

# Evaluating the usefulness of neural machine translation for the Polish translators in the European Commission

**Karolina Stefaniak**

European Commission, Directorate General for Translation

1-7 Boulevard Pierre Frieden, L-1543 Luxembourg

Office: T2 00/B076

[karolina.stefaniak@ec.europa.eu](mailto:karolina.stefaniak@ec.europa.eu)

## Abstract

The mission of the Directorate General for Translation (DGT) is to provide high-quality translation to help the European Commission communicate with EU citizens. To this end DGT employs almost 2000 translators from all EU official languages. But while the demand for translation has been continuously growing, following a global trend, the number of translators has decreased. To cope with the demand, DGT extensively uses a CAT environment encompassing translation memories, terminology databases and recently also machine translation. This paper examines the benefits and risks of using neural machine translation to augment the productivity of in-house DGT translators for the English–Polish language pair. Based on the analysis of a sample of NMT-translated texts and on the observations of the working practices of Polish translators it is concluded that the possible productivity gain is still modest, while the risks to quality are quite substantial.

## 1 Introduction

Machine translation arrived at the European Commission in 1976 with the purchase of Systran, a rule-based technology, for the English–French language pair. This initial version was then developed and adapted as EC Systran to respond to the Commission’s needs; specialized terminology and dictionaries were imported, and, over time, other language pairs were added. By

2010, some 2 million pages were translated per year with Systran, used both by EU and Member States’ officials to provide quick drafts of texts in unfamiliar languages, and by EU translators. However, while the quality was fine for getting the gist of short, repetitive texts with standardized structure and terminology, the system was rather unsuitable for translating legislation, and so its use among translators never really caught up (Petritis, 2001; Eisele, 2017a).

EC Systran was discontinued in December 2010 and since then the Commission has been working on its own machine translation system, developed by the Directorate-General for Translation. MT@EC, based on Moses, an open-source statistical machine translation (SMT) toolkit, and improved by rule-based pre- and post-processing, went operational in June 2013. By 2017 it offered 78 direct language pairs, covering all EU official languages (Eisele 2017b). It proved to be quite helpful for certain language pairs (English–French or English–Portuguese, for example) and quite unusable for other (like English–Hungarian or English–Finnish). Polish, with its free word order, rich inflectional morphology and complex orthography, was also quite challenging for the system, which produced very mixed results, from acceptable translations to unintelligible nonsense. As a result, few Polish translators tried to use it as a resource in the translation process. An internal evaluation performed in the Polish Language Department of DGT in 2017 revealed that SMT was useful only for certain text types and that only half of the SMT output was suitable for post-editing and thus likely to bring some productivity gains. Moreover, the post-editing speed depended strongly on the typing speed of the translator in question: translators with good

typing skills benefited much less from using SMT than those with poorer skills<sup>1</sup>.

In November 2017 the Directorate-General for Translation launched eTranslation, a neural machine translation system, as part of the Connecting Europe Facility. The aim of eTranslation is not only to deliver raw machine translation to the public administration or to interested SMEs in EU Member States, but also to provide MT as a tool for translators in EU institutions, to be embedded within their CAT workflow. With the introduction of eTranslation the question arose as to whether NMT actually provides better results than SMT as far as productivity and quality of translation is concerned. Based on the literature on the subject (e.g. Bentivogli et al., 2016), one would expect for example less lexical, inflectional morphology and word order errors in the NMT output when compared to SMT, and overall less editing effort, measured by automatic metrics such as BLEU and TER. However, studies on the performance of machine translation involving the Polish language are very limited (e.g. Skadiņš et al., 2014; Wołk and Marasek, 2015), and since the specific types of errors are dependent on the particular language pair involved and are influenced by the morphosyntactic features of the target language, a simple extrapolation of the results obtained with one language pair to another language pair is not possible. Therefore in the Polish Language Department of DGT it was decided to conduct an evaluation on the benefits and risks of using NMT produced by eTranslation as a translation aid, beside translation memory, for the English–Polish language pair, concentrating in particular on the post-editing efficiency and the risks to quality.

The paper is organized as follows. Section 2 provides a description of the data and methods used in the evaluation. Section 3 reports on the results and in section 4 the outcomes are discussed and final observations are offered.

## 2 Data and methods

DGT's aim is to provide high-quality translations that are fit for publication. To this end translators have at their disposal a number of tools, most

notably translation memories (Euramis) and a terminology database (IATE), integrated in a CAT environment. Machine translation is provided during pre-processing and can be included as a resource to complement the translation memory. Hence, machine translation is presented for editing only when no TM match is found. The threshold for TM matches is set at 75%. When opting to use MT, translators can choose whether they want to use it in the Autosuggest mode only (which is a feature that can speed up typing by presenting words and phrases from the MT translation memory after a few characters have been typed in an empty segment) or whether they want to have MT suggestions inserted in the segment every time no TM match is found.

For the purpose of the present evaluation a group of 9 translators was recruited. They worked in their usual way, assisted by translation memory, but were instructed to always choose MT when downloading translation resources and to use it for all 'new' segments, i.e. segments that did not have a TM match, consistently and for all their translation assignments. They were also asked to put down any comments and opinions regarding the quality of the NMT output, including examples of mistakes, for each translation assignment. Each of them translated between 1 and 13 texts. The texts varied in length from 1 page to over 150 pages and reflected well the text types and subject domains usually translated by DGT. The text types covered both legislative texts (Commission regulation, Commission decision, proposal for a Council decision, proposal for a regulation of the European Parliament and the Council, proposal for a directive of the European Parliament and the Council, report from the Commission, communication from the Commission, impact assessment to a proposal for a regulation) and non-legislative texts (public consultation, report of an audit, notification of a concentration, list of phrases for a database, press release, description of a game for children, letter to a citizens, letter to the national authorities, text on the e-Justice portal, text on the Europa portal). The subject domains included: agriculture, climate, human health, maritime affairs, fisheries, internal market, industry, transport, competition, taxation, customs union, justice, trade, regional policy, banking, finance, external relations, internal affairs and migration. The test period lasted three months (July-September 2018).

---

<sup>1</sup> That evaluation followed a similar approach to the one presented here, however, a comparison of the performance of SMT and NMT is beyond the scope of this article.

In total, during the test period the testing group translated 57 texts. 48 texts (22 legislative and 26 non-legislative) were used for further analysis (9 translations were discarded for reasons such as very few or no MT segments, shared project and problems with TER processing). Also, for each raw MT segment (3178 MT segments in total) the TER score was calculated using the final translation as a reference.

Since it was not possible to record the post-editing time automatically during translation, small-scale productivity tests on isolated sentences were conducted. Six translators (out of the 9 participating in the evaluation) were asked to perform the test. A subsample of 12 sentence pairs was randomly selected from the texts translated during the test period with the aim to collect sentences with increasing TER scores to see whether post-editing speed depended on the quality of the NMT sentence as indicated by the TER score. The source sentences contained 29 words on average. Translators were divided into two groups and each translator was asked to translate 6 sentences from scratch and to edit 6 different NMT-produced sentences. In this way each sentence from the subsample was translated or post-edited by three translators. They worked directly in a Word document, but were asked to proceed in their usual way (consult terminology database, translation memory database etc.). The time needed for both activities was measured by the author of this evaluation with a stop watch separately for each sentence and then averaged for each sentence and for each translator.

### 3 Results

#### 3.1 Quality of the NMT output

The usefulness of machine translation can be assessed by analyzing the type of errors found in the raw MT output, as some errors have more impact on the quality of the final product than other. This data was gathered by means of feedback from translators performing the post-editing. They reported that NMT produced rather fluent sentences, with few linguistic errors, but at the same time lacking consistency, which made ensuring textual coherence more difficult. Moreover, the accuracy mistakes produced by NMT were often difficult to detect and only a careful comparison with the original could reveal the mistake. For example, in the translation below the phrase ‘must be respected’ is missing entirely

from the NMT translation, which is nevertheless still fluent and grammatically correct:

EN: The capacity ceilings set out in Annex II to Regulation (EU) No 1380/2013 on the CFP must be respected and any granting of aid for purchasing a new vessel must not lead to exceeding these capacity ceilings.

NMT: Pułapy zdolności połowowej określone w załączniku II do rozporządzenia (UE) nr 1380/2013 w sprawie WPRyb oraz przyznawanie pomocy na zakup nowego statku nie mogą prowadzić do przekroczenia tych pułapów zdolności.

[The capacity ceilings set out in Annex II to Regulation (EU) No 1380/2013 on the CFP and any granting of aid for purchasing a new vessel must not lead to exceeding these capacity ceilings.]

Errors in terminology were also reported to be common, ranging from deprecated or obsolete terms chosen in place of the preferred ones, to a wrong equivalent in the given context. For example:

EN: For the purpose of this flexibility exercise, the eligibility of requested stock transfers and the state of exploitation of these stocks have been taken into account.

NMT: Do celów tej elastyczności uwzględniono kwalifikowalność wnioskowanych transferów zapasów oraz stan eksploatacji tych stad.

[For the purpose of this flexibility exercise, the eligibility of requested *stock* transfers and the state of exploitation of these *fishstocks* have been taken into account.]

In this translation, the second occurrence of the term ‘stocks’ has been wrongly translated as ‘fishstocks’ (fisheries term) and not as ‘inventories’ (financial term). Mistakes of that kind were common especially for single-word homonyms and are also an example of errors in consistency; both types of mistakes are related as terms need to be translated consistently. Inconsistencies could occur even in the same sentence, like in the example above, but were most often found across sentences. All testers agreed that using NMT made keeping terminology coherent in the translation more difficult.

These findings are consistent with the literature on the subject (see section 1 above) and so these two types of errors could be considered typical for NMT irrespective of the language pair and the text type. The issues mentioned in the feedback that seem specific for the translation from

English into Polish were wrong word order, calqued from the original sentence, and errors in verbs forms (tense, voice, aspect or mood). Pronouns, too, were often mistranslated due to their ambiguity in the English original, which had to be resolved in the translation into Polish, like in the following example:

EN: The plastic gives the article its essential character as its [plastic's] presence is predominant in quantity and because of its determinant role in relation to the use of the article.

NMT: Tworzywo sztuczne nadaje artykułowi jego zasadniczy charakter, ponieważ występuje on w przeważającej ilości oraz z powodu jego decydującej roli w odniesieniu do użytkowania artykułu.

[The plastic gives the article its essential character as its [article's] presence is predominant in quantity and because of its determinant role in relation to the use of the article.]

Here the masculine pronoun ‘on’ is used in the NMT output, which refers to ‘the article’; to refer to the ‘the plastic’ the neuter pronoun ‘ono’ should have been selected. Thus, in spite of being grammatical, the translation is wrong.

Specific in the context of DGT were frequent mistranslations of the titles of legal acts, since they should not have been retranslated and needed to be quoted verbatim the way they had been published in the Official Journal. The same was true for quotations. Also, when confronted with abbreviations, proper names, including given names, Latin names, chemical nomenclature etc., as well as other infrequent words the NMT engine could get very creative, from misplacing letters (‘Łukasz Brasszek’ instead of ‘Łukasz Brzenczek’) to producing new ‘words’ (‘femzabójstwa’, a non-existing word as an equivalent of ‘femicides’), to creating unintended comical effects (‘newborns’ translated as ‘nowe borówki’, literally ‘new berries’). This may be explained by the fact that eTranslation had been trained on corpus of predominantly legal texts, as they constitute the majority of texts translated by DGT.

### 3.2 Post-editing effort

The quality of the MT output is reflected, too, in the technical effort of post-editing, i.e. in the number of insertions, deletions and word shifts that the translator has to perform to produce a translation of the required quality. The technical effort can be measured by means of automatic evaluation metrics, such as TER (translation edit

rate) (Snover et al., 2006). TER scores range from 0 (best) and 1 (worst). The score can be greater than 1, if the number of edits between the MT and reference segment is greater than the number of words in the reference segment.

TER scores were obtained for all sentences in the sample and then averaged for individual texts and for the sample as a whole. The average TER score for individual texts in the sample varied from 0.14 to 1.1; the average TER score for the whole sample was 0.42. The median TER score varied from 0.10 to 0.67. The median for the whole sample was 0.33. The share of MT segments with TER=0 (i.e. segments that did not require any editing) varied from 0% to 40.9% (12.7% on average). The first quartile was at the level of 0.31 and the 3rd quartile at 0.49.

Significant differences between legislative and non-legislative texts were observed. The average TER score for individual legislative texts varied from 0.14 to 0.61 (average: 0.34, median 0.32). The average TER score for non-legislative texts varied from 0.2 to 1.1 (average: 0.49, median: 0.42). The summary of the results is presented in Table 1.

	All	Legislative	Non-Legislative
Average TER	0.42	0.34	0.49
Median TER	0.33	0.32	0.42
1 <sup>st</sup> Quartile	0.31	0.28	0.38
3 <sup>rd</sup> Quartile	0.49	0.38	0.64
TER=0	12.7%	11.7%	13.5%

Table 1. Comparison of legislative vs. non-legislative texts

These quantitative results clearly show that for the English–Polish language pair NMT performs much better for legislative texts in comparison to non-legislative texts. This may be explained by the fact that in general MT performs better for standard, repetitive texts featuring characteristic terms and phrases (which is typical for legislative texts), while it does not give equally good results for texts containing new terminology or rare words, including idioms, metaphors or proper names (which occur more often in non-legislative texts).

When interpreting the results, one has to remember that metrics like TER largely ignore notions of semantic equivalence and say nothing about the reason of the edits. Neither do TER scores fully capture the cognitive effort of post-editing, as some corrections may be more demanding than other, depending on the type and

severity of the errors (see also Koponen et al., 2012). This is particularly problematic in the evaluation of NMT, which produces fluent, grammatical sentences that may nonetheless contain serious accuracy mistakes (see section 3.1 above). Another problem with metrics relying on the post-editing distance is that even minor errors might require substantial changes to the MT output, or the other way round, minor edits may suffice to correct severe mistakes (see also Burchardt and Lommel, 2017). In other words, the technical effort may not necessarily correlate with the temporal effort, i.e. the speed at which the translator processes the MT output. This is discussed in the next section 3.3.

### 3.3 Productivity gain

Although it seems intuitive to predict that MT-produced segments with low TER scores, i.e. segments that require little or no intervention, require also short editing times, this correlation is by no means straightforward. In particular, ‘establishing the exact threshold on HTER scores above which translations should be considered too bad to be post-edited is a complex problem in itself’ (Specia and Farzindar, 2010: 38). The research on this subject is inconclusive. For example, Gaspari et al. (2014) reported only a weak correlation between the evaluation metrics (BLEU, TER and METEOR) and the post-editing time. On the other hand, de Gibert Bonet (2018) found out that the higher the TER score, the longer translators needed to correct the MT-produced sentence. The TER threshold she established for productivity gain was 0.33. Also Parra Escartín and Arcedillo (2015) reported a productivity gain for segments with  $TER \leq 0.3$ .

To determine such productivity threshold for the purpose of the present evaluation, small-scale productivity tests with 6 translators were conducted. The results per sentence are shown in Table 2.

	TER	Average translation speed	Average post-editing speed
Sentence 1	0.11	0.21	0.16
Sentence 2	0.15	0.23	0.48
Sentence 3	0.21	0.32	0.53
Sentence 4	0.22	0.17	0.21
Sentence 5	0.33	0.18	0.32
Sentence 6	0.34	0.24	0.52
Sentence 7	0.41	0.21	0.25
Sentence 8	0.46	0.16	0.39

Sentence 9	0.50	0.32	0.29
Sentence 10	0.61	0.17	0.20
Sentence 11	0.63	0.18	0.30
Sentence 12	0.71	0.20	0.25

Table 2. Average translation and post-editing speed (in words/second) per sentence.

The post-editing speed varied greatly between sentences with the same TER. No clear correlation between the TER score and the post-editing speed could have been established and no clear productivity threshold. Rather, based on the observations of translators during the productivity tests, the processing speed seemed to depend more on the syntactic complexity of the source sentence, its terminological density and the number of references it contained that needed to be checked. Consider the following sentence:

Article 14(1)(b) of Commission Regulation (EC) No 2535/2001 provides that licence applications lodged from 1 to 10 June may be used for imports during the period from 1 July to 31 December following.

To make sure the MT output is correct, the translator has to look up the regulation in question, find the appropriate article and compare the source text and the translation as published in the Official Journal to the text under translation and the machine translation output, respectively. Only then can they make the decision on the accuracy of the MT. This may take as much time as translating from scratch, or more, because there is no text to compare when translating from scratch. In this case, indeed, post-editing took more time than translation from scratch (on average 0.16 vs. 0.21 words/second).

These observations are consistent with the conclusions of Tatsumi (2009), who suggests that there may not be a linear relationship between the post-editing speed and the differences measured by automatic metrics, and that other variables like sentence length or error types influence the processing time.

The post-editing speed varied greatly also among the 6 translators. It could be observed, for example, that translators who felt unfamiliar with the subject domain needed more time for post-editing in comparison to their colleagues specializing in that subject. Still, on average, all translators were faster when post-editing the NMT output than when translating from scratch, even though the difference was sometimes minimal. The average translation speed was 0.22 words/second; the average post-editing speed

was 0.32 words/second. This is shown in Table 3.

	Average translation speed	Average post-editing speed
Translator A	0.23	0.40
Translator B	0.29	0.37
Translator C	0.20	0.32
Translator D	0.17	0.32
Translator E	0.26	0.27
Translator F	0.18	0.24

Table 3. Average translation and post-editing speed (in words/second) per translator.

Using the average processing speed for translating from scratch of 0.20 words/second and the post-editing speed for NMT of 0.32 words/second the potential productivity gain could be calculated. A productivity gain is the difference between the time necessary to translate a page with the help of translation memory (TM) matches only and the time needed to translating the same page using TM supplemented with MT suggestions:

$$\begin{aligned} \text{productivity gain} &= (\text{time to edit TM} \\ &\text{matches} + \text{time to translate from scratch}) \\ &- (\text{time to edit TM matches} + \text{time to post-} \\ &\text{edit MT}) \end{aligned}$$

Because translation and post-editing speed are expressed in words/second, a standard page of 350 words was assumed. In the sample, the average share of ‘empty’ segments that did not yield any TM matches and which could thus potentially benefit from using NMT was 44.9%. Also, since the processing speed of TM matches was not measured, it was assumed to equal the post-editing speed. The productivity gain thus calculated was only 4 minutes or 17 % per page on average. Similarly modest results, when post-editing speed is measured in an actual working environment and when TM matches is taken into account, are reported in the literature so far. For example, Castilho et al. (2017), who compared the translation of texts from educational domain from English into German, Greek, Portuguese and Russian, also found no clear improvement with regard to productivity, suggesting that ‘NMT for production may not as yet offer more than an incremental improvement in temporal PE effort’ (Castilho et al., 2017: 127).

## 4 Conclusions

The initial driving force behind the development of machine translation back in the 1940s was the firm belief that high-quality fully automated translation is not only possible, but is a matter of a few years. After seven decades of research one needs to face the fact that when it comes to MT there is no one-size-fits-all solution. Machine translation engines have to be customized to accommodate the desired terminology, style, domain and other requirements, including whether the MT translation is meant for publication and dissemination or rather for short-lived internal use. In other words, ‘the degree of human involvement required (...) will depend on the purpose, value and shelf-life of the content’ (Way 2013).

The requirements placed on DGT translators, especially regarding the quality of the translation of legal acts, are even higher than the usual requirements on the translation market for texts meant for publication. This is because mistakes in legal texts impact not only on DGT’s image, they also have legal consequences. Beside accuracy, consistency within the text and with any related texts is of particular importance, e.g. terminological consistency with the acts in the same domain or lexical and terminological consistency with the basic legal act. Hence the usefulness of machine translation must be evaluated in view of these particular requirements.

For the Polish language, neural machine translation usually produces rather well-formed sentences suitable for post-editing. Hence, correcting the NMT output was not perceived to be very cumbersome by translators participating in the evaluation. On the other hand, on average only less than 20 % of NMT segments did not contain any errors; and most of the mistakes in the remaining segments were mistakes in accuracy or terminology, which poses serious challenges to the quality of the final translation. Legal texts seem to benefit more from NMT than non-legal texts, probably because of their repetitive and standard character. In non-legal texts NMT suggestions often need extensive adaptation of style and register and therefore are in general perceived to be less useful.

There seems to be only a weak correlation between the TER score and the post-editing time, although a bigger sample is necessary to corroborate this finding. The calculated

productivity gain when NMT is used to complement TM matches is still modest. However, this finding needs to be confirmed with more data obtained under more controlled conditions. Also, observations of the working practices of the Polish translators at DGT point out to the possibility that there might be a stronger relationship between other variables and the post-editing speed, such as the experience of the translator in the subject domain and the number of terms and references or quotations in the sentence. This hypothesis, too, would require further testing.

## References

- Bentivogli, Luisa, Arianna Bisazza, Mauro Cettolo and Federico Marcello. 2016. Neural versus Phrase-Based Machine Translation Quality: a Case Study. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 257-267.
- Burchardt, Aljoscha and Arle Lommel. 2017. Quality Management for Translation. In Jörg Porsiel (ed.). 2017. *Machine Translation. What Language Professionals Need to Know*. Berlin: BDÜ Weiterbildungs- und Fachverlagsgesellschaft GmbH, 128-147.
- Castilho, Sheila, Federico Gaspari and Maria Gialama. 2017. A Comparative Quality Evaluation of PBSMT and NMT using Professional Translators. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP 2016)*, 257-267.
- Eisele, Andreas. 2017a. Machine Translation at the European Commission. In Jörg Porsiel (ed.). 2017. *Machine Translation. What Language Professionals Need to Know*. Berlin: BDÜ Weiterbildungs- und Fachverlagsgesellschaft GmbH, 209-220.
- Eisele, Andreas 2017b. From MT@EC to eTranslation in CEF. Overview for the DGT QT21 Workshop. <http://www.qt21.eu/wp-content/uploads/2018/02/02-20170315-DGT-overview-March-2017.pdf> (consulted on 24.3.2019).
- Gaspari, Federico, Antonio Toral, Sudip Kumar Naskar, Dede Groves and Andy Way. 2014. Perception vs Reality: Measuring Machine Translation Post-Editing Productivity. Sharon O'Brien, Michel Simard and Lucia Specia (eds). 2014. *Proceedings of the 11th conference of the Association for Machine Translation in the Americas: workshop on post-editing technology and practice (WPTP3)*, Vancouver, 60-72.
- de Gibert Bonet, Ona. 2018. *To post-edit or to translate... That is the question. A case study of a recommender system for Quality Estimation of Machine Translation based on linguistic features*. MA Thesis. University of Basque Country.
- Koponen, Maarit, Wilker Aziz, Luciana Ramos and Lucia Specia. 2012. Post-editing time as a measure of cognitive effort. *AMTA 2012 Workshop on Post-Editing Technology and Practice (WPTP 2012)*, 11-20.
- Parra Escartín, Carla and Manuel Arcedillo. 2015. A fuzzier approach to machine translation evaluation: A pilot study on post-editing productivity and automated metrics in commercial settings. *Proceedings of the Fourth Workshop on Hybrid Approaches to Translation (HyTRA@ACL 2015)*, 40-45.
- Petritis, Angeliki. 2001. *EC Systran: The Commission's Machine translation system*. European Commission Translation Service. <http://mt-archive.info/Petritis-2001.pdf> (consulted on 27.03.2019).
- Skadiņš, Raivis, Mārcis Pinnis, Andrejs Vasiljevs, Inguna Skadiņa and Tomáš Hudík. 2014. Application of Machine Translation in Localization into Low-Resourced Languages. *Proceedings of the Seventeenth Annual Conference of the European Association for Machine Translation (EAMT 2014)*, 209-216.
- Specia, Lucia and Atefeh Farzindar. 2010. Estimating Machine Translation Post-Editing Effort with HTER. *Proceedings of the AMTA-2010 Workshop Bringing MT to the User: MT Research and the Translation Industry*, 33-41.
- Snover, Matthew, Bonnie Dorr, Richard Schwartz, Linnea Micciulla and John Makhoul. 2006. A Study of Translation Edit Rate with Targeted Human Annotation. *Proceedings of Association for Machine Translation in the Americas*, 223-231.
- Tatsumi, Midori. 2009. Correlation Between Automatic Evaluation Metric Scores, Post-editing speed, and Some other Factors. *Proceedings of MT Summit XII*, 332-339.
- Way, Andy. 2013. Traditional and emerging use-cases for machine translation. *Proceedings of Translating and the Computer 35*, London.
- Wolk, Krzysztof and Krzysztof Marasek. 2015. Neural-based machine translation for medical text domain. Based on European Medicines Agency leaflet texts. *Procedia Computer Science 64*, 2-9.