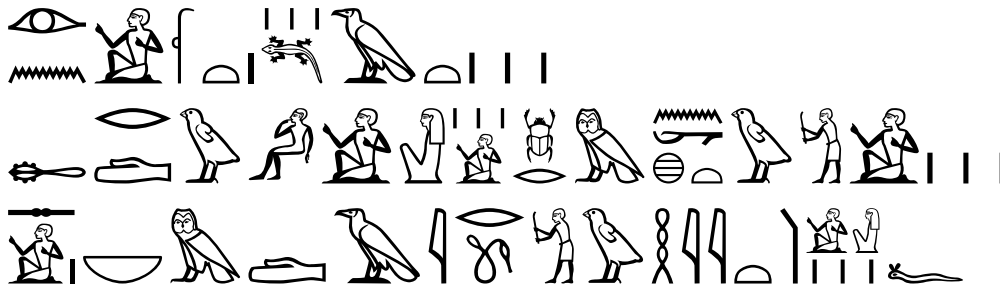
50 Grapow, *op. cit.*, p. 14.

51 Allen, *Middle Egyptian Literature* cit., p. 65.

Table 29: Phase 2, example 7. Prediction from transliteration and comparison.

Here in Table 29, we can observe that the same translations, lemma IDs and POS tags already analyzed have been proposed again. This implies that, at least for this sentence, the model has the same precision and understanding of the sentence, and therefore translates it in the same way regardless of whether the input is the Gardiner code or the transliteration.

For the eighth example, taken from the *Story of Sinuhe*⁵², we analyzed:



jrj.n =j rnp.wt 'š'.wt

hrd.w =j hpr m nht.w

z nb m d'r wh.jt =f

I have spent (lit. have done) many years,
my children have become strong,
and each of them (lit. every man) was the head of his own family.

This is a sentence made up of three clauses. The first is verbal in the *sdm.n =f* form. The second is composed of a pseudo-verbal construction with the subject *hrd.w* followed by the stative form of the verb, and therefore translated with a past value; the preposition *m* introduces *nht.w* (noun adjective) and is used to indicate identity/quality. The third is an adverbial clause, in which *z nb* is the subject and the preposition *m* (indicating quality/identity) introduces the adverbial part.

We subjected the texts to the model text in two steps. The first, in which the division into the three propositions was maintained, and the second, in which the entire text was presented.

The sequence of hieroglyphics in Gardiner code is as follows:

D4 N35 A1 M4 X1 Z1 Z2 I1 G1 X1 Z2

F32 D21 D46 G43 A17 A1 B1 Z2 A1 L1 D21 G17 N35 M3 Aa1 X1 G43 A24 A1 Z2

O34 A1 Z1 V30 G17 D46 G1 M17 D21 T12 A24 G43 V28 M17 M17 X1 T14 A1 B1 Z2 I9

⁵² In particular: pBerlin 3022, 92-94. Blackman, *op. cit.*, p. 24.

By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:

source → target	Uni No.	Model prediction – Separated units	Model prediction – Full text	Check
ea → de	1	Ich verbrachte viele Jahre	Ich verbrachte viele Jahre und meine	Ich verbrachte viele Jahre, Meine Kinder
	2	Meine Kinder sind zu Müdern geworden	Kinder, die entstanden sind durch	wurden Starke, Ein jeder war ein
	3	jedermann vernichtet seine Bedürfnisse	Zufriedenheit aller Männer, durch Bitternis seiner Kinder	Bezwinger seines Stammes ⁵³
ea → en	1	I spent many years in my life	I spent many years of my children, having	When I had spent many years, my boys grew
	2	My children have become weapons	become a strong man, and every man has a	into strongmen, each man his tribe's
	3	Every man has a desire for his tribe	strong man, and every man has a strong man	suppressor ⁵⁴
ea → tnt	1	jrj n =j rnp,t š.t	jrj.n =j rnp,t š.t hrd.w =j	jrj.n =j rnp.wt š'.wt
	2	hrd.w =j pr m nt,w	pr m n	hrd.w =j hpr m nht.w
	3	z nb m djr why,t =f		z nb m d'r wh.jt =f
ea → lKey	1	28550 400055 10030 94920 41011	851809 10030 94920 41011 5 854539	28550 10030 94920 41011 124480 10030
	2	124480 10030 116230 500292 87580	124480 10030 116230 500292 87620 12	116230 500292 87580 125290 500292 177740
	3	125010 81660 500292 177740 48730 10050		48730 10050
ea → wordClass	1	verb_3-inf personal_pronoun substantive_fem adjective	verb_3-inf personal_pronoun substantive_fem adjective verb_3-lit	verb_3-inf personal_pronoun substantive_fem adjective
	2	substantive_masc verb_3-lit preposition substantive	preposition substantive_masc	substantive_masc personal_pronoun verb_3-lit preposition
	3	substantive_masc adjective preposition verb_3-lit substantive_fem personal_pronoun		substantive_masc preposition verb_3-lit substantive_fem personal_pronoun




Table 30: Phase 2, example 8. Prediction from hieroglyphics in Gardiner code and comparison.



⁵³ Grapow, *op. cit.*, p. 44.

⁵⁴ Allen, *Middle Egyptian Literature cit.*, p. 93.

With Gardiner code as input (Table 30), you can see many differences and a few misunderstandings.

In the separated unit stage, the first clause presents *jri n* without the dot, *rnp,t* without the *w* of the plural, and *š.t* instead of *š'.wt*. Although these errors are not reflected in the English translation, we can note that the *n* detached from the verb has instead been recognized in the lemma ID 400055, i.e. the preposition *n* with the meaning of dative, while this addition does not appear in the POS tags. The second clause gives us *pr* and *nt,w*, both missing the *h*. Another difference can be noted in the POS tags where it does not recognize *=j* as "personal_pronoun".

The English translation, has an error: the model translates  as "weapons" instead of "strongmen". In the third clause transliteration, it reports *djr* instead of *d'r* and *why,t* instead of *wh.jt*, but within the TLA database they are also transliterated in this way; in the English translation it presents "has a desire for his tribe" instead of "each man his tribe's suppresser". The lemma IDs and POS tags vary slightly, because the model recognizes  as a noun modified by an adjective (125010 and 81660) instead of reporting both as a single noun (125290), although this cannot be considered an error; furthermore, it does not include the gender of  and reports it only as "substantive".

In the full text stage, the same way as it was for the previous example, the model encounters difficulties due to the length of the input. The English translation begins to loop repeating the phrase "and every man has a strong man,"; the transliteration stops just beyond the middle of the second sentence; the lemma IDs, although they include some variations of  such as 851809 instead of 28550, also have the ID 5 which corresponds to , or "moment, instant", and furthermore the writing of the IDs stops halfway; finally, even the POS tags do not exceed half of the second sentence.

This text was rather complex and we noted the model's difficulties in recognizing some particular terms. The analysis of the separate sentences highlighted some small errors, especially in the translation and transliteration, while the result was much more precise in the lemma IDs and POS tags. Regarding the whole text, however, we demonstrated the model's limits of input length.

By inserting the transliteration *jri.n =j rnp.wt š'.wt, hrd.w =j hpr m nht.w e z nb m d'r wh.jt =f* as input (first maintaining the unit subdivisions and then inserting the entire text), we obtained the following results:

source → target	Uni No.	Model prediction – Separated units	Model prediction – Full text	Check
ea → de	1	Ich verbrachte viele Jahresfeste	Ich verbrachte viele Jahresfeste und meine	Ich verbrachte viele Jahre, Meine Kinder wurden Starke, Ein
	2	Meine Kinder sind zu	Kinder, die zu einem	

source → target	Uni No.	Model prediction – Separated units	Model prediction – Full text	Check
	3	<p>einem Starken geworden</p> <p>Jeder Mann hat einen Schwachen, er abweicht</p>	<p>Starken geworden waren jedermann in einem Zäpfchen, das er</p>	<p>jeder war ein Bezwinger seines Stammes⁵⁵</p>
ea → en	1	I spent many years in my life	I spent a period of many years my children, having	When I had spent many years, my boys grew into strongmen, each
	2	My children become strong	become a strong man, and every man has a	man his tribe's suppressor ⁵⁶
	3	Every man has a shortage and a shortage	strong man, and a strong man has a	
ea → tnt	1	28550 10030 94920 41011	28550 10030 94920 41011 124480 10030 116230 500292 87580 125290 50029	28550 10030 94920 41011 124480 10030 116230 500292 87580 125290 500292 177740 48730 10050
	2	124480 10030 116230 500292 87580		
	3	125010 81660 500292 177740 48730 10050		
ea → lKey	1	verb_3-inf personal_pronoun substantive_fem adjective	verb_3-inf personal_pronoun substantive_fem adjective	verb_3-inf personal_pronoun substantive_fem adjective
	2	124480 10030 116230 500292 87580	substantive_masc personal_pronoun verb	substantive_masc personal_pronoun verb_3-lit preposition
	3	125010 81660 500292 177740 48730 10050		substantive_masc substantive_masc preposition verb_3-lit substantive_fem personal_pronoun
ea → wordClass	1	verb_3-inf personal_pronoun substantive_fem adjective	verb_3-inf personal_pronoun substantive_fem adjective verb_3-lit	verb_3-inf personal_pronoun substantive_fem adjective
	2	substantive_masc personal_pronoun verb_3-lit preposition substantive	preposition substantive_masc	substantive_masc personal_pronoun verb_3-lit preposition substantive_masc substantive_masc
	3	substantive_masc adjective preposition verb_3-lit verb_3-lit personal_pronoun		preposition verb_3-lit substantive_fem personal_pronoun

Tabella 31: Fase 2, esempio 8. Predizione dalla traslitterazione e confronto.

With Gardiner code as input (Table 31), you can see several differences between the two stages and difficulty of the text.

In the separated units stage, we note improvements in each field. In the English translation the model encountered problems in the third clause, when it reached *m d'r wh.it =f*, which translated

55 Grapow, *op. cit.*, p. 44.

56 Allen, *Middle Egyptian Literature* cit., p. 93.

“has a shortage and a shortage”; in the POS tags he recognized *nḥt.w* as "substantive" but was unable to identify the masculine gender, and finally he was wrong to consider *wḥjt* as "verb_3-lit" instead of "substantive_fem". The rest of what the model reported is very accurate, but we want to underline the slight variation in the lemma IDs also encountered when Gardiner code was submitted, which divides *z nb* (should have unique ID 125290) in two different IDs, the first 125010 for *z*, and the second 81660 for *nb*.

In the full text stage, we observe the usual inconvenience of providing long input. The English translation stops at *m ḍr wḥjt =f* of the third clause, reporting “every man has a strong man, and a strong man has a”; in the lemma IDs the model interrupts without providing the last three, i.e. 177740 48730 10050; while, in the POS tags, it stops at the term *ḥpr*, writing it only as "verb" and leaving out everything else.

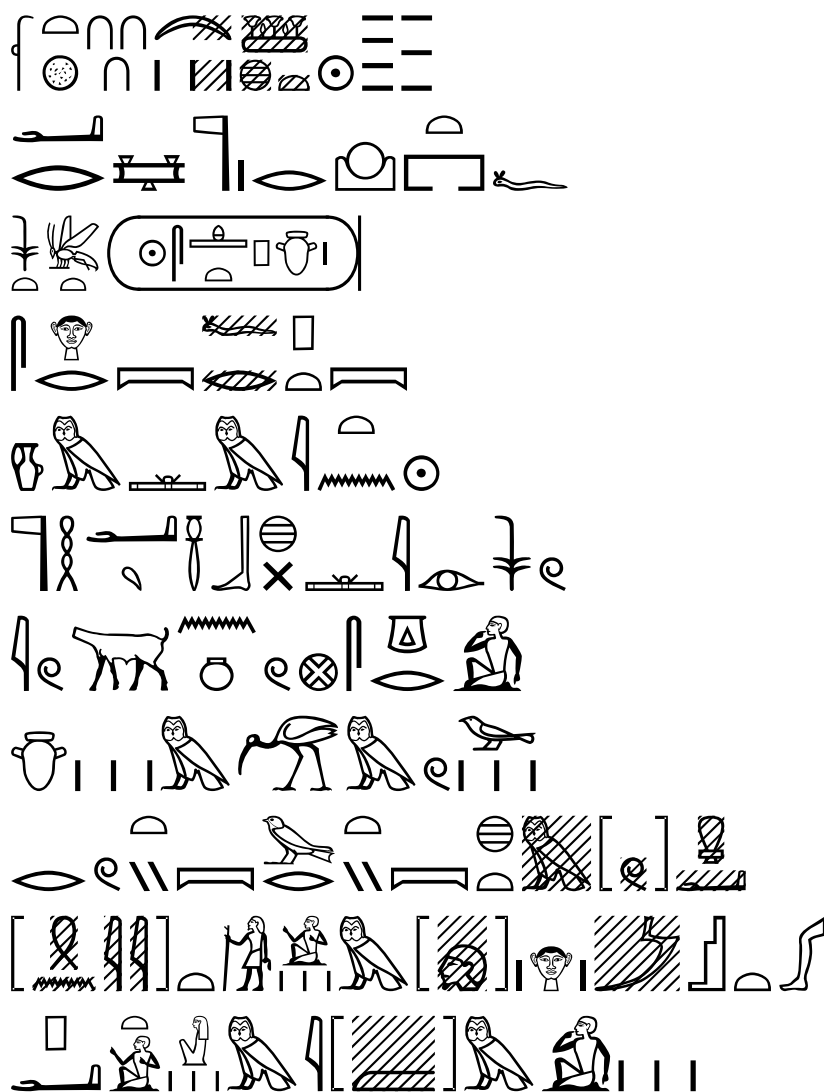
If regarding the full text stage, we can see an improvement in accuracy, in the full text stage we can say that the model is in difficulty. English translation and lemma IDs are accurate until the model gets far enough into the text that it causes a series of errors and stops writing. We can deduce that submitting a text longer than three sentences to the model is counterproductive for the correctness of the results.

3 – Third phase: Stress Test

In the last testing phase, we subjected to HIERO-TRANSFORMER a single passage taken from the Story of Sinuhe⁵⁷. This was chosen, albeit arbitrarily, for its length, in order to test the model's ability to translate a greater quantity of sentences.

The transliteration, in addition to using the same convention of the TLA, also takes into consideration the use of the comma instead of the dot to indicate the feminine ending. We chose to do it because in the previous phases we observed that the model understands the input better.

The passage we analyzed was the following:



rnpt-zp 30 jbd 3 ʰt sw 7

ʿr ntr r ʰt, t =f

nswt bjty shtp jb rʿ

shrj =f r p, t

57 In particular: pRamesseum A (Berlin 10499, verso), 5-11. Blackman, *Middle-Egyptian Stories*, Bruxelles, 1972, p. 3-4.

hnm m jtn

h' ntr 'bh m jr sw

jw hnw m sgr

jb.w m gm.w

rw,ty wr,ty htm.w

šnw,t m tp hr m's,t

p',t m jm.w

Regnal year 30, month 3 (of the) Inundation, day 7.

The god's ascent to his Akhet,

King of the Upper and Lower Egypt Sehetepibre.

He had ascended to the sky,

and had united with the sun-disk,

the god's body had mingled with the one of his creator.

The Residence was silent,

the hearts were sad,

the great double gate was sealed,

the court had its head on its knees,

and the people were mourning.

We divided this passage into smaller units for ease of reference. The passage begins with the indication of the dating (*rnpt-zp 30 jbd 3 'ht sw 7*); follows a verbal clause with the narrative infinitive (*r ntr r 'h,t =f*), which describes the death of the sovereign, later identified by the title *nswt bjty* and by his throne name *štp jb r'*. Continuing, we find a verbal clause in *sdm =f* form with causative stem (*šhrj =f r p,t*), followed by two clauses with verbs in stative form (*hnm m jtn & h' ntr 'bh m jr sw*), all which narrate the king's ascent towards the sky and its union with the sun-disk. Further, except for *rw,ty wr,ty htm.w* which is a subject-stative construction, the text continues with non-verbal adverbial clauses in which the preposition *m* introduces the predicate. These clauses describes the consequent sadness for the death of the sovereign.

The sequence of hieroglyphics in Gardiner code is as follows:

M4 X1 O50 V20 V20 V20 N11 Z2 M8 Aa1 X1 N5 Z4A Z4A Z2

D36 D21 N31 R8 Z1 D21 N27 X1 O1 I9

M23 X1 L2 X1 N5 S29 R4 X1 Q3 F34 Z1

S29 D2 D21 N1 I9 D21 Q3 X1 N1

W9 G17 Y1 G17 M17 X1 N35 N5

R8 V28 D36 F51 U23 D58 Aa1 Z9 Y1 M17 D4 M23 Z7
M17 Z7 F26 N35 W24 Z7 O49 S29 W11 D21 A2
F34 Z2 G17 G28 G17 Z7 G37 Z2
D21 Z7 X1 Z4 N1 G36 D21 X1 Z4 N1 Aa1 X1 G17 Z7 S20 D36
V7 N35 M17 M17 X1 A21 A1 Z2 G17 D1 Z1 D2 Z1 U1 Q1 X1 D56
Q3 D36 X1 A1 B1 Z2 G17 M17 Aa15 G17 A2 Z2

We gave this passage to the model in two different waves. In the first one, we inserted the entire text, and in the second, we split the units.

By entering this input, we manually compared the prediction provided by the model and the correct response, as you find in the following table:

source → target	Model prediction – Full text	Check ⁵⁸
ea → de	Regierungsjahr 30, Monatsletzter-Jahres-Jahres-Jahres-Jahres-	Im Jahre 30 im dritten [Monat] der Überschwemmungszeit, Tag 7
ea → en	Government year 30, month 2, day 2, day 2, day 2, day 2, day 2, day 2	Regnal year 30, 3 Inundation 7
ea → tnt	rnpt,t zp 30 bd,w š,tj r ntr r,t =f ns	rnpt-zp 30 jbd 3 3ht sw 7 'r ntr r 3h,t =f
ea → lKey	850581 400085 93 91901 90260 91901 227 10050 88060 141120 23290 10050	850581 850814 93 850814 216 129460 850814 21770 90260 91901 227 10050
ea → wordClass	substantive numeral substantive_masc nisbe_adjective_preposition substantive_masc preposition substanti	substantive cardinal substantive_masc cardinal substantive_fem substantive_masc cardinal verb_3-lit substantive_masc preposition substantive_fem personal_pronoun

Table 32: Phase 3. Prediction from hieroglyphics in Gardiner code of the full text and comparison.

With Gardiner code as input (Table 32), you can see the model finds great difficulties both in understanding the passage and in getting to the second line of text. This example serves to demonstrate that HIERO-TRANSFORMER cannot work on long inputs. Furthermore, if the input increases, the model commits more errors in the initial sentences.

However, we can observe that it manages to predict a few terms among the various fields correctly, but they are too few to make it worth analyzing the result.

The units were then subjected to the model individually. We manually compared the prediction provided by the model and the correct response, as you find in the following table:

source → target	Unit no.	Model prediction – Separated units	Check ⁵⁹
ea → de	1	Regierungsjahr 30, Monat 2 der	Im Jahre 30 im dritten [Monat] der

⁵⁸ For German translation: Grapow, *Der stilistische Bau der Geschichte des Sinuhe*, Berlin, 1952, p. 11. For English translation: Allen, *Middle Egyptian Literature*, Cambridge, 2015, p. 59.

source → target	Unit no.	Model prediction – Separated units	Check
		Überschwemmungszeit, Tag 4	Überschwemmungszeit, Tag 7
	2	Gott war in seinen Horizont eingegangen	stieg der Gott empor zu seinem Horizont
	3	König von Ober- und Unterägypten Sehetepib	Der König von Ober- und Unterägypten Sehetep- ib-re
	4	Er fährt zum Himmel	er wurde entfernt zum Himmel
	5	Werde nachts dem Tau ausgesetzt	indem er vereint wurde mit der Sonne
	6	Der Gottesleib ist verstopft, wenn er gemacht wird	indem der Gottesleib sich verband mit dem der ihn gemacht hatte
	7	Da schwieg die Residenz	die Residenz (war) in Schweigen
	8	Die Herzen waren in Gram	die Herzen (waren) in Trauer
	9	Das große Doppeltor war verschlossen	die beiden großen Tore waren geschlossen
	10	Der Hofstaat war in Trauerhaltung	der Hofstaat (war) im (Zustand des) Kopf auf Knie
	11	Die Pfeile waren in Verzögerung	das Volk (war) in Klage
ea → en	1	Government year 30, month 2 of the Schemu year period	Regnal year 30, 3 Inundation 7
	2	God had reached his horizon	The god's ascent to his Akhet
	3	The King of the Upper and Sub-Egypt Sehetepib	Dual King Sehetepibre
	4	He ascends to heaven	going off to the sky
	5	It is in the sun	united with the sun-disk
	6	The neck of God is broken	the god's body mingled with the one who made it
	7	The residence was silent	The residence was in stillness
	8	The hearts were in Gram	minds in grief
	9	The double door was closed	the great double gate shut
	10	The court state was in sadness	the circle with head on lap
	11	The pilgrims were in trouble	the elite in mourning
ea → tnt	1	rnpt,t zp 30 bd 2 š,tj	rnpt-zp 30 jbd 3 ʾḥt sw 7
	2	r ntr r,t=f	ʾr ntr r ʾḥ,t=f
	3	nswt bj,tj Šḥtp jb rw	nswt bjty šḥtp jb rʾ
	4	shrj =f r p,t	shrj =f r p,t
	5	ḥnm m jtn	ḥnm m jtn
	6	ḥ,w ntr b jrj sw	ḥʾ ntr ʾbḥ m jr sw
	7	jw ḥnw =s gr	jw ḥnw m sgr
	8	jb m gm,w	jb.w m gm.w
	9	rw,tj wr,tj tm.w	rw,ty wr,ty ḥtm.w
	10	šny,t m tp ḥr ms,t	šnw,t m tp ḥr mʾs,t
	11	p,t m jm	pʾ,t m jm.w
ea → IKey	1	850581 850814 93 850814 152880	850581 850814 93 850814 216 129460 850814
	2	21770 90260 91901 227 10050	21770 90260 91901 227 10050
	3	88060 400449	88060 400449

The third unit (*nswt bjty štp jb rʿ*) is translated into English as “The King of the Upper and Sub-Egypt” instead of “Dual King,” but is still correct. The model transliterates *nswt bj.tj Štp jb rw*, which can also be considered correct given that Ra is transliterated *Rʿw* in the TLA database. Lemma IDs and POS tags are completely correct.

From here on, the model translates the sentences both in present and past tenses because, with each new insertion of input, it loses track of what it had previously done. So, if a passage were to be translated in the past tense, because the beginning of the text makes us understand that it is a narration, the model would not know it since it does not keep track of this information. For this reason, subsequent translations into English can be considered accurate if reported in present or past tenses.

The fourth unit (*šhrj =fr p,t*) is translated into English with a slight variation (“ascends” instead of “going off”). Transliteration, lemma IDs and POS tags are correct, except for a small inaccuracy where the model reports “verb_3-lit” instead of “verb_3-inf”.

The fifth unit (*hnm m jtn*) is not well understood by the model, at least for the initial verb, and in fact, it is translated into English as “It is in the sun” without considering *šꜥꜥ hnm* with the meaning of joining. Transliteration, lemma IDs and POS tags are completely correct.

The sixth unit (*hʿ ntr ʿbh m jr sw*) is translated into English incorrectly (“The neck of God is broken”), and although the model understands that there is something that is part of the god, does not comprehend the rest. The transliteration *h,w ntr b jri sw* presents some inaccuracies, and we can note the absences of the radicals ʿ, ʿ and *h* similarly to what the model had reported in the second unit. The first two lemma IDs, i.e. 90260 (*ꜥ ntr*) and 81660 (*ꜥ nb*) are not exact, especially because there should be three of them, i.e. 854816 89 64366. The POS tags present the same error as the lemma IDs: the first two tags are wrong (“substantive_masc adjective”) while three tags should be there (gods_name verb_3-lit preposition).

The seventh unit (*jw hnw m sgr*) is translated into English with a slight variation (“silent” instead of “in stillness”). The model transliterates *jw hnw =s gr* omitting the preposition *m* and making a mistake because it treats the last noun as a suffixed pronoun and another term. Lemma IDs partly follow the transliteration error: they are correct except for the missing 64362. POS tags also follow the transliteration error reporting “personal_pronoun verb” instead of “preposition substantive_masc”.

The eighth unit (*jb.w m gm.w*), apart from “Gram”, a term that we also find in the German translation, is translated into English with a slight variation (“hearts” instead of “minds”). The transliteration is almost correct (*jb m gm,w*), and in fact, the model only leaves out the ending of the plural of *jb.w*. Lemma IDs and the POS tags are correct, except for a small inaccuracy in the last POS tag, in which the model does not include the masculine gender.

The ninth unit (*rw,ty wr,ty htm.w*) is translated into English in a variant (“door was closed” instead of “gate shut”), and “great” is omitted. The model transliterates with a small inaccuracy: it omits the *h* in *htm.w*. The lemma IDs and POS tags are correct, except for an inaccuracy in the first term of the POS tag where the model reports “substantive_masc” instead of “substantive_fem”.

The tenth unit (*šnw,t m tp hr m's,t*) is translated into English quite correctly, in fact, having the head on the knees or on the lap, is a position that in Middle Egyptian indicates a sad and lamenting person⁶⁰; while we can consider “court” a variant of “circle”. The model transliterates with a few inaccuracies: we find *šny,t* instead of *šnw,t* and *ms,t* in which it omits the *ʕ*. Lemma IDs and POS tags are completely correct.

The eleventh unit (*p',t m jm.w*) is not understood since it is translated into English “The pilgrims were in trouble” instead of “the elite in mourning”. In the transliteration (*p,t m jm*) we can observe the usual lack of *'* and the absence of the plural ending of *jm.w*. The last lemma ID is 24660 (𓂏𓂛𓂏𓂛 *jm*) instead of 25980. In the POS tags there are some inaccuracies about nouns' genders: the model reports “substantive_masc preposition substantive” instead of “substantive_fem preposition substantive_masc”.

As we noted, the inaccuracies and errors are various, but examining them we can say that the model is quite accurate. Some differences can be considered translation variations, others are misunderstandings regarding the gender of the nouns, while only a few are errors. Furthermore, we want to underline that when HIERO-TRANSFORMER generates a single wrong word that gives meaning to a clause (for example, a verb), the clause undergoes variations in meaning to align with the translation that the model meant. Finally, we noted that the model struggles to transliterate some consonants, such as *'*, *ʕ* and *h*. The reason for their absence is not understood and *ad hoc* experiments are necessary to verify the causes.

For the same sentences, we also submitted to the model the transliteration of the entire passage: *rnpt-zp 30 jbd 3 ʕht sw 7 'r ntr r ʕh,t =f nswt bjty šhpt jb r' shrj =f r p,t hnm m jtn h' ntr ʕbh m jr sw jw hnw m sgr jb.w m gm.w rw,ty wr,ty htm.w šnw,t m tp hr m's,t p',t m jmw*. We obtained the following results:

60 TLA ID 850576, <<https://thesaurus-linguae-aegyptiae.de/lemma/850576>> (consultato il 20/10/2023).

source → target	Model prediction – Full text	Check ⁶¹
ea → de	Regierungsjahr 30, Monat 3 der Opferplatte, Tag 7 durch den Gott für seine Aktion, den König von Ober- und Unter	Im Jahre 30 im dritten [Monat] der Überschwemmungszeit, Tag 7, stieg der Gott empor zu seinem Horizont. Der König von Ober- und Unterägypten Schetep-ib-re
ea → en	Year 30, month 3 of the time	Regnal year 30, 3 Inundation 7
ea → lKey	500284 850814 93 850814 121200 850814 119590 90260 91901 217 10050 8	850581 850814 93 850814 216 129460 850814 21770 90260 91901 227 10050
ea → wordClass	substantive_masc cardinal substantive_masc cardinal substantive_masc cardinal preposition substantive_	substantive cardinal substantive_masc cardinal substantive_fem substantive_masc cardinal

Table 34: Phase 3. Prediction from transliteration of the entire passage and comparison.

In Table 34, we can observe that the model has the same major difficulties previously encountered. Even for transliteration, we see that HIERO-TRANSFORMER cannot work on long inputs.

The units were then subjected to the model individually. We manually compared the prediction provided by the model and the correct response, as you find in the following table:

source → target	Unit no.	Model prediction – Separated units	Check ⁶²
ea → de	1	Regierungsjahr 30, Monat 3 der Achet-Zeit, Tag 7	Im Jahre 30 im dritten [Monat] der Überschwemmungszeit, Tag 7
	2	Der Gott war in seinen Horizont eingetreten	stieg der Gott empor zu seinem Horizont
	3	Der König von Ober- und Unterägypten Schetepibre	Der König von Ober- und Unterägypten Schetep-ib-re
	4	Er wird zum Himmel aufsteigen	er wurde entfernt zum Himmel
	5	Umfange mit der Sonnenscheibe!	indem er vereint wurde mit der Sonne
	6	Die Gottesglieder waren verschmolzen mit dem, der ihn geschaffen hatte	indem der Gottesleib sich verband mit dem der ihn gemacht hatte
	7	Die Residenz war in Schweigen	die Residenz (war) in Schweigen
	8	Die Herzen waren in Gram	die Herzen (waren) in Trauer
	9	Das große Doppeltor war verschlossen	die beiden großen Tore waren geschlossen
	10	Die Scheune war in Trauerhaltung	der Hofstaat (war) im (Zustand des) Kopf auf Knie
	11	Die p.t-Leute waren in Ziegenbrei	das Volk (war) in Klage
ea → en	1	Year 30, 3rd month of the flood time, day 7	Regnal year 30, 3 Inundation 7
	2	God had fallen to his horizon	The god's ascent to his Akhet

61 Per la traduzione in tedesco: Grapow, *op. cit.*, p. 11. Per la traduzione in inglese: Allen, *Middle Egyptian Literature* cit., p. 59.

62 Per la traduzione in tedesco: Grapow, *op. cit.*, p. 11-12. Per la traduzione in inglese: Allen, *Middle Egyptian Literature* cit., p. 59.

source → target	Unit no.	Model prediction – Separated units	Check
	3	The King of the Upper and Sub-Egyptian Sehetepibre	Dual King Sehetepibre
	4	He will rise to heaven	going off to the sky
	5	Close to the sun	united with the sun-disk
	6	The members of God were mixed with the one who created him	the god's body mingled with the one who made it
	7	The residence was in silence	The residence was in stillness
	8	The hearts were in Gram	minds in grief
	9	The double gate was closed	the great double gate shut
	10	A storage with the head and the knee	the circle with head on lap
	11	The nobles were in	the elite in mourning
ea → lKey	1	500284 850814 93 850814 216 129460 850814	850581 850814 93 850814 216 129460 850814
	2	119610 90260 91901 227 10050	21770 90260 91901 227 10050
	3	88060 400449	88060 400449
	4	140980 10050 91901 58710	140980 10050 91901 58710
	5	123420 64360 33080	123420 64360 33080
	6	101950 90260 89 400082 400112 129490	854816 89 64366 28550 129490
	7	21881 123280 64362 147270	851515 123280 64362 147270
	8	23290 64362 167180	23290 64362 167180
	9	93510 121710	93510 121710
	10	155970 64362 850576	155980 64362 850576
	11	58820 500292 25130	59610 64362 25980
ea → wordClass	1	substantive cardinal substantive_masc cardinal substantive_fem substantive_masc cardinal	substantive cardinal substantive_masc cardinal substantive_fem substantive_masc cardinal
	2	verb_3-lit substantive_masc preposition substantive_fem personal_pronoun	verb_3-lit substantive_masc preposition substantive_fem personal_pronoun
	3	substantive kings_name	substantive kings_name
	4	verb personal_pronoun preposition substantive_fem	verb personal_pronoun preposition substantive_fem
	5	verb_3-lit preposition substantive_masc	verb_3-lit preposition substantive_masc
	6	substantive_masc verb_3-lit preposition verb_3-inf personal_pronoun	gods_name verb_3-lit preposition verb_3-inf personal_pronoun
	7	particle substantive_masc preposition substantive_masc	particle substantive_masc preposition substantive_masc
	8	substantive_masc preposition verb_3-inf	substantive_masc preposition substantive_masc
	9	substantive_masc adjective verb_3-lit	substantive_fem verb_3-lit
	10	substantive_fem preposition substantive	substantive_fem preposition substantive
	11	substantive_fem preposition substantive	substantive_fem preposition substantive_masc

Table 35: Phase 3. Prediction from transliteration of the separate units and comparison.

By receiving the transliteration in single sentences (Table 35), one can notice a few inaccuracies and errors, but we can also observe the great accuracy of the model.

The first unit (*rnpt-zp 30 jbd 3 3ht sw 7*) is translated into English with a corrected variant. The first lemma ID is the only different one, i.e. 500284 (𓂏𓂏𓂏𓂏 *rnpt.t*), which translates as “ahropfer; Jahresfest (allg.); yearly sustenance; festival of the year”⁶³. The POS tags are completely correct.

The second unit (*r ntr r 3h,t =f*) is translated into English partially incorrect, as we find “had fallen” instead of “ascent”. The lemma IDs agree with the translation error and the model reports 119610 (𓂏𓂏𓂏 *hr*) instead of 21770. The POS tags are completely correct.

The third unit (*nswt bjty shtp jb r'*) is translated by the model in an extended variant, but still correct, while everything else is reported exactly.

The fourth unit (*shrj =f r p,t*) is translated into English with a couple of inaccuracies: the model translates in the future tense, and uses the verb “rise” instead of “going off”. The lemma IDs and POS tags are completely correct.

The fifth unit (*hnm m jtn*) is not understood by the model, and is translated into English as “Close to the sun” instead of “united with the sun-disk”. However, lemma IDs and POS tags are completely correct.

The sixth unit (*h' ntr 3bh m jr sw*) is translated into English with a difference that can be considered a variant. The lemma IDs present some differences, which can still be regarded as exact: we find 101950 (𓂏𓂏𓂏 *h'.w*) and 90260 (𓂏𓂏 *ntr*) instead of 854816 (*h'.w-ntr*), then 400082 (𓂏𓂏 *m*) instead of 64366, and finally 400112 (𓂏𓂏 *jrj*), or “zeugen; gebären; to engender; to bear”⁶⁴, instead of 28550 (𓂏𓂏 *jrj*), or “machen; tun; fertigen; fungieren als; handeln; to do; to make; to create; to act as”⁶⁵. In the par-of speech, the only difference is the first tag, in which the model reports “substantive_masc” instead of “gods_name”.

The seventh unit (*jw hnw m sgr*) is translated into English correctly. In the lemma IDs, there is a slight difference in the first, i.e. 21881, which is a correct variation of 851515. The POS tags are correct.

The eighth unit (*jb.w m gm.w*) has the same English translation as that of the Gardiner code, and once again we find “Gram” borrowed from the German translation. Lemma IDs are correct, while the last POS tag is mistaken: the model reports “verb_3-inf” instead of “substantive_masc”.

The ninth unit (*rw,ty wr,ty htm.w*) is translated into English with the same variation previously found, although this time it correctly reports “gate”. Lemma IDs are correct. In the POS tags,

63 TLA ID 500284, <<https://thesaurus-linguae-aegyptiae.de/lemma/500284>> (consultato il 20/10/2023). WB II, 435.1.

64 TLA ID 400112, <<https://thesaurus-linguae-aegyptiae.de/lemma/400112>> (consultato il 20/10/2023). WB I, 111.1-5.

65 TLA ID 28550, <<https://thesaurus-linguae-aegyptiae.de/lemma/28550>> (consultato il 20/10/2023). WB I, 108.5-111.22.

however, the model divides *rw,ty* and *wr,ty* into “substantive_masc” and “adjective”, but this could still be considered an inaccuracy rather than an error.

The tenth unit (*šnw,t m tp hr m's,t*) is only understood in its final part: the model translates the initial part into English as “A storage” instead of “the circle”, and reports “and the knee” instead of “on lap”. The lemma IDs agree with the translation, reporting 155970 (𓏏𓏲𓏲𓏲 *šnw,t*), or “Scheune; Kornspeicher; granary; storeroom; grenier”⁶⁶ instead of 155980 (𓏏𓏲𓏲𓏲𓏲𓏲𓏲𓏲 *šnw,t*), or “Umgebung; Hofstaat; Umkreis; court; entourage (of the king, of a god)”⁶⁷. The POS tags are completely correct. We want to underline that we did subsequent tests on this unit, inserting *šn,wt* instead of *šnw,t*. The comma displacement led the model to understand it as the court. the granary/storage translation is technically correct since it is translated in the same way. We could find the difference while observing the meaning of the clause or while comparing it with the hieroglyphics. In any case, we can explain this error by the fact that with the same transliteration, the model trained more with the meaning of granary/storage. However, the cause of the correction when the comma is displaced is not understood.

The eleventh unit (*p',t m jm.w*) is understood only in its initial part, as we find “The nobles were in” instead of “the elite in mourning”. Among the lemma IDs, there is 58820 (𓏏𓏲𓏲𓏲 *p'w,t*) instead of 59610, then 500292 instead of 64362 (in this case they are the same term *m*) and finally 25130 (𓏏𓏲𓏲𓏲𓏲 *jm,j*) instead of 25980. The POS tags are almost correct, but the model reports the last tag as “substantive” without specifying the masculine gender.

Once again, although there are several imperfections and errors, it can be said that the model is accurate. In a few units transliterations, the errors the model makes are in line with what it has learned, and therefore if it finds a term that could be translated in two different ways, for example, *šnw,t* (granary/storage and court), he will choose the translation that he thinks is most likely. In this case, we can reasonably assume that the increasing probability for the model is reflected by the number of terms it received during training, i.e. greater for granary/storage and smaller for circle (court).

66 TLA ID 155970, <<https://thesaurus-linguae-aegyptiae.de/lemma/155970>> (consultato il 20/10/2023). WB IV, 510.1-16.

67 TLA ID 155980, <<https://thesaurus-linguae-aegyptiae.de/lemma/155980>> (consultato il 20/10/2023). WB IV, 511.1-512.7.

References

- GARDINER Alan Henderson, *Egyptian Grammar, Being an Introduction to the Study of Hieroglyphs*, Third edition, Griffith Institute, Oxford, 1957.
- ALLEN James Peter, *Middle Egyptian Literature, Eight Literary Works of the Middle Kingdom*, Cambridge University Press, Cambridge, 2015.
- GARDINER Alan Henderson, VOGELSANG Friedrich, *Die Klagen des Bauern*, in: *Literarische Texte des Mittleren Reiches*, edited by ERMAN Adolf, vol. 1, J.C. Hinrichs'sche Buchhandlung, Leipzig, 1908.
- BRESCIANI Edda, *Letteratura e poesia dell'Antico Egitto*, preface by DONADONI Sergio, Giulio Einaudi Editore, Turin, 1969.
- VOGELSANG Friedrich, *Kommentar zu den Klagen des Bauern*, in: *Untersuchungen zur Geschichte und Altertumskunde Ägyptens*, edited by SETHE Kurt, vol. 6, Lipsia, 1913, <<https://archive.org/details/UGAAe6/page/n47/mode/2up>> (accessed: 10 october 2023).
- BLACKMAN Aylward Manley, *Middle-Egyptian Stories*, in: *Bibliotheca Aegyptiaca II*, Second edition, Fondation Égyptologique Reine Élisabeth, Bruxelles, 1972.
- GRAPOW Hermann, *Der stilistische Bau der Geschichte des Sinuhe*, in: *Untersuchungen zur ägyptischen Stilistik*, vol. 1, in coll.: Deutsche Akademie der Wissenschaften zu Berlin, Institut für Orientforschung; 10, Akademie#Verlag, Berlino, 1952, <<https://archive.org/details/Grapow1952/mode/2up>> (accessed: 10 october 2023).
- GOLENISCHIEFF Wladimir Semenovich, *Les papyrus hiératiques No. 1115, 1116 A et 1116 B de l'Ermitage Impérial à St-Petersbourg*, Bibliotheca Aegyptiaca 2, Manufacture des papiers de l'Etat, San Pietroburgo, 1913, <<https://meretsegerbooks.cld.bz/golenischeffpapyrusermitage/2/>> (accessed: 10 October 2023).
- GARDINER Alan Henderson, *The Admonitions of an Egyptian Sage from a Hieratic Papyrus in Leiden (pap. Leiden 344 recto)*, Georg Olms Verlag, Hildesheim, 1969.
- Thesaurus Linguae Aegyptiae* <<https://thesaurus-linguae-aegyptiae.de>>, Corpus issue 18, Web app version 2.1.5, 7/26/2023, ed. by Tonio Sebastian Richter & Daniel A. Werning on behalf of the Berlin-Brandenburgische Akademie der Wissenschaften and Hans-Werner Fischer-Elfert & Peter Dils on behalf of the Sächsische Akademie der Wissenschaften zu Leipzig (accessed: 2 August 2024).