HistoCrypt 2023

6th International Conference on Historical Cryptology

20-22 June 2023, Munich

Proceedings of the 6th International Conference on Historical Cryptology HistoCrypt 2023

Editors
Carola Dahlke and Matthias Göggerle

Published by:

NEALT Proceedings Series No. 50
ISSN: 1736-8197 (print), 1736-6305 (online)
ISBN 978-91-8075-134-6 (PDF)
https://doi.org/10.3384/ecp195
Linköping Electronic Conference Proceedings No. 195
ISSN 1650-3686 (print), 1650-3740 (online)
NEALT Proceedings Series No. 50
ISSN 1736-8197 (print), 1736-6305 (online)

SPONSORS

Deutsches Museum

Universität der Bundeswehr München

Freundes- und Förderkreis
Deutsches Museum e.V.

Preface

We are very pleased to introduce the proceedings of the 6th International Conference on Historical Cryptology (HISTOCRYPT 2023), held at the Deutsches Museum in Munich, Germany from June 20–22, 2023.

Traditionally, HISTOCRYPT is open to scientific contributions from historical cryptology and cryptography, as well as connected disciplines such as linguistics, image processing, machine learning, and historical research. Typical subjects include, but are not limited to, the development and application of cryptography throughout history, the analysis of historical ciphers and encryption machines with modern computerized methods, the decryption of previously unsolved historical cryptograms, and teaching and promoting cryptology in schools, universities, and the public.

The scientific program of HISTOCRYPT 2023 was curated by an international program committee, consisting of researchers in cryptology, history, intelligence and language technology. The main goal was to deliver a high quality program with a wide variety of topics. In the double-blind synchronous review process at least three reviewers evaluated each submission and gave their recommendations. Ambiguous results were thoroughly discussed among the reviewers and the senior members of the program committee, who then made the final selection.

The program committee welcomed submissions in two tracks: *regular papers* up to 10 pages (including references) on substantial, original, and unpublished research, including evaluation results, where appropriate, and *short papers* up to 4 pages (including references) on smaller, focused contributions, work in progress, negative results, surveys, tutorials, or opinion pieces.

The conference received 26 submissions from all over the world, including from Austria, Cyprus, the Czech republic, Finland, France, Germany, Hungary, Italy, Netherlands, Poland, Slovakia, Spain, Sweden, the UK, Australia, Brazil, Israel and the United States. We were able to accept 18 full papers, 2 short papers and 2 posters. All accepted papers are collected in this volume in alphabetical order of the first author's surname.

In addition, we were very fortunate to have two high profile keynote speakers: *Jörn Müller-Quade*, expert on IT security and the certification of AI-systems; and *George Lasry*, an internationally renowned expert on the computerized cryptanalysis of historical algorithms.

Organizing a conference and a peer-review process always relies on the goodwill and support of many colleagues to take their valuable time and contribute to an interesting and fruitful program. First of all, I would like to thank Carola Dahlke for the great collaboration and the great conference location.

My special thanks goes to all area chairs of the program committee, Bernhard Esslinger, Carola Dahlke, Marek Grajek, Benedek Láng, Beáta Megyesi, Klaus Schmeh and Dermot Turing for their substantial support and for many constructive online-meetings. It was a pleasure working with you.

I also want to thank the 26 members of our program committee for their time, effort, and constructive feedback during the review process. In addition, I would like to thank all the authors for once again making these proceedings interesting, diverse and impressive. Furthermore, many thanks go to Christoph Ruhl and Alexandra Suckow for managing the conference website and the registration process. Last but not least, my sincere gratitude goes to the local organization, Carola Dahlke and Mathias Göggerle from Deutsches Museum München, for carrying the burden of the local organization, and for publishing these proceedings.

I wish you all a joyful time while exploring the papers in this volume!

Arno Wacker
Program Chair of HISTOCRYPT 2023

Program Committee

- Arno Wacker (program chair), University of the Bundeswehr Munich, Germany
- Carola Dahlke (general chair) , Deutsches Museum, Germany
- Bernhard Esslinger (area chair), University of Siegen, Germany
- Benedek Láng (area chair), Budapest University of Technology and Economics, Hungary
- Marek Grajek (area chair), Poland
- Beáta Megyesi (area chair), Uppsala University, Sweden
- Nadine Akkermann, Leiden University, The Netherlands
- Eugen Antal, Institute of Computer Science and Mathematics, Slovakia
- Jörgen Dinnissen, private researcher, the Netherlands
- Ekaterina Domnina, Moscow State Lomonosov University, Russia
- John F. Dooley, Knox College, US
- Joachim von zur Gathen, University of Bonn (Emeritus), Germany
- Otokar Grosek, Slovak University of Technology, Slovakia
- Henner Heck, University of the Bundeswehr Munich, Germany
- Olga Kieselmann, University of the Bundeswehr Munich, Germany
- Jozef Kollár, Slovak University of Technology in Bratislava, Slovakia
- Nils Kopal, University of Siegen, Germany
- George Lasry, DECRYPT and CrypTool projects, Israel
- Maximilian Marketsmüller, University of the Bundeswehr Munich, Germany
- Jakub Mírka, State Regional Archives in Pilsen, Czech Republic
- Diego Navarro, University Carlos III of Madrid, Spain
- Ingo Niebel, Historian and journalist, Germany
- Stefan Porubsky, The Czech Acadamey of Sciences, Czech Republic
- Christoph Ruhl, University of the Bundeswehr Munich, Germany
- Mathias Schlolaut, University of the Bundeswehr Munich, Germany
- Klaus Schmeh, Private researcher, Germany

- Gerhard F. Straßer, Pennsylvania State University (Emeritus), Germany (US)
- Betsy Rohaly Smoot, Independent scholar, US
- Jörg Ulbert, Université Bretagne Sud, France
- Serge Vaudenay, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
- Michelle Waldispühl, Göteborgs Universitet, Sweden
- Frode Weierud, Crypto Cellar Research, Norwegian
- René Zandbergen, European Space Agency, Germany

Local Organizing Committee

- Carola Dahlke (general chair) , Deutsches Museum, Germany
- Matthias Göggerle, Deutsches Museum, Germany
- Maximilian Marketsmüller, University of the Bundeswehr Munich, Germany
- Christoph Ruhl, University of the Bundeswehr Munich, Germany
- Alexandra Suckow, University of the Bundeswehr Munich, Germany

Steering Committee

- Klaus Schmeh, Cryptovision, Germany
- Dermot Turing, Kellogg College, UK

Contents

Encrypted Documents and Cipher Keys From the 18th and 19th Century in the Archives of Aristocratic Families in Slovakia

Eugen Antal, Pavol Marák, Pavol Zajac
Slovak University of
Technology in Bratislava
Slovakia
<firstname.secondname>@stuba.sk

Tünde Lengyelová, Diana Duchoňová
Institute of History of
the Slovak Academy of Sciences
Slovakia
tunde.lengyelova@savba.sk,
histdidi@savba.sk

Abstract

In this article, we present encrypted documents and cipher keys from the 18th and 19th century, related to central-European aristocratic families Amade-Üchtritz, Esterházy, and Pálffy-Daun. In the first part of the article, we present an overview and analysis of the available documents from the archives with examples. We provide a short historical overview of the people related to the analyzed documents to provide a context for the research. In the second part of the article, we focus on the digital processing of these historical manuscripts. We developed new tools based on machine learning that can automate the transcription of encrypted parts of the documents, which contain only digits as cipher text alphabet. Our digit detection and segmentation are based on YOLOv7. YOLOv7 provided good detection precision and was able to cope with problems like noisy paper background and areas where digits collided with the text from the reverse side of the paper.

1 Introduction

Many historical encrypted documents and cipher keys have survived in various archives all over the world. Collecting materials from a particular time period or geographic location can bring insights into the cipher design from a specific time/location. A valuable collection of historical encrypted documents and cipher keys can be found in Austria (Láng, 2020), Germany (Antal and Mírka, 2022), etc. It is thus desirable to investigate and compare various collections from a wide time range and location (Megyesi et al., 2022), because it is essential for understanding how ciphering evolved in the past. The encrypted documents and keys are easy to recognize but are difficult to locate in archives. This fact was already observed by Láng (2020) and other researchers in historical cryptography.

Systematic collection, annotation, and sharing of these materials among researchers are thus very important. For this reason, we bring to the reader's attention to two interesting projects that focus on such collections. The *DECRYPT*[1] project (Megyesi et al., 2020; Megyesi et al., 2022) contains 4360 records at the time of writing this paper. Perhaps the only disadvantage of this database is that most of the documents are not publicly available[2]. The *HCPortal*[3] project (Antal and Zajac, 2020; Antal and Zajac, 2021) contains at the time of the writing of this document 763 records. Every record is freely available to everyone. One of the data collections that is planned to be added to HCPortal in the near future is the content of this article.

The aim of this paper is to provide the first detailed description of encrypted documents and cipher keys from the 18th and 19th century, related to central-European aristocratic families that lived in the territory of modern Slovakia. The investigated collection is deposited in the Slovak National Archive in Bratislava. During the initial research phase, we found documents related to historical cryptology in the fonds of three different aristocratic families:

- Amade-Üchtritz,
- Esterházy,
- Pálffy-Daun.

In these fonds, we identified a total of 13 card-

[1] https://de-crypt.org/

[2] The probable reason is the problem with granting permissions from archives.

[3] https://hcportal.eu

board boxes containing 96 encrypted documents[4] and 9 cipher keys.

The encrypted documents are from the 18th and 19th century. The earliest document is from 1711, and the latest is from 1834. The most commonly used encryption system from this time period is called *nomenclator*. Several publications have already described this cipher system in detail, including its large variety. For more information about this encryption system see (Antal and Mírka, 2022; Megyesi et al., 2022) or other publications published in the proceedings of the HistoCrypt[5] conference.

2 Analysis of Available Materials

The used symbol set (cipher text alphabet) of the investigated collection, both in encrypted documents and cipher keys consist of only numbers (e.g. on Figures 1, 3), and markups of numbers (Figure 2). It also confirms that using digits became more common/standard in the 17th/18th centuries as stated in (Megyesi et al., 2022). A nomenclator system can be used with and without separators based on its design. Separators are mostly required to clearly distinguish/split the cipher text units (Antal and Mírka, 2022). In total, 37 encrypted documents contain cipher text with a separator, and 59 are without a separator character.

Some documents consist of cipher text only, some contain the corresponding (decrypted) plain text. In 8 cases, only the cipher text is preserved. Plain text is written above/under the encrypted lines of the text in 16 documents. In 58 cases, the plain text is available in a separate document. Finally, in 14 cases we were unable to verify whether the plain text is preserved.

Encrypted documents were preserved in all three fonds, however, cipher keys were only found in two of them. An interesting fact about this collection is that the preserved encryption keys do not match[6] the preserved cipher texts. We also investigated some (possibly related) cipher keys from Austria (Haus-, Hof- und Staatsarchiv of Vienna) and from Hungary, however, these keys also did not match our encrypted documents.

A more detailed description of the available encrypted documents and cipher keys are present in the following subsections.

2.1 Encrypted Manuscripts

The **Esterházy**[7] family archive contains most of the encrypted documents (fifty-eight), spread in four boxes. In cardbox no. 631 there is (only) one document related to cryptography, a four-page-long encrypted letter from 1744 written in the German language. The most promising materials[8] are deposited in cardbox no. 634, which contains an extensive communication between Count Nicolaus Esterházy and Wenzel Anton Prince of Kaunitz-Rietberg from the years 1756 and 1757. The whole communication is in the German language. Interestingly, they used at least two different encryption keys in the communication - one where the cipher text units are separated with a separator char (Figure 1) and one without a separator (Figure 3). Moreover, the cipher texts with a separator contain special digit markups (Figure 2). There are two additional documents sent to Count Nicolaus Esterházy in this box, one sent by Graf Colloredo and one by Empress Maria Theresa. For most of the encrypted documents in this box, the plain text is also available in separate documents.

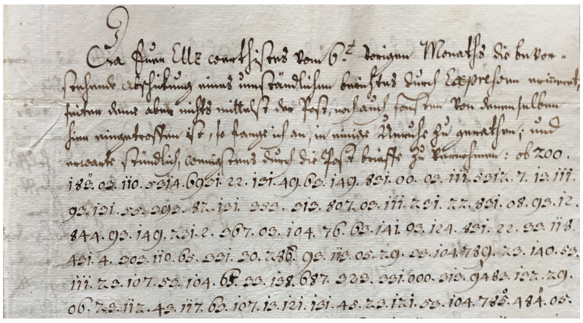


Figure 1: Cipher text example with separator chars (Slovak National Archives, fond Esterházi - čeklíska vetva, box n. 634)

Cardbox no. 635 contains letter concepts from the years 1741 and 1744 with encrypted parts in nine of them. It should be noted, however, that only small parts of these letter concepts are encrypted. These encrypted parts are mainly written below the corresponding plain text parts (Figure 4), or inserted next to the non-encrypted text parts (Figure 5). Again, the text is written in the German language.

[4] The documents consist of nearly 300 pages containing encrypted text.

[5] https://histocrypt.org

[6] Except for one key matching the encrypted diary of Emil Üchtritz, see Section 2.1.

[7] Čeklís family branch.

[8] Fourty one documents with encrypted content.

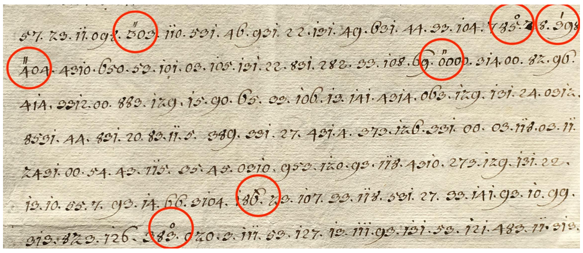


Figure 2: Special digit markups (Slovak National Archives, fond Esterházi - čeklíska vetva, box n. 634)

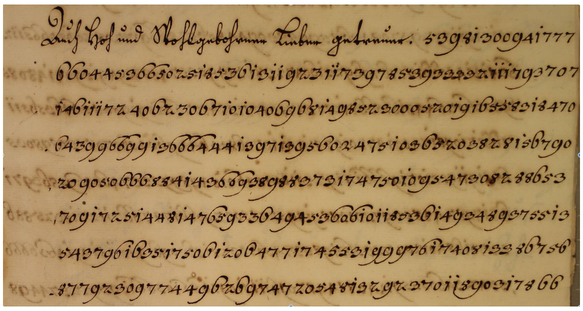


Figure 3: Cipher text example without separator chars (Slovak National Archives, fond Esterházi - čeklíska vetva, box n. 634)

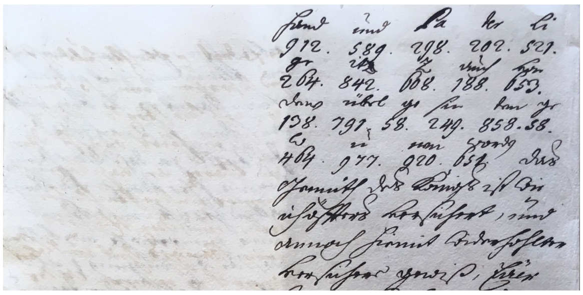


Figure 4: Message concept with encrypted part (Slovak National Archives, fond Esterházi - čeklíska vetva, box n. 635)

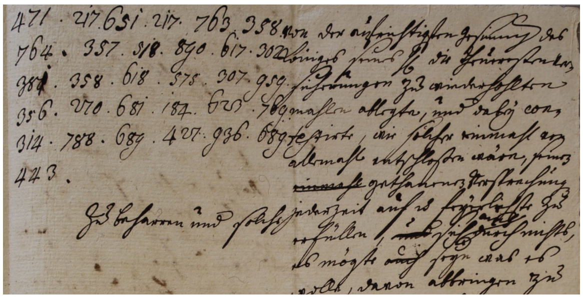


Figure 5: Message concept with encrypted part (Slovak National Archives, fond Esterházi - čeklíska vetva, box n. 635)

Cardbox no. 636 contains six encrypted documents: one fully encrypted concept and five letters, of which there are five documents written in French (Figure 6), and one in the German language. All of the documents are dated to 1744. One encrypted letter was sent by Empress Maria Theresa to Count Nicolaus Esterházy. In the remaining documents, we do not recognize the participants of the communication. The corresponding plain text parts are not available for these documents.

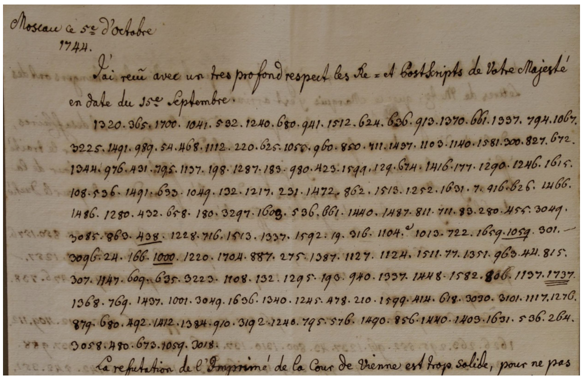


Figure 6: Cipher text example (Slovak National Archives, fond Esterházi - čeklíska vetva, box n. 636)

In the **Amade-Üchtritz** family archive, we have found thirty-seven encrypted documents. Most of the documents are from the beginning of the 19th century and are related to Emil Üchtritz in the diplomatic service of the King of Saxony and to various Saxon envoys. Cardbox n. 136 contains five encrypted messages from 1814-1834, written in German and French. Three documents were sent by Detlev Graf von Einsiedel, and two by Minckwitz, Saxon Minister for Foreign Affairs. Four documents contain the corresponding plain text parts written above/below the cipher text. In one case, only the cipher text is preserved (Figure 7).

Cardbox n. 139 contains six encrypted French messages from 1806, correspondence of Senfft von Pilsach, Saxon envoy in Paris. In all cases, the corresponding plain text is written above the cipher text (Figure 8).

Cardbox n. 140 contains twenty-one documents written in French and German languages dated between 1816-1823. These messages are communication of Detlev Graf von Einsiedel with Schulenburg, the Saxon envoy in Vienna, and to Griesinger, the Saxon Legation Counsellor in Vi-

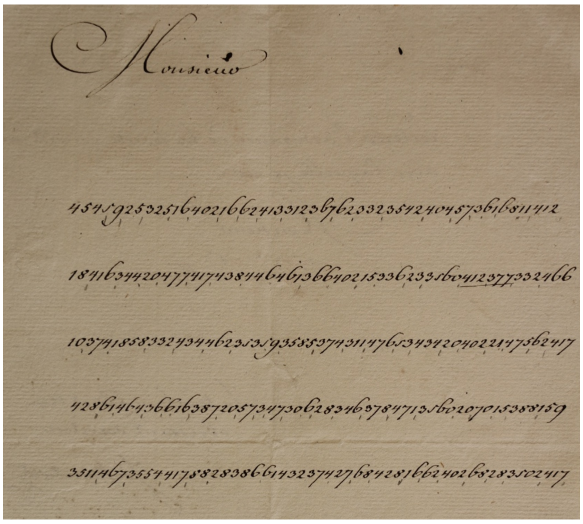


Figure 7: Cipher text example (Slovak National Archives, fond Amade-Üchtritz, box n. 136)

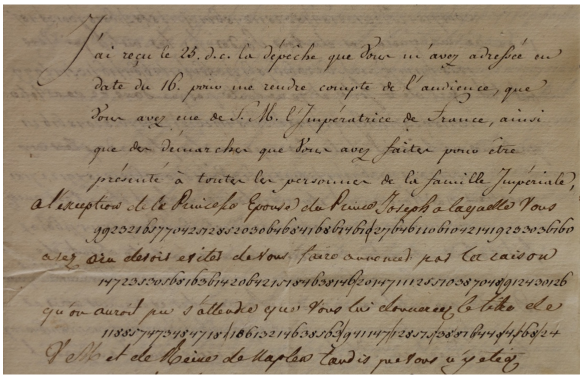


Figure 8: Cipher text example (Slovak National Archives, fond Amade-Üchtritz, box n. 139)

enna. One message was sent by Minckwitz to Schulenburg. At least two different cipher keys were used. In all cases, the corresponding plain text is available. Cardbox n. 141 contains four encrypted documents written in German and French languages sent to Fleming, the Polish-Saxon envoy in Vienna (Figure 9). These documents are from 1758, and the corresponding plain text is written above the cipher text in all cases.

Very interesting materials are located in cardbox n. 150, which contains the diary of Emil Üchtritz from October 1804 to August 1805 written mainly in the German language. The diary has 123 pages, containing encrypted entries on 27 pages (Figure 10). Analysis and detailed description of the diary will be the content of a future publication.

In the fonds of the **Pálffy-Daun** family, we found only one encrypted document written in the French language sent by Emperor Charles VI to Wirich Filip von Daun. The message contains only a few lines of encrypted text, where the corresponding plain text is written above the cipher text.

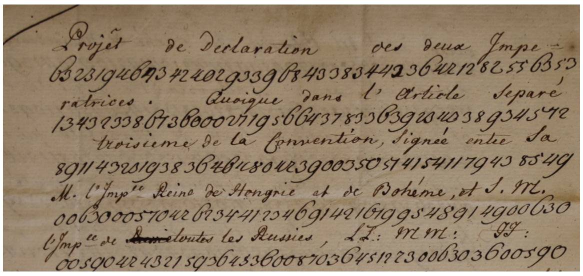


Figure 9: Cipher text example (Slovak National Archives, fond Amade-Üchtritz, box n. 141)

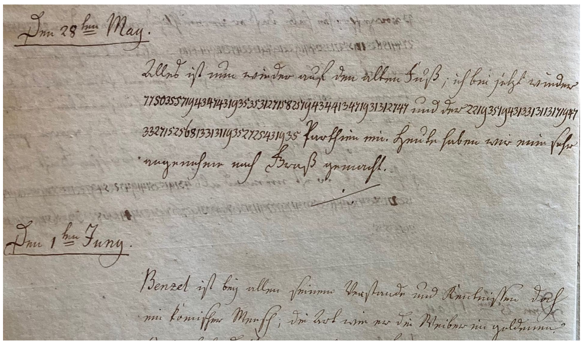


Figure 10: Encrypted diary parts (Slovak National Archives, fond Amade-Üchtritz, box n. 150)

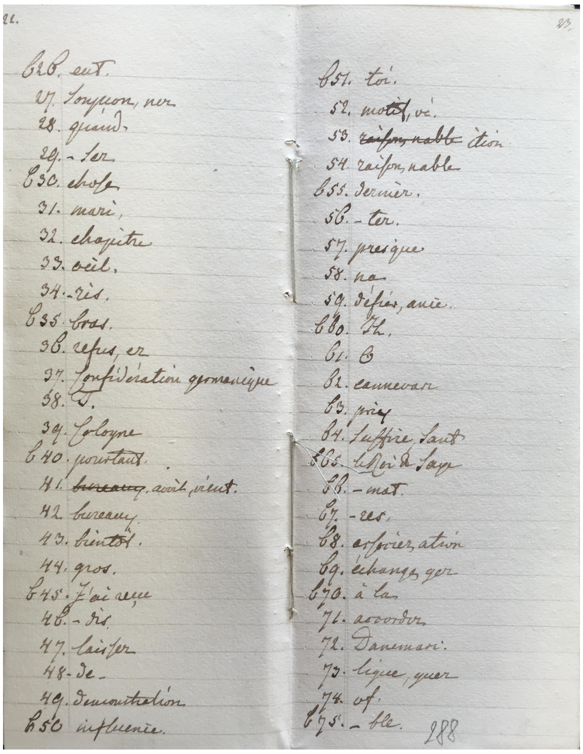
2.2 Cipher Keys

The examined fonds contain nine cipher keys. Seven cipher keys in the **Pálffy-Daun** family archive are (probably) related to the War of Spanish Succession. These (Spanish and Italian) cipher keys were used by Wirich Daun to communicate with several Counts and Marquises. The structure of these keys is very similar and copies the structure of a classical nomenclator system (Figure 15), each key is drawn on a large paper sheet and consists of simple/homophonic substitution[9], bigram (and trigram) substitution, codes, and nulls. The individual sub-encryption systems are graphically separated in a table.

There are two cipher keys in the **Amade-Üchtritz** family archive. One cipher key is for a small (German) nomenclator system, containing letter substitution and a few codes, which was used

[9] Homophones were commonly used only for vowels in these cipher keys.

by Emil Üchtritz to encrypt parts of his diary. The second (French) cipher key was also used by Emil Üchtritz and it is dated to 1831. This key consists of two small notebooks, *Chiffrant* for encryption and *Déchiffrant* for decryption (Figure 11), and contains three-digit numbers in a range 100 - 999.



22.

626.	eut.
27.	Soupçon, ner
28.	quand.
29.	- ser
630.	chose
31.	mari,
32.	chapitre
33.	oeil.
34.	- ris.
635.	bras.
36.	refus, er
37.	Confédération germanique
38.	D.
39.	Cologne
640.	pourtant.
41.	~~bureaux~~, avoit, vient.
42	bureaux
43.	bientôt.
44.	gros.
645.	J'ai reçu
46.	- dis.
47.	laisser
48-	de -
49.	demonstration
650	influence.

23.

651.	toi.
52.	motif, ve.
53.	~~raison, nable~~ ition
54.	raison, nable
655.	dernier.
56.	- ter.
57.	presque
58.	na
59.	défier, ance.
660.	Th.
61.	B
62.	cannevas
63.	prix
64.	suffire, sant
665.	le Roi de Saxe
66.	- mat.
67.	- res.
68.	associer, ation
69.	échange, ger
670.	à la
71.	accorder
72.	Danemark.
73.	ligue, guer
74.	of.
675.	- ble.

288

Figure 11: Cipher key example (Slovak National Archives, fond Amade-Üchtritz, box n. 138)

3 Historical Context

The encrypted documents in the **Esterházy** family archive date from the reign of *Maria Theresa* (1740–1780). At the beginning of her reign, she fought the so-called War of the Austrian Succession 1740–1748, and subsequently the so-called Seven Years' War (1756–1763). A significant role was played by diplomacy, and skilled diplomats, who managed to defend Maria Theresa's position in European politics. In this section, we give historical background on the prominent figures related to the analyzed encrypted documents.

Count Nicolaus Esterházy (1711–1764) was an imperial and royal chamberlain, counselor, guardian of the crown, and diplomat (Great Britain, Madrid, St. Petersburg). In Madrid, he was to conclude and sign the new treaty of alliance ending the War of the Austrian Succession, but he contracted severe poisoning from the infected water, from which he almost died. The Spanish-Austrian treaty of alliance was eventually signed without Esterházy's presence. (Khavanova, 2019) In the autumn of 1753, he was given the position of ambassador to St. Petersburg, but because of the Seven Years' War, Esterházy's mission was extended indefinitely. His task was to ensure the greatest possible integration of Russia into the Austro-French alliance contracted in 1753. On his return to Vienna in June 1762, Esterházy was appointed to the rank of captain of the Hungarian personal guard, at the same time a general. He died at the age of 53 in 1764 in Karlovy Vary. (Khavanova, 2016; Khavanova, 2017)

Wenzel Anton Prince of Kaunitz-Rietberg (1711–1794) was an envoy to Turin, Brussels, and Paris, and in 1748 he played an important role in the Peace of Aachen, which ended the War of the Austrian Succession. From his return to Vienna (1753) until his death, he served as a court and state chancellor of the Habsburg monarchy. He initiated the Theresian reforms in the civil service and the establishment of the Council of State (1760). Kaunitz was one of Maria Theresa's closest advisers. He resigned from his post after Franz II ceded some Polish territories to Prussia in 1793 and wanted to exchange Austrian possessions in the Netherlands for Bavaria. (Encyclopaedia Beliana, 2017; Schilling, 1994)

Wirich Philipp von Daun (1669–1741) was an Austrian field marshal in the Imperial Army in the War of the Spanish Succession under the command of Eugene of Savoy. He became famous for the successful defense of Turin in 1706. Daun became viceroy of Naples (1707–1708 and 1713–1719), and governor of the Austrian Netherlands (1724-1725). (Schmidt-Brentano, 2006; Kubeš et al., 2018) He was an imperial privy councilor and chamberlain. For his services on the Italian battlefields during the War of the Spanish Succession, he received the Italian noble titles of Marquis of Rivoli (1706) and Duke of Teano (1710). Italian and Spanish cryptographic materials from the **Pálffy-Daun** family archives are also linked to these titles.

The encrypted documents in the **Amade-Üchtritz** family archives relate to personalities representing the Kingdom of Saxony at the Viennese court. Most of them date from the first two decades of the 19th century, especially from the end of the Napoleonic Wars, when Saxony found itself as an ally of Napoleon among the defeated countries, and through its envoys sought to regain lost positions, especially the revision of territorial losses.

Detlev Graf von Einsiedel (1773-1861) was a Saxon businessman, minister, and confidant of the Saxon kings Frederick Augustus I and his successor Anton. In foreign policy, he advocated a close alliance with Austria, seeking to strengthen Saxony's sovereignty by brokering new dynastic alliances with European ruling families. (Wetzel, 2014) In this he was supported by his brother-in-law *Emil von Üchtritz* (1783-1841), a long-time Saxon envoy to the Viennese court. He stood in the service of King Frederick Augustus of Saxony, especially during his captivity after the Battle of Leipzig in 1813, and worked for the restoration of the Saxon kingdom. He then served as an envoy to France and, from 1830 to Vienna. His son Emil was an officer, married into the Hungarian **Amade** family. (Wurzbach, 1883)

4 Digital Processing of Historical Manuscripts

Our valuable collection of historical manuscripts was converted into a digital form so it can be processed by modern computer algorithms to analyse statistical properties and even solve the documents encrypted by nomenclator. The image resolution of the original images is 4160 × 6240 pixels. Documents vary in quality and readability. Some of the documents contain encrypted and non-encrypted parts on a single page. The portion of the collection contains documents with digits formed in groups delimited by separators. Automated image processing needs to take all these factors into consideration.

Antal and Marák (2022) proposed an automated software system for historical handwritten digit detection, recognition, and transcription based on deep learning. This way, a large number of historical documents can be analyzed and solved in an automated manner. They provided a detailed description of a digit detector based on Mask R-CNN (He, 2017) method along with technical details related to the dataset, manually created annotations, training, inference, and transcription algorithm. Their Mask R-CNN based detector works well in general, however, it has some limitations which need to be addressed.

Limitations of the digit detector proposed in (Antal and Marák, 2022):

- Mask R-CNN detector along with ResNet as its backbone network require a large amount of GPU memory which led to unavoidable optimizations. They needed to resize their original images to a smaller resolution to be able to fit data into memory and run training.

- Handwritten digit detection is a challenging task where small and densely grouped objects (digits) are located in a large document. Most detectors, where Mask R-CNN is no exception, fail to reliably detect such objects. For this reason, they had to employ a strategy of dividing an input image into smaller 128 × 128 blocks. Training Mask R-CNN on the image blocks led to significant improvement in digit detection. Unluckily, it has introduced a serious downside. The image division may lead to splitting of a single digit into two different image blocks. As a result, this solution misses digits positioned near the border of each image block.

- Mask R-CNN was introduced in 2017 and is no longer considered as state-of-the-art detector. According to the benchmark (Papers with code, 2023), there are new families of modern multi-class object detectors achieving average precision approximately 25 % higher than Mask R-CNN. Moreover, modern detector architectures make detection of small objects possible without having to split the image into small blocks.

These problems motivated us to design, train, and test a new digit detector based on YOLOv7 (Wang et al., 2022), which is a faster and more accurate deep learning model than Mask R-CNN. Furthermore, YOLOv7 can easily be configured to detect small objects using its auto-anchor algorithm. Its official implementation and documentation can be found at (YOLOv7 GitHub, 2023). In its essence, YOLOv7 is an object detector producing rectangular bounding boxes for object instances. However, it has a special architectural tweak which allows it to run in instance segmentation mode to produce high precision polygonal masks for detected objects.

Handwritten digit transcription based on YOLOv7 is a procedure consisting of several steps:

1. Preparing the software environment and dependencies for YOLOv7.

2. Obtaining the YOLOv7 model pretrained on COCO dataset (Tsung-Yi Lin et al., 2014).

3. Preparing the training, validation and testing dataset.

4. Converting digit annotations from COCO format to YOLOv7 PyTorch format.

5. Configuring the model and setting hyperparameters.

6. Training the model using transfer learning.

7. Testing the model.

8. Digit transcription.

4.1 Preparation Stage

All the experiments were carried out on a computer with the following specification: AMD Ryzen 9 5900HX, GeForce RTX 3080 16 GB, 32GB DDR4 3200MHz. In order to have a robust detector and minimize training time, we employed a transfer learning technique. We obtained a pretrained model on the COCO dataset. The pretrained model accepts images of size 1280 × 1280, so we needed to downscale our original images. Our image dataset has 18 images and was split into training (12 images), validation (3 images) and testing (3 images) sets. We collected more than twelve thousand[10] polygonal annotations of digits 0-9 which are stored in well-known COCO format. More details about the annotations can be found in (Antal and Marák, 2022). Since our YOLOv7 performs digit detection on the entire image rather than on image blocks, we needed to transform geometric properties of annotations so they correspond to images of size 1280 × 1280. The annotations were converted from COCO format to YOLOv7 PyTorch format using Roboflow framework (Roboflow, 2023).

4.2 Model Configuration, Hyperparameters and Training

The training took 300 epochs and the batch size was set to 1 due to GPU memory limitations. The learning rate was set to 0.0025 using a stochastic gradient descent optimizer. We also performed slight image augmentation to enhance the variability of our dataset. Namely, we generated augmented images by modifying hue, saturation, and brightness. Geometric transformations were omitted. Table 1 shows the values of standard performance indicators measured on the test dataset for the best model after 300 epochs of training.

Masks are polygons produced by the detector which determine the area of the object. Bounding boxes are rectangular areas defining the object location. Compared to masks, bounding boxes are less accurate estimations of object location when the object has an irregular shape. $mAP_{@.50}$ and $mAP_{@.50:.95}$ indicators represent mean average precision which is a common metric to evaluate the performance of the object detector. *mAP* indicator quantifies the detection accuracy measured for different IoU (Intersection over Union) values. IoU measures the degree of overlap between the predicted bounding box/mask and the ground truth box/mask. *mAP* is the average of AP (Average Precision) values where AP corresponds to the area under the precision-recall curve. The precision is calculated as the ratio of correctly predicted positive examples divided by the total number of positive examples that were predicted. The recall is a metric that quantifies the number of correct positive predictions made out of all positive predictions that could have been made.

Mask accuracy			
$mAP_{@.50}$	$mAP_{@.50:.95}$	Precision	Recall
91.3 %	60.9 %	88.2 %	91.0 %
Bounding box accuracy			
$mAP_{@.50}$	$mAP_{@.50:.95}$	Precision	Recall
91.6 %	66.8 %	88.6 %	91.1 %

Table 1: YOLOv7 digit detection performance on the test dataset after 300 training epochs

4.3 Testing

We evaluated the model inference using the confidence threshold of 50 % (i.e. only digits with at least 50% confidence are included in the final set of detections). We experimented with standard quality images to judge the general performance. The detector turned out to have only minor problems with missed out or misclassified digits. Figure 12 shows the detected bounding boxes and masks in a sample document.

[10] 12,433

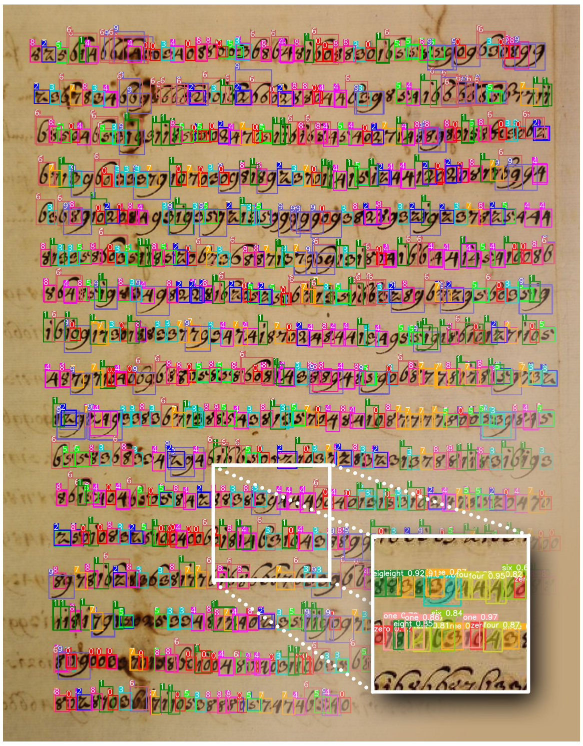


Figure 12: Result of digit detection and segmentation

Since our solution currently does not perform any image preprocessing, we also investigated the detection performance on images of lower quality. YOLOv7 was able to cope with problems like noisy paper background and areas where digits collided with the text from the reverse side of the paper. There are documents with encrypted and non-encrypted parts on a single page. In such situations we need to avoid false digit detections in non-encrypted parts. As we can see in Figure 13, this is where YOLOv7 performed quite well producing only minimum detections in areas outside the encrypted regions.

Last but not least, we sometimes encounter digit separators (dots) which are placed in digit sequences. We investigated the impact of the separators on detection performance. We found that separators do not have adverse effects on detection accuracy. Sample output from the aforementioned testing scenario is depicted in Figure 14.

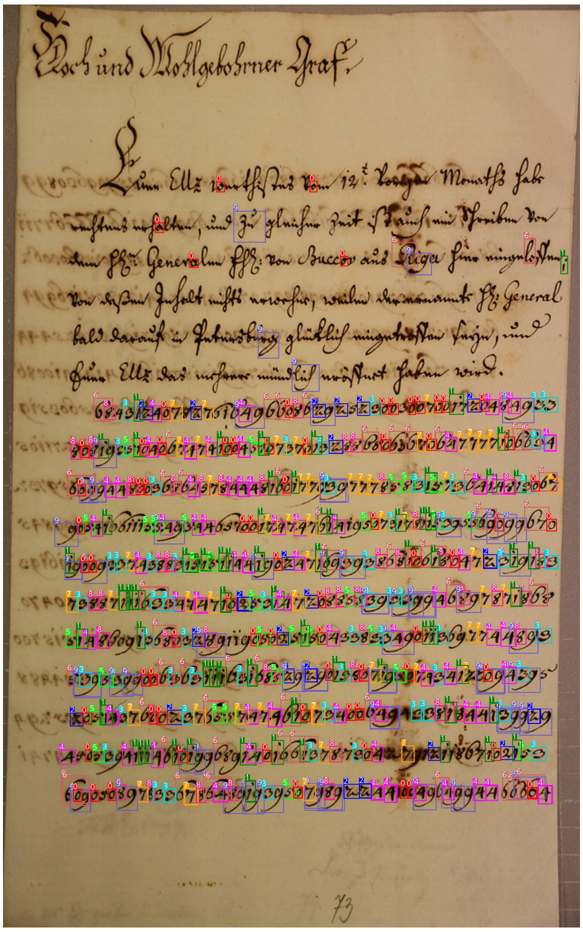


Figure 13: Digit detection performance in a document combining encrypted (at the bottom) and non-encrypted parts (at the top)

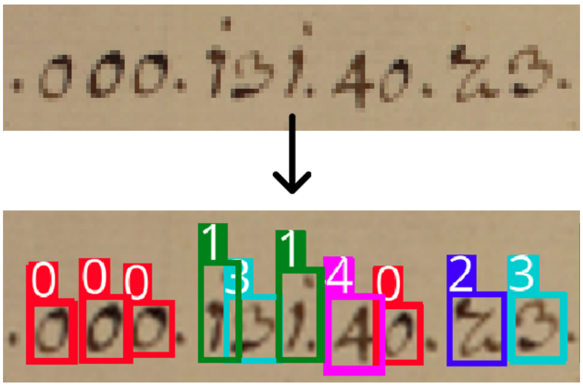


Figure 14: Digit detection performance in a document containing digit separators

4.4 Transcription

Once digits are detected by YOLOv7, all bounding boxes and their corresponding class labels are passed to the automated transcription algorithm presented in (Antal and Marák, 2022) which detects lines and extracts the digits. Digits are then stored in a text file. Figure 16 shows the intermediate steps of automated transcription of the encrypted document.

4.5 Future Work

In this paper we have presented our YOLOv7 based handwritten digit detector which is able to detect and classify large number of digits on a paper with acceptable performance. In addition, besides standard bounding boxes, it also supports polygonal masks for even higher precision. We have also eliminated the problem of Mask R-CNN detector where input images must be divided into small blocks leaving digits at the borders undetected. YOLOv7 solves this using its auto-anchor algorithm and detects digits in the entire document.

We plan to further improve our digit detection algorithm by extending the training dataset and incorporating geometric augmentations. It is crucial to train the detector on digits written using different writing styles and images of various quality. Currently, we deal with more research problems including segmentation of handwritten text area within the document, detection of special digit markups and glyphs. It is equally important to apply deep learning methods to segmentation and extraction of structured information from nomenclator keys.

5 Conclusions

During our research of historical encrypted documents in the Slovak National Archive in Bratislava we found several encrypted documents and keys in the archives of the aristocratic families Amade-Üchtritz, Esterházy, and Pálffy-Daun. Some of the documents are still encrypted, without known decryption or preserved corresponding key. To facilitate the potential decryption of these documents we focus on different tasks:

- systematic collection and examination of the materials, trying to match the documents with keys, or at least with similar documents with known decryption;
- historical study of the related time period, persons, and their relations, to provide insight on potential location of keys and potential key words for decryption attempts;
- and finally providing a supporting automation, mainly in the transcription of documents.

We believe that only with proper computer automation the researchers would be able to properly examine the large number of historical documents contained in the archives. As presented, new machine learning methods have promising results with good precision. However, the automated transcription of documents with machine learning does always lead to some detection and transcription errors (FAR, FRR). Unlike in plain text documents, such errors are difficult to detect and correct, because the ciphertext is seemingly a random sequence of characters. The ultimate goal of automatic decryption would require a further research in decryption methods that can tolerate some amount of transcription failures.

Acknowledgments

This work was supported by grant VEGA 2/0072/20: *Modern processing methods of encrypted archival documents*, and grant APVV-21-0371: *The glamour and fall of nobility. Strategies of noble representation in the History of Slovakia*.

We thank Jakub Mírka for his help, and Benedek Láng for providing materials from Austria and Hungary.

References

Eugen Antal and Pavol Zajac. 2020. HCPortal Overview. In *Proceedings of the 3rd International Conference on Historical Cryptology, HistoCrypt 2020*, pages 18 - 20. Linköping University Electronic Press.

Eugen Antal and Pavol Zajac. 2021. HCPortal Modules for Teaching and Promoting Cryptology. In *Proceedings of the 4th International Conference on Historical Cryptology, HistoCrypt 2021*, pages 1 - 11. Linköping University Electronic Press.

Eugen Antal and Pavol Marák. 2022. Automated transcription of historical encrypted manuscripts. In *Tatra Mountains Mathematical Publications, 82 (2022)*, pages 65 - 86. Slovak Academy of Sciences.

Eugen Antal and Jakub Mírka. 2022. Wrong Design of Cipher Keys: Analysis of Historical Cipher Keys From the Hessisches Staatsarchiv Marburg Used in the Thirty Years' War. In *Proceedings of the 5th International Conference on Historical Cryptology, HistoCrypt 2022*, pages 1 - 11. Linköping University Electronic Press.

Encyclopaedia Beliana. 2017. Kaunitz, Wenzel Anton. Available online: `https://beliana.sav.sk/heslo/kaunitz-wenzel-anton`. ISBN 978-80-89524-30-3.

Kaiming He et al.. 2017. Mask R-CNN. In *IEEE International Conference on Computer Vi-*

sion (ICCV), 2017, pages 2980-2988. DOI: 10.1109/ICCV.2017.322.

Olga Khavanova. 2016. Farewell to Sent-Petersburg: count Nicolas Esterházy leaves the Russian capital (orig. Búcsú Szentpétervártól: Gróf Esterházy Miklós elhagyja az orosz fővárost). In *Aetas* 31 / 4, pages 188 - 199.

Olga Khavanova. 2017. The career of count Nicolas Esterházy Ambassador of St. Petersburg (orig. Gróf Esterházy Miklós szentpéteri nagykövet karrierje). In *Az Esterházyak fraknoi ifjabb ága.*, pages 60 - 75. Mesto Senec.

Olga Khavanova. 2019. From Madrid To St. Petersburg. Count Miklós Esterházy, the first Hungarian Career Diplomat of Maria Theresa (orig. Madridtól Szentpétervárig. Gróf Esterházy Miklós, Mária Terézia első magyar karrierdiplomatája). In *Századok* 153/6, pages 1123 - 1128.

Jiří Kubeš et al. 2018. On behalf of the Emperor. Czech and Moravian aristocracy in Habsburg diplomacy 1640-1740 (orig. V zastoupení císaře. Česká a moravská aristokracie v habsburské diplomacii 1640–1740). ISBN: 9788074225741 Praha : NLN.

Benedek Láng. 2020. Was it a Sudden Shift in Professionalization? Austrian Cryptology and a Description of the Staatskanzlei Key Collection in the Haus-, Hof- und Staatsarchiv of Vienna. In *Proceedings of the 3rd International Conference on Historical Cryptology, HistoCrypt 2020*, pages 87 - 95. Linköping University Electronic Press.

Beáta Megyesi, Bernhard Esslinger, Alicia Fornés, Nils Kopal, Benedek Láng, George Lasry, Karl de Leeuw, Eva Pettersson, Arno Wacker and Michelle Waldispühl. 2020. Decryption of historical manuscripts: the DECRYPT project. In *Cryptologia*, volume 44, number 6, pages 545-559. Taylor & Francis.

Beáta Megyesi, Crina Tudor, Benedek Láng and Anna Lehofer. 2021. Key Design in the Early Modern Era in Europe. In *Proceedings of the 4th International Conference on Historical Cryptology, HistoCrypt 2021*, pages 121 - 130. Linköping University Electronic Press.

Beáta Megyesi, Crina Tudor, Benedek Láng, Anna Lehofer, Nils Kopal, Karl de Leeuw and Michelle Waldispühl. 2022. Keys with nomenclatures in the early modern Europe. In *Cryptologia*, DOI: 10.1080/01611194.2022.2113185. Taylor & Francis.

Papers with code. 2023. Object Detection on COCO test-dev. `https://paperswithcode.com/sota/object-detection-on-coco`.

Roboflow. 2023. Roboflow. `https://roboflow.com/`.

Lothar Schilling. 1994. Kaunitz und das renversement des alliances. Studien zur außenpolitischen Konzeption Wenzel Antons von Kaunitz. In *Historische Forschungen*, volume 50. Berlin : Duncker & Humblot.

Antonio Schmidt-Brentano. 2006. Kaiserliche und k.k. Generale (1618-1815) Österreichisches Staatsarchiv/A. Schmidt-Brentano.

Tsung-Yi Lin et al.. 2014. Microsoft COCO: Common Objects in Context. In *CoRR*, abs/1405.0312 `http://arxiv.org/abs/1405.0312` Dataset URL: `https://cocodataset.org/#home`

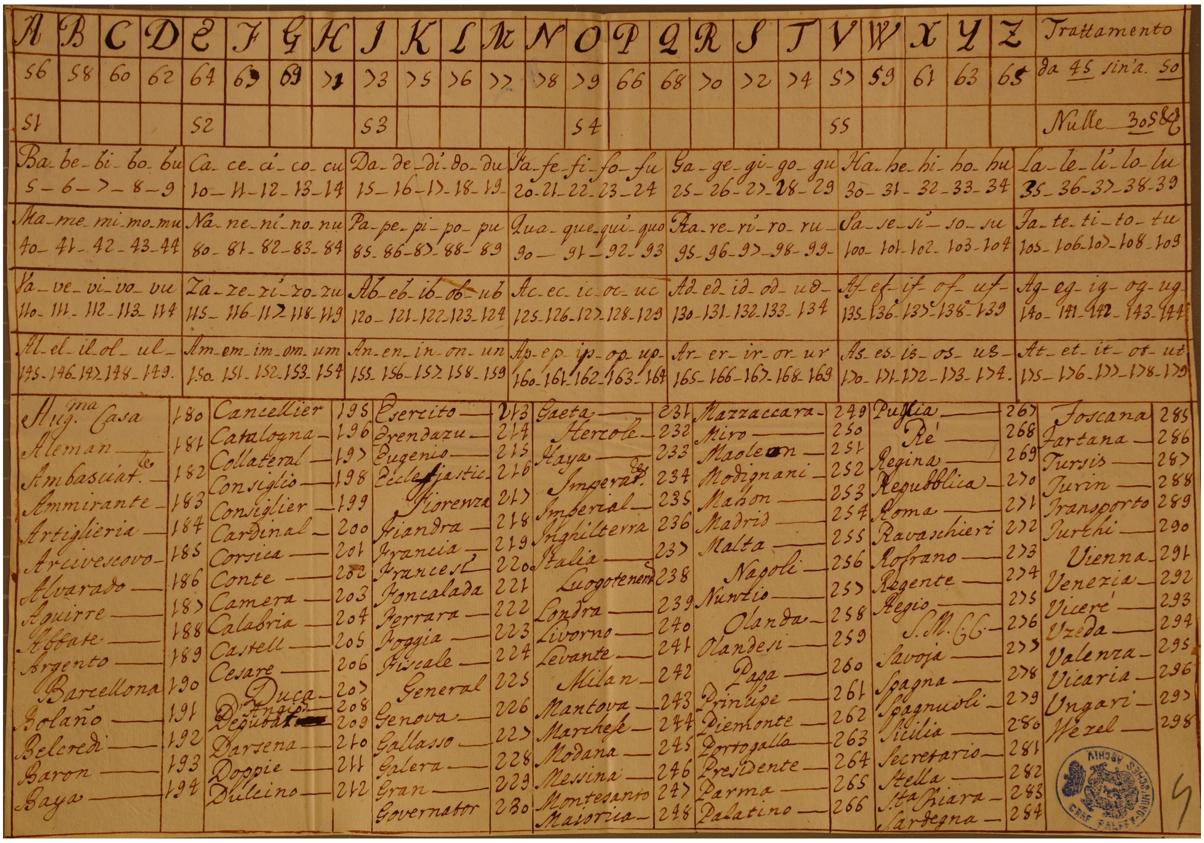
Wang Chien-Yao, Bochkovskiy Alexey, Liao Hong-Yuan Mark. 2022. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *ArXiv*, pages 2980-2988. DOI: 10.48550/ARXIV.2207.02696 `https://arxiv.org/abs/2207.02696`.

Michael Wetzel. 2014. Detlev Graf von Einsiedel. Available online: `https://saebi.isgv.de/biografie/Detlev_Graf_von_Einsiedel_(1773-1861)`.

Constant von Wurzbach. 1883. Üchtritz, Emil von. In *Biographisches Lexikon des Kaiserthums Oesterreich Trzetrzewinsky-Ullepitsch*, Band 48. Verlag der Universitäts-Buchdruckerei von L. C. Zamarski. Wien.

YOLOv7 GitHub. 2023. Official YOLOv7 GitHub repository. `https://github.com/WongKinYiu/yolov7`.

Appendices



A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	V	W	X	Y	Z	Trattamento
56	58	60	62	64	67	69	71	73	75	76	77	78	79	66	68	70	72	74	57	59	61	63	65	da 45 sin'a. 50
51				52				53					54						55					Nulle 305 &c

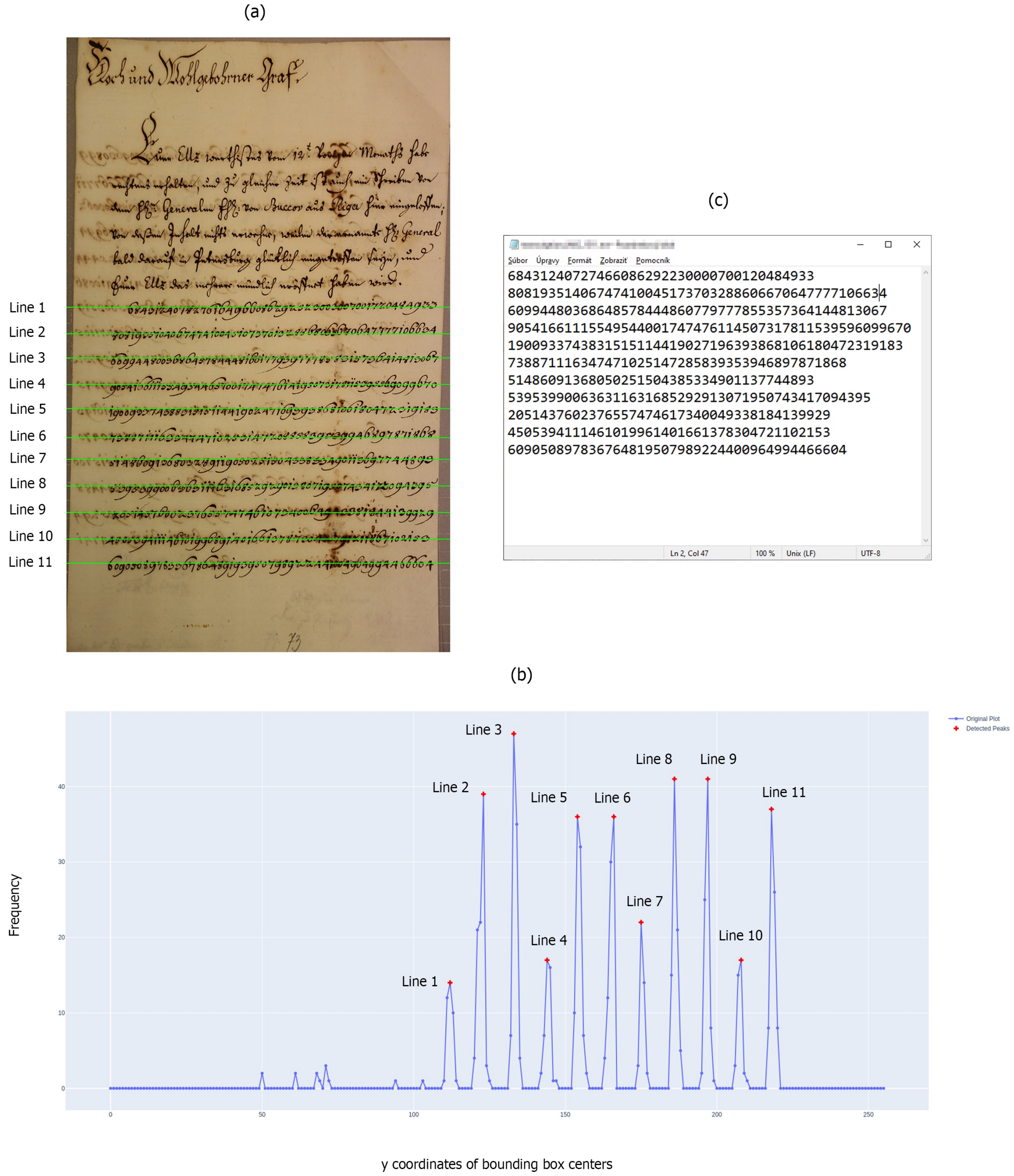
Ba-be-bi-bo-bu 5-6-7-8-9	Ca-ce-ci-co-cu 10-11-12-13-14	Da-de-di-do-du 15-16-17-18-19-	Fa-fe-fi-fo-fu 20-21-22-23-24	Ga-ge-gi-go-gu 25-26-27-28-29	Ha-he-hi-ho-hu 30-31-32-33-34	La-le-li-lo-lu 35-36-37-38-39
Ma-me-mi-mo-mu 40-41-42-43-44	Na-ne-ni-no-nu 80-81-82-83-84	Pa-pe-pi-po-pu 85-86-87-88-89	Qua-que-qui-quo 90-91-92-93	Ra-re-ri-ro-ru- 95-96-97-98-99-	Sa-se-si-so-su 100-101-102-103-104	Ta-te-ti-to-tu 105-106-107-108-109
Va-ve-vi-vo-vu 110-111-112-113-114	Za-ze-zi-zo-zu 115-116-117-118-119	Ab-eb-ib-ob-ub 120-121-122-123-124	Ac-ec-ic-oc-uc 125-126-127-128-129	Ad-ed-id-od-ud- 130-131-132-133-134	Af-ef-if-of-uf- 135-136-137-138-139	Ag-eg-ig-og-ug 140-141-142-143-144
Al-el-il-ol-ul- 145-146-147-148-149.	Am-em-im-om-um 150-151-152-153-154	An-en-in-on-un 155-156-157-158-159	Ap-ep-ip-op-up- 160-161-162-163-164	Ar-er-ir-or-ur 165-166-167-168-169	As-es-is-os-us- 170-171-172-173-174.	At-et-it-ot-ut 175-176-177-178-179

Aug.ma Casa	180
Aleman	181
Ambasciat.re	182
Ammirante	183
Artiglieria	184
Arcivescovo	185
Alvarado	186
Aguirre	187
Abbate	188
Argento	189
Barcellona	190
Bolaño	191
Belcredi	192
Baron	193
Baya	194
Cancellier	195
Catalogna	196
Collateral	197
Consiglio	198
Consiglier	199
Cardinal	200
Corsica	201
Conte	202
Camera	203
Calabria	204
Castell	205
Cesare	206
Duca	207
D'Angiò	208
Deputat	209
Darsena	210
Doppie	211
Dulcino	212
Esercito	213
Erendazu	214
Eugenio	215
Ecclesiastic	216
Fiorenza	217
Fiandra	218
Francia	219
Francesi	220
Foncalada	221
Ferrara	222
Foggia	223
Fiscale	224
General	225
Genova	226
Gallasso	227
Galera	228
Gran	229
Governator	230
Gaeta	231
Hercole	232
Haya	233
Imperat.ce	234
Imperial	235
Inghilterra	236
Italia	237
Luogotenen.te	238
Londra	239
Livorno	240
Levante	241
Milan	242
Mantova	243
Marchese	244
Modana	245
Messina	246
Montesanto	247
Maioria	248
Mazzaccara	249
Miro	250
Maoleon	251
Modignani	252
Mahon	253
Madrid	254
Malta	255
Napoli	256
Nuntio	257
Olanda	258
Olandesi	259
Papa	260
Principe	261
Piemonte	262
Portogallo	263
Presidente	264
Parma	265
Palatino	266
Puglia	267
Re'	268
Regina	269
Repubblica	270
Roma	271
Ravaschieri	272
Rofrano	273
Regente	274
Regio	275
S.M.CC.	276
Savoja	277
Spagna	278
Spagnuoli	279
Sicilia	280
Secretario	281
Stella	282
S.ta Chiara	283
Sardegna	284
Toscana	285
Tartana	286
Tursis	287
Turin	288
Transporto	289
Turchi	290
Vienna	291
Venezia	292
Vicere'	293
Uzeda	294
Valenza	295
Vicaria	296
Ungari	297
Wezel	298

GRAF PÁLFFY-DAUN'SCHES ARCHIV

5

Figure 15: Cipher key example (Slovak National Archives, fond Pálffy-Daun, Klasse XXXIII - Wierich Daun, box n. 38, fasc. 22)



```
684312407274660862922300007001204849 33
```

Figure 16: Intermediate results of our automated handwritten document transcription: (a) original document with highlighted positions of detected lines, (b) histogram of *y* coordinates of bounding box centers (peaks indicate line positions), (c) text file containing the transcribed digits

The cipher of Emperor Rudolf II's "Alchemical Hand Bell"

Richard Bean
School of Information Technology
& Electrical Engineering
University of Queensland
Australia 4072
r.bean1@uq.edu.au

Corinna Gannon
Städel Museum
Dürerstrasse 2
60596 Frankfurt am Main
Germany
gannon@staedelmuseum.de

Sarah Lang
University of Graz
Elisabethstraße 59/III
8042 Graz
Austria
sarah.lang@uni-graz.at

Abstract

We examine a cipher found inscribed in the so-called "Alchemical Hand Bell" from the Kunstkammer of Emperor Rudolf II. We provide insight into the bell's history, a correction for an existing published transcription, perform statistical analysis of the ciphertext, and look at possible encryption methods and plaintext languages. Given the analysis, we examine the possibilities of digraphic and polyphonic ciphers and give a brief overview of how these were used in the historical context.

1 Introduction and Description

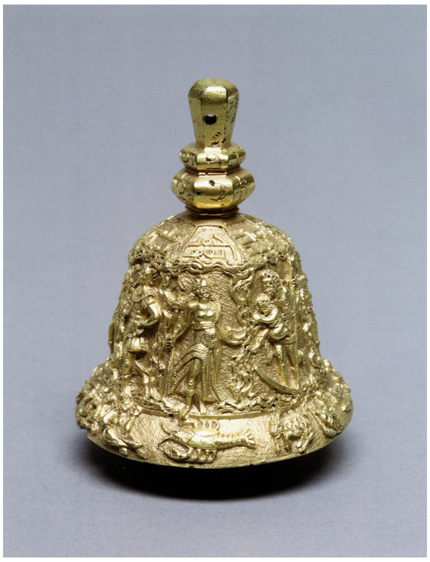


Figure 1: Hans de Bull, "Alchemical Hand Bell" of Emperor Rudolf II, ca. 1600, h. 7,8 cm; d. 6,3 cm, Vienna, Kunsthistorisches Museum, inv. no. Kunstkammer, 5969. https://www.khm.at/objektdb/detail/91976/. Source of images: Gannon (2019).

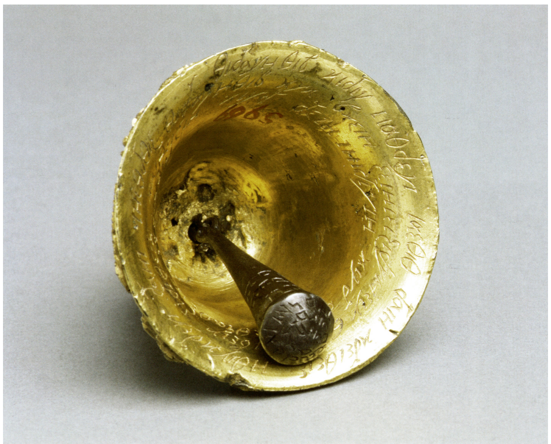


Figure 2: See figure 1.

Around 1600, the Prague goldsmith Hans de Bull cast two hand bells for Emperor Rudolf II (1552–1612). One of them has survived the past four centuries and is nowadays on display in the Kunsthistorisches Museum (KHM) in Vienna (figures 1 and 2).[1] This little *Kunstkammer* piece is fascinating in many ways. From a letter by the artist we know that it was cast from an alloy of the seven planetary metals – gold, silver, copper, iron, lead, tin and mercury – before it was gilt.[2] This was confirmed in a recently conducted XRF-analysis at the KHM.[3] Such a sevenfold alloy had been described by the Swiss physician and alchemist Paracelsus (1493/4–1541) who called it *Electrum* and provided astrological and alchemical instructions for creating efficacious artefacts out of it in his text corpus *Archidoxis magica* (Huser, 1590, Appendix, pp. 115–130).

[1] Hans de Bull, "Alchemical Hand Bell" of Emperor Rudolf II, ca. 1600, h. 7,8 cm; d. 6,3 cm, Vienna, Kunsthistorisches Museum, inv. no. Kunstkammer, 5969. https://www.khm.at/objektdb/detail/91976/.

[2] Prague, archive Pražského hradu, Dvorská komora, box 5, no. 698 http://documenta.rudolphina.com/Regesten/A1612-10-00-02669.xml

[3] The results of this analysis will appear in Gannon (2024).

The sound of bells made from *Electrum* was supposed to summon planetary spirits and deities in order to provide their user insight into the secrets of the cosmos and therefore wisdom and power.[4] Emperor Rudolf II, who admired the Paracelsian philosophy and fostered the study of alchemy and natural magic at his court, must have been fascinated by such a promising material and likely appreciated a successful realization for his exquisite collection.

Whereas the "Alchemical Hand Bell's" intellectual and art historical background have been studied and reconstructed extensively (Bukovinská and Purš, 2010; Tilton, 2015; Gannon, 2019; Gannon, 2023b; Gannon, 2023a; Gannon, 2024)[5], a cryptological riddle remains to be solved. The decoration of the bell's mantle allows for a straightforward interpretation – the seven full-figure planetary deities, the corresponding signs of the zodiac, the symbols of the seven planets and metals, as well as a number of pseudo-Chaldean and pseudo-Arabic letters, visualize the interdependence between macro- and microcosm. A mystery is the spiraling Greek inscription that was carved into the inside of the bell's mantle. Each of the 163 letters can be identified; however, the inscription seems to contain no meaning. The iron clapper was also engraved with a spiraling Hebrew script, yet, the Hebrew letters are hardly legible and cannot fully be transcribed. Even though magical artifacts are often inscribed with nonsense script, for example another talisman created for Emperor Rudolf II whose reverse is adorned with similarly meaningless Hebrew letters (Gannon, 2020) – preferably corrupted Greek or Hebrew – it is tempting to suggest that at least the Greek sequence of letters may contain an encrypted message. It was not uncommon to hide alchemical recipes under a cipher (Piorko et al., 2023).

Also, since the bell was supposed to be used for summoning supernatural beings by calling their names, it is not unlikely that the inscription contains a list of such names. The English magus John Dee (1527–1608/09), for example, who sojourned in Prague with his 'scryer', the alchemist Edward Kelly (1555–1597), and tried to win the emperor's favor, practiced a comparable form of angelic magic. Similar artifacts and names to communicate with celestial beings were involved in his "angelic conversations" (Harkness, 1999; Clucas, 2006). In many cases, they appear to be a random stringing together of letters (Turner, 1986, p. 73). The question of whether this is also the case with the Greek inscription inside the "Alchemical Hand Bell" or whether it can indeed be deciphered as a legible plaintext remains to be answered. In the following, possible approaches will be presented.

2 Statistical analysis

The first published transcription of the ciphertext by Gannon (2019) contained a slight mistake which can be corrected as follows.

> ϑιδαγΗ ϑιβ κιδιγ ιιαϑδεγι ιαεϑιϑ δαιΗ κδειϑειζ ΗϑιγκδεγιδαΗι ιΗεϑδϑιζ ϑιδαγ Ηϑιβ κγκ βκειΗ ζειΗιει ζιδγΗειγ ϑιβ ιγαιβειγ ζιδιϑειΗ καιϑειζιΗ κιγδ δειΗ ιΗιδιγιΗ κιγδ δειΗ ΗεϑιαϑζειγζεϑιΗϑιΗ

Notably, each of the Greek letters in the ciphertext is from the first ten letters in the Greek alphabet. A transliteration of this into numerals would be:

> 783026 781 98382 88073428 804787 3086 93487485 678293428 3068 86473785 78302 6781 929 19486 5486848 58326482 781 82081482 58387486 908748586 9823 3486 86838286 9823 3486 6478075482 54786786

The frequency count is as follows (table 1).

From here on, we will use digits to represent each character, as contemporaneous ciphers with ten different characters often used the digits 0 through 9 and the letters are the initial letters of the Greek alphabet. We will also describe each block as a "word". Preserving the spacing, there are three repeated words: 781, 3486 and 9823. We see that in fact "9823 3486" is a two-word phrase repeated twice. Excluding spaces, we note two ten-number repetitions: 7830267819 and 8698233486. There are 53 unique bigrams (of a possible 100) with, for example, 86 and 48 occurring 12 times, and 78 occurring 11 times. The average word length is 6.04 with standard deviation

[4] On the ritual accompanying the use of the bell: Gannon (2019), Gannon (2024). On the connection between music and magical practices see Gannon (2023a).

[5] A detailed study of the "Alchemical Hand Bell" is part of Corinna Gannon's dissertation submitted in November 2022 at Goethe University, Frankfurt am Main and will be published in the future.

Greek Letter	Digit	Count
Α α	0	9
Β β	1	5
Γ γ	2	15
Δ δ	3	16
Ε ε	4	18
Ζ ζ	5	8
Η η	6	18
Θ ϑ	7	18
Ι ι	8	47
Κ κ	9	9

Table 1: Frequency count of ciphertext letters.

2.19. However, as the ciphertext is very short, this does not assist much with plaintext language identification, or for distinguishing between plaintext language candidates. In a cipher from around the same time period, Bean, Lang and Piorko (2022) noted that an observed average word length of 5.8 was too long for English but in line with Latin.

The calculated index of coincidence (IC) of the ciphertext is 3745/163/162 = 0.142. For the digraphic index of coincidence (DIC), it is 0.0269. Compared to Mason's table (Mason, 2005) we see that random digit ciphers, naturally, have an IC and DIC of 0.1 and 0.01. These values are quite distinct from those observed here. We note that 18 of the 27 words contain an even number of Greek letters. To check how common this is if the plaintext were ordinary Latin text, we performed sampling from 93 Project Gutenberg Latin books.[6] We generated one million examples of 27 word texts with the same proportions as from the books; about 1 in 12 samples had at least 18 of 27 words with even lengths.

We estimate the probabilities of the plaintext language as: Latin 60%, Greek 30% and German 10%. These estimates are based on the vernacular of Rudolf's court, the context of the cipher, and the contemporary context of objects of similar vintage. Another less likely option is Czech. Also, Hebrew inscriptions, as found on the clapper of the bell, are quite common, given Rudolf's interest in Kabbalah.[7]

3 Cipher type Diagnosis

These basic observations are a good starting point to try to diagnose the cipher type (Callimahos, 1977, Chapter XI). Two diagnosis tools, based on ACA ciphers and using machine learning techniques are currently available online: Mason (BION)[8] and the tool from Leierzopf et al. (2021) known as "NCID" [9]. Both tools provide a "probability" score in percentage terms ranking various possible ciphers. Using an input of digits, Mason's tool suggests the two most likely ciphers are Monome-Dinome (73) and Tridigital (23). With the ciphertext input as English letters, the top two outputs are Bazeries (25) and CheckerBoard (20). These are clearly anachronistic suggestions. The NCID tool has also been trained on ciphertext from the "key-phrase" cipher. Using English letter input, two reasonable suggestions are Checkerboard (48%) and key-phrase (5%). Other ciphers from the ACA list with similar index of coincidence statistics are shown in the following table; that is, selected rows from Mason's table. These four ciphers have output in the form of numbers and use a key square, a matrix, or a 5x5 Polybius square in the encipherment process. Note that these statistics are based on enciphered English plaintext using English keywords, which would have different statistical properties to Latin, Greek or German plaintext. The statistics are given as two values (table 2): the IC is multiplied by 1,000 while the DIC is multiplied by 10,000, and each value is given as a mean / standard deviation pair.

ACA Cipher Type	IC	DIC
Grandpré	128/3	179/15
Monome-dinome	124/7	249/36
Tridigital	122/8	195/29
Nihilist substitution	144/11	218/33

Table 2: Monographic and digraphic index of coincidence statistics for selected ACA ciphers. **IC** = Index of coincidence (mean/sd) times 1,000; **DIC** = Digraphic index of coincidence (mean/sd) times 10,000.

The Grandpré cipher was first introduced in 1905; the monome-dinome cipher is believed to date from the Spanish Civil War (c1936), while the Nihilist cipher dates from 19th century Russia. Thus, these ciphers in their current form cannot have been used for the hand bell. Other ACA ciphers with numerical output such as the "Pol-

[6] `https://www.gutenberg.org/`.

[7] See Gannon (2020).

[8] `https://bionsgadgets.appspot.com/gadget_forms/refscore_extended.html`.

[9] `https://www.cryptool.org/en/cto/ncid`.

lux" or "Morbit" cannot be in use as they are derived from Morse code which was developed in the 1840s. As noted, the table from Mason does not provide statistics for the "key phrase" cipher. The observed digraphic index of coincidence (0.0269) is quite high, as is the number of words with even length. Thus, a digraphic cipher cannot be excluded. However, the simplest interpretation would be that spaces indicate plaintext word divisions. Historically, the first digraphic cipher was described in della Porta (1563) which provided a 20 x 20 table mapping every combination of two letters to a unique symbol. A simpler process was not developed until the "Playfair" cipher of 1854 which mapped letter pairs using a 5 x 5 Polybius square.

4 Polyphonic ciphers

Given these statistics and observations, the cipher may well be a so-called "polyphonic cipher" where any single ciphertext letter can map to many plaintext letters. One example of a polyphonic cipher is the so-called "key-phrase" cipher described in Kahn (1996, 787), Gaines (1956, 103) and used as a common cipher in American Cryptogram Association challenges. The ACA website gives an example as follows, using the phrase "Give me liberty or give me death".[10]

```
pt alphabet: abcdefghijklmnopqrstuvwxyz
CT alphabet: GIVEMELIBERTYORGIVEMEDEATH

pt: aciphertextlettermaystandfor
CT: GVBGIMVMMAMTMMMMVYGTEMGOEERV
pt: morethanoneplaintextletter.
CT: YRVMMIGOROMGTGBOMMAMTMMMMV.
```

However, Kahn seems to indicate that this cipher was limited to one time period, around 1832, when it was used by the Duchess of Berry. Solution methods for longer polyphonic ciphers using simulated annealing are discussed briefly in Lasry, Megyesi, and Kopal (2021): They examine papal ciphers from the 16th century, which use digits. Various cipher examples from Meister (1906) are given by Tomokiyo (2019a; 2019b; 2020). Another simple and obvious basis for a ten-digit polyphonic cipher is with a ten-letter keyword, with all letters different. For example, with the keyword "artichokes":

```
0 1 2 3 4 5 6 7 8 9
A R T I C H O K E S
B D F G J L M N P Q
U V W X Y Z
```

5 Conclusion

Emperor Rudolf II was an avid collector of alchemical paraphernalia. The cipher from the hand bell is quite short and yet provides a considerable challenge in terms of diagnosis as there is little context, unlike the contemporaneous papal ciphers. In recent years, the research paradigm called the 'New Historiography of Alchemy' (Principe and Newman, 2001) has promoted the use of so-called 'RRR methods' to replicate historical recipes experimentally.[11] Discoveries such as that by Bean et al. (2022) demonstrate that, in the context of ciphers as well, alchemical secrets are not necessarily 'empty secrets' as had been claimed in the past (Eco, 2016). Yet the riddle of the hand bell cipher shows that we still lack understanding of some of the many different alchemical practices of secrecy (Lang, 2023). For instance, until successful cryptanalysis is achieved, it will be hard to tell if this is 'a real cipher' or some other form of symbolism which was meaningful to its creators, yet whose meaning we do not yet understand.

Further research on the cipher could include calculating the unicity distance of the ciphers discussed here in the suggested three plaintext languages. This would give some idea of whether a solution is even possible at this ciphertext length.

Acknowledgments

We thank Jim Reeds for his helpful comments and contributions.

References

Richard Bean, Sarah Lang, and Megan Piorko. 2022. Solving an alchemical cipher in a shared notebook of John and Arthur Dee. In *Proceedings of the 5th International Conference on Historical Cryptology HistoCrypt 2022*, number 188, pages 12–21. Linköping University Electronic Press.

Beket Bukovinská and Ivo Purš. 2010. Die Tischglocke Rudolfs II: Über ihren Urheber und ihre Bedeutung. *Studia Rudolphina*, 10:89–104.

[10] https://www.cryptogram.org/downloads/aca.info/ciphers/KeyPhrase.pdf.

[11] Reconstruction, replication, re-enactment; see Hendriksen (2020, 314). Corinna Gannon also applied this method and, in collaboration with Christoph Jäggy, tried to reconstruct the sevenfold alloy *Electrum* based on various recipes from Paracelsian sources. Results will be published in her PhD dissertation (submitted November 2022).

Lambros Callimahos. 1977. *Military Cryptanalytics Part III.* NSA, Fort Meade.

Stephen Clucas. 2006. John Dee's angelic conversations and the *Ars notoria*: Renaissance magic and mediaeval theurgy. *John Dee: Interdisciplinary Studies in English Renaissance Thought*, pages 231–273.

Giambattista della Porta. 1563. *De Furtivis Literarum Notis vulgo. De ziferis Libri IIII.* Scotus, Naples.

Umberto Eco, 2016. *Il discorso alchemico e il segreto differito*, pages 97–116. La nave di Teseo.

Helen F Gaines. 1956. *Cryptanalysis: A Study of Ciphers and Their Solution*, volume 97. Courier Corporation.

Corinna Gannon. 2019. The alchemical hand bell of Rudolf II: A touchstone of art and alchemy. In Štěpán Vácha and Sylva Dobalová, editors, *Studia Rudolphina 19. Bulletin of the Research Center for Visual Arts and Culture in the Age of Rudolf II*, pages 81–97, Prag. Artefactum.

Corinna Gannon. 2020. The amulet of Rudolf II – Kabbalistic talisman and pansophic collectible. In Štěpán Vácha, editor, *Studia Rudolphina 20. Bulletin of the Research Center for Visual Arts and Culture in the Age of Rudolf II*, pages 83–101, Prag. Artefactum.

Corinna Gannon. 2023a. Der Klang von sieben Metallen. Die Alchemistische Tischglocke Kaiser Rudolfs II. In Philippe Cordez, Rebecca Müller, and Joanna Olchawa, editors, *Rhythms and Resonances: Sounding Objects in the Middle Ages*, Paris [forthcoming].

Corinna Gannon. 2023b. *Electrum* in the Kunstkammer of Rudolf II. objects made from seven metals. In Sarah Lang, Michael Fröstl, and Patrick Fiska, editors, *Alchemistische Labore. Praktiken, Texte und materielle Hinterlassenschaften / Alchemical Laboratories. Practices, texts, material relics*, pages 87–101, Graz. Grazer Universitätsverlag [forthcoming].

Corinna Gannon. 2024. Die Alchemistische Tischglocke Kaiser Rudolfs II. – Kunstkammerstück und magisches Artefakt. *Wiener Jahrbuch für Kunstgeschichte [forthcoming].*

Deborah E. Harkness. 1999. *John Dee's Conversations with Angels: Cabala, Alchemy, and the End of Nature*. CUP, Cambridge.

Marieke M. A. Hendriksen. 2020. Rethinking performative methods in the history of science. 43:313–322.

Johannes Huser, editor. 1590. *Husersche Quartausgabe. Dieser Theil (welcher der Dritte unter den Philosophischen Schrifften) begreifft fürnemlich das treffliche Werck Theophrasti, Philosophia Sagax, oder Astronomia Magna genannt: Sampt ettlichen andern Opusculis, und einem Appendice*, volume 10.

David Kahn. 1996. *The Codebreakers: The comprehensive history of secret communication from ancient times to the internet*. Simon and Schuster.

Sarah Lang. 2023. Situating ciphers among alchemical techniques of secrecy. In *Proceedings of the 6th International Conference on Historical Cryptology (HistoCrypt 2023)*. Linköping University Electronic Press.

George Lasry, Beáta Megyesi, and Nils Kopal. 2021. Deciphering papal ciphers from the 16th to the 18th century. *Cryptologia*, 45(6):479–540.

Ernst Leierzopf, Nils Kopal, Bernhard Esslinger, Harald Lampesberger, and Eckehard Hermann. 2021. A massive machine-learning approach for classical cipher type detection using feature engineering. In *Proceedings of the 4th International Conference on Historical Cryptology HistoCrypt 2021*, number 183, pages 111–120. Linköping University Electronic Press.

William Mason. 2005. Reference statistics for ACA ciphers. *The Cryptogram*, May-June:4–6.

Aloys Meister. 1906. *Geheimschrift im Dienste der Päpstlichen Kurie von ihren Anfängen bis zum Ende des XVI. Jahrhunderts.* Ferdinand Schöningh, Paderborn.

Megan Piorko, Sarah Lang, and Richard Bean. 2023. Deciphering the *Hermeticae Philosophae Medulla*: Textual cultures of alchemical secrecy. *Ambix*, May.

Lawrence M. Principe and William R. Newman. 2001. Some problems with the historiography of alchemy. In William R. Newman and Anthony Grafton, editors, *Secrets of Nature: Astrology and Alchemy in Early Modern Europe*, pages 385–432, Cambridge/Massachusetts. MIT Press.

Hereward Tilton. 2015. Bells and spells: Rosicrucianism and the invocation of planetary spirits in early modern Germany. *Culture and Cosmos: A Journal of the History of Astrology and Cultural Astronomy*, 19/1:5–26.

Satoshi Tomokiyo. 2019a. Identifying Italian ciphers from continuous-figure ciphertexts (1593). *Cryptologia*, 43(1):23–46.

Satoshi Tomokiyo. 2019b. Polyphonic substitution in Italian Numerical Ciphers.

Satoshi Tomokiyo. 2020. A polyphonic substitution cipher of the Catholic League (1592–1593).

Robert Turner. 1986. *The heptarchia mystica of John Dee*. Aquarian Press.

Ottavian Medici and the decline of Venetian cryptography

Paolo Bonavoglia
Mathesis Venezia c/o Liceo Foscarini Cannaregio 4942 I 30121 Venezia
paolo.bonavoglia@mathesisvenezia.it

Abstract

The recent discovery, in the State Archives of Venice, of a 1621 final account from a committee of three noblemen charged to evaluate cipher services, sheds new light on the decline of cryptography in Venice in the subsequent centuries. The committee produced not only its evaluation, but also a new, interesting cipher, by the young Ottauian Medici. But, after Medici, stagnation and decline set in, until the final collapse of the Republic of Venice in 1797.

1 Introduction

The literature about cryptography in the 17th and 18th centuries in Venice is rather poor; Pasini in his booklet[1] does provide some limited insights, but his main subject is a set of cryptograms from around 1550; Meister in his chapter about Venice,[2] covers a limited period of research up to 1550, because his main goal was to study the various roots of modern cryptography; with only limited time to stay in Venice, he did not go beyond 1550. Other authors such as Preto[3] and Iordanou,[4] also wrote about cryptography in Venice, but their main interest was for its application for espionage and secret services, and less for technical considerations.

Four years of research at the State Archives of Venice, among other things, shed a little more light on this period.

In the 16th century Venice boasted a formidable team of cryptanalysts working under the control of the Council of Ten, henceforth abbreviated to CX:[5] Giacomo Soro, Alvise Borghi, Giambattista Ludovici, and Gianfrancesco Marin, who boasted to be able to break [almost] any cipher; Marin, eventually found himself in tears that he was the last one capable of decrypting foreign ciphers. The CX, also concerned about this situation, requested him to instruct his son, Ferigo in the art of cryptanalysis; and, then, when Gianfrancesco died unexpectedly in 1578, the CX ordered the requisition of all his books in the hope that these would be enough for Ferigo to learn alone the art of *leuar le ziffre senza scontro*.[6]

It was an unfullfilled hopes; before 1578 one finds many references from the CX praising the great cryptanalytic accomplishments of Soro, Borghi, Ludovici and Marin. After 1578, in spite of extensive research, I found no further mention of any successes in decrypting foreign dispatches. Obviously, this lack of evidence is not conclusive, after all, cryptanalysis is perhaps the only science where it is best not to boast or to publicize one's successes—in fact, it is often recommended (and done) to destroy any data about successful decryptions.

In a different consideration, that of designing ciphers, the 16th century in Venice had seen an evolution of ciphers, from those using fancy symbols, letters from exotic alphabets, geometric figures, etc., to ciphers consisting of letters followed by one or two numbers, usually written raised up, as an exponent. These were

[1] (Pasini, 1872 2019)

[2] (Meister, 1902)

[3] (Preto, 1994 1999)

[4] (Iordanou, 2019)

[5] The Council of Ten, often abbreviated to Cons^o of X, or even shorter CX, was a powerful, perhaps the most powerful body of the Venetian Republic, and was also in charge of the secret services, among which was also the cryptographic service, entrusted to the so-called deputies of ciphers

[6] English: deciphering the cryptograms without the cipher sheet

usually nomenclators, consisting of an alphabet almost always with homophones, a certain number of nulls, and a dictionary of words encrypted with a single sign; increasingly the use of syllabaries (groups of letters formed by one or more consonants followed by a vowel), became widespread. Nothing out of the ordinary, the nomenclator was the most widely-used cipher in Europe for professional users, namely by the military, and especially for diplomatic purposes.

During the same century several polyalphabetic ciphers were invented by ingenious amateurs,[7] in particular the polyalphabetic ciphers of: Leon Battista Alberti,[8] one of the foremost architects of the Italian Renaissance; Johannes Trithemius,[9] an abbot; Giovan Battista Porta,[10] a playwright; and Blaise de Vigenère,[11] a diplomat; not to mention, Giambattista Bellaso, a secretary to various cardinals, who was in charge of their ciphers and published some very ingenious ciphers.

As early as the second half of the 16th century, a trend had begun to make ciphering and deciphering nomenclators easier to use and faster, which raised concerns about the safety of nomenclators, primarily by the two most brilliant designers of ciphers in the final decades of the 16th century, Hieronimo di Franceschi,[12] secretary of the Senate and deputy of ciphers for the Council of Ten from 1576 to 1600, and Pietro Partenio[13] a private notary from 1563 to 1610, with a great skill for ciphers. They all shared a common concern: other rulers would have their own skilled cryptanalysts, therefore the Venetian ciphers were no longer as secure as believed; the proposed remedy, however, was radically different, as we will see.

Franceschi considered nomenclators unreliable and proposed instead adopting the *uere ziffre* true ciphers, as he called the polyalphabetic ciphers—such was the *cifra delle caselle*, a polyalphabetic cipher based on arithmetic, subtraction and addition.[14] While Partenio despised letter-by-letter ciphering, as in mono- and polyalphabetic ciphers, and aimed rather at strengthening the nomenclators by increasing the size of the dictionary and syllabary, but above all by super-encrypting the nomenclator to keep it safe, even in case of theft of the cipher sheet. Such was his *cifra n. 5*, the only one that was used in 1595 by the Paris embassy, albeit for a very short time.

The polyalphabetic ciphers were presented as ciphers that were absolutely indecipherable unless one knew the key, for individual letters were not always encrypted with the same sign or at most with a choice of equivalent signs as in nomenclators, but the same letter could be encrypted with different letters, making frequency analysis, a statistical tool useful for of forcing monoalphabetic ciphers, useless. Indeed this is true only with a random and non-reusable key.

Nevertheless, the nomenclator continued to reign supreme for centuries, as David Kahn proposes in Codebreakers,[15] wondering why the cipher offices were so hostile to the polyalphabetic ciphers.[16] Franceschi's cipher was, to my knowledge, the only case of a polyalphabetic cipher that was actually used in the real world for diplomatic messages.

2 1600, year of the turning point

In 1596 the CX had resolved to elect a committee of five nobles to find a solution to the dispute between Franceschi and Partenio, specifically between the previously mentioned, Franceschi's *cifra delle caselle*, and Partenio's *cifra 5.*

Around 1599 or 1600 the dispute came to an inglorious end, with a final contest between the two ciphers and their inventors their

[7] Naturally I mean *amateurs* in the cryptographic field, people whose main profession was not in the cryptographic field

[8] (Alberti, 1511)

[9] His best known work is (Trithemius, 1507 1613).

[10] Author of a huge review of ciphers in (Porta, 1606)

[11] (de Vigenère, 1587)

[12] 1540 - 1600 . The name was spelled Hieronimo up to the end of the century, thereafter Girolamo.

[13] 1538 - 1620; the surname is also spelled Parthenio, and in Latin, Parthenius

[14] (Bonavoglia, 2019)

[15] (Kahn, 1967 1996)

[16] Agostino Amadi wrote in his treatise of ciphers, about the polyalphabetic ciphers of Bellaso: *[...] le quali tutte sono nobilissime inuentioni, ma non da da Principi che uogliono il sodo, il uero et saper ancora loro che quella zifra [sie li?] per quella forma and not another one.* English: [...] all of them are very noble inventions, but not [used] by Princes (rulers) that want the practical, the true, and to know that one cipher has that meaning (word or letter) and not another one

adversaries, along with their respective assistants; it was held despite the decision by Pietro Amai, Franceschi's assistant, to excuse himself, because he did not feel skilled enough in adding and subtracting, a basic skill with the *caselle*! The five noblemen wrote a draft report[17] that ended with a verdict of parity between the two ciphers, with both rated as very strong, although with a slight preference for Franceschi's *caselle*; in the end they recommended using both ciphers, by alternating the two.

As it turns out, the final report does not appear to have ever been delivered to the CX, since only the draft could be found, more than likely because Franceschi had died in the first half of 1600. Therefore the above recommendations were completely ignored

3 Pietro Amai takes over from Franceschi

When Franceschi died, the role of chief deputy for ciphers passed into the hands of Pietro Amai, Franceschi's main collaborator, who was for some time joined by Ferigo Marin son of Zuan Francesco.

Although the son of the more-celebrated Agostino Amadi, author of the latest treatise on ciphers produced by Venetian cryptography, Pietro, like Ferigo Marin, comes across as a rather weak and lazy character, as was already evident from his inglorious withdrawal from the final round of the Franceschi—Partenio dispute, cited above.

Amai was careful not to reintroduce the *caselle*, much less Partenio's superencrypted systems. The current cipher, since 1599, has been the A = Z10,[18] a simplified nomenclator without homophones and nulls and with a total of 300 cipher signs. This is almost certainly the cipher that Partenio claimed to have easily decrypted in his 1606 letter, in which he denounced the weakness of the ciphers being used during that period.

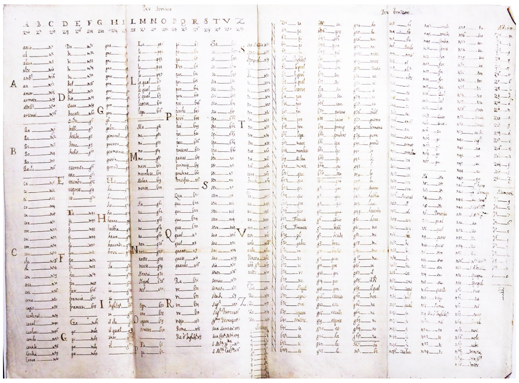


Figure 1: The Z10 cipher. Original cipher sheet. *ASVe Cifre, chiavi e scontri di cifra ...busta 1, f. 3*

Indeed, the cipher has several weaknesses: 1) there are no homophones and nulls, but only one cipher sign for each letter; 2) there is a large syllabary, and although that should have been a strength, instead, there is a weakness, clearly visible in figure 2, the syllable ciphers are ordered by vowels: A = 1, E = 2, I = 3, O = 4, U = 5, following a medieval scheme, e.g.: the 1226 Liber Plegiorum.[19]

letter	a	e	i	o	u
cipher	1	2	3	4	5

Obviously, the cryptanalyst trying to crack this cipher, upon realizing that the syllables are ordered, would be greatly aided in reconstructing the syllabary, which is the backbone of the cipher, and would have had little difficulty finding the solution.

The CX was well aware of this situation and sent reprimands to the cipher deputies, complaining about the serious disorder in the ciphers office and the fact that for years and years the current cipher, the Z10, was never changed.

4 A committee of three noblemen is elected

Finally, in 1619, the CX approved the election of a committee of three noblemen charged with reforming the ciphers and to find a new cipher to replace the old Z10, which after twenty

[17] This long sought report, was found in 2022 as two almost identical minutes in poor condition due to oozing inks in an envelope in the State Archives of Venice, henceforth abbreviated to *ASVe*. *ASVe Cifre, chiavi e scontri di cifra ...busta 6.*

[18] The classification of ciphers using the encryption of the letter **A** is due to Luigi Pasini (1835 1885) an archivist of the Archives of Venice, who reordered the cryptographic papers and wrote a booklet about the ciphers of the Republic of Venice, focusing on some encrypted letters around 1550 (Pasini, 1872 2019).

[19] A medieval register of chancery records digitized in *ASVe Collegio Minor Consiglio Liber Plegiorum, Reg-12231229, c. 48r, 84-v, 117r*

ba m^{1}	be m^{2}	bi m^{3}	bo m^{4}	bu m^{5}	ca m^{11}	ce m^{12}	ci m^{13}	co m^{14}	cu m^{15}	cra m^{21}	cre m^{22}	cri m^{23}	cro m^{24}	cru m^{25}	da m^{31}	de m^{32}	di m^{33}	do m^{34}	du m^{35}
fa m^{41}	fe m^{42}	fi m^{43}	fo m^{44}	fu m^{45}	fra m^{51}	fre m^{52}	fri m^{53}	fro m^{54}	fru m^{55}	ga m^{61}	ge m^{62}	gi m^{63}	go m^{64}	gu m^{65}	gna g^{1}	gne g^{2}	gni g^{3}	gno g^{4}	gnu g^{5}
gra g^{11}	gre g^{12}	gri g^{13}	gro g^{14}	gru g^{15}	ha g^{21}	he g^{22}	hi g^{23}	ho g^{24}	hu g^{25}	ia g^{31}	ie g^{32}	ii g^{33}	io g^{34}	iu g^{35}	la g^{41}	le g^{42}	li g^{43}	lo g^{44}	lu g^{45}
ma g^{51}	me g^{52}	mi g^{53}	mo g^{54}	mu g^{55}	na g^{61}	ne g^{62}	ni g^{63}	no g^{64}	nu g^{65}	pa b^{1}	pe b^{2}	pi b^{3}	po b^{4}	pu b^{5}	pra b^{11}	pre b^{12}	pri b^{13}	pro b^{14}	pru b^{15}
qua b^{21}	que b^{22}	qui b^{23}	quo b^{24}	quu b^{25}	ra b^{31}	re b^{32}	ri b^{33}	ro b^{34}	ru b^{35}	sa b^{41}	se b^{42}	si b^{43}	so b^{44}	su b^{45}	sca b^{51}	sce b^{52}	sci b^{53}	sco b^{54}	scu b^{55}
spa b^{61}	spe b^{62}	spi b^{63}	spo b^{64}	spu b^{65}	sta n^{1}	ste n^{2}	sti n^{3}	sto n^{4}	stu n^{5}	stra n^{11}	stre n^{12}	stri n^{13}	stro n^{14}	stru n^{15}	ta n^{21}	te n^{22}	ti n^{23}	to n^{24}	tu n^{25}
tra n^{31}	tre n^{32}	tri n^{33}	tro n^{34}	tru n^{35}	ua n^{41}	ue n^{42}	ui n^{43}	uo n^{44}	uu n^{45}	za n^{51}	ze n^{52}	zi n^{53}	zo n^{54}	zu n^{55}					

Figure 2: The Z10 cipher syllabary.

years had passed through too many hands to be considered safe.

Procurator Girolamo Giustinian, and noblemen Francesco Morosini and Ottaviano Valiero were elected. In the meantime, two young men Ottavian[20] Medici and Giambattista Lionello had passed the exams required to become deputies of ciphers, and therefore were able to collaborate with the newly-elected committee.

One of the committee's first acts was to consult the octogenarian Pietro Partenio, although these were no longer the years when the CX enthusiastically praised his ciphers. Here is an excerpt from the committee's final report of 1621:[21]

> Several times we had meetings with a diligent examination of a great variety of ciphers, by the *ziffristi* secretaries and by the late Pietro Parthenio very skilled in that profession; of this person we could tell Your Serenity, with our usual sincerity, that we saw, while he was in life, very witty inventions, of equal safety, and worthy of commendation, but balanced these requirements with some difficulty in the use and slowness in deciphering and enciphering, when a multiplicity of tasks are presented almost instantly every day on all sides, we have judged, for these causes alone, that we cannot recommend their use.

Ultimately an elegant way to dismiss Partenio, and generally an epitaph for overly-complicated and slow ciphers; indeed, finding the balance between security and speed of use is an age-old problem in cryptography.

As the text suggests, Partenio had since died in 1620, according to Tassini,[22] and there are no wives or children reported. He therefore had no direct heirs, but in his letter of January 1606 (1605 m.v.)[23] he designated one with these words:

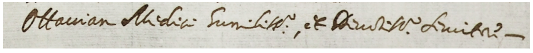
> [...] so that we may instruct Ottauiano Medici, extraordinary of the Cancellaria, which is to me like a son, as everyone knows, of excellent hope.

Partenio was right, as we shall see, the already-mentioned Medici would prove to be the dominant figure in Venetian cryptography in the first half of the 17th century; and in any case the last cipher deputy of any depth in the history of the Republic of Venice.

5 The cipher of the three noblemen, by Medici and Lionello

It took the committee more than two years to arrive at a consensual proposal, a report which, after rejecting Partenio's ciphers as too difficult and slow, proposed the adoption of a new cipher that assimilated some important innovations from the past.

The report attributes the design of this cipher to the deputies of ciphers in service at that time. Attached to the report are the enciphered and deciphered text of several pages, such as were typically used by ambassadors, and, in this case, used as an exam; on April

[20] The name is variously spelled as *Ottauio, Ottauian, Ottauiano.* I prefer the form *Ottauian* used in his signature, slike this one:

Ottauian Medici Humil.mo, et Deuot.mo Seruitor

[21] 17th century Italian original text: *Più uolte siamo stati insieme con essaminatione dili[gentissi]ma sopra una gran uarietà de scontri, che ci sono stati presentati et dalli secretari ziffristi e dal già Pietro Parthenio peritissimo in tal professione; di questo soggetto potemo con la [nostra?] solita sincerità a dire a Vostra Serenità di hauer ueduto, mentre egli uiueua inuentioni molto spiritose, di pari sicurtà, et degne di comendatione ma bilanciati questi requisiti con qualche difficoltà nel uso et tardità nel trazer et scriuer, quando alla giornata occorre che che quasi a tempi presenti risorge la multiplicità da ogni parte habbiamo giudicato per queste sole cause di non poter determinare la loro essercitatione*

[22] citeTasCit88

[23] m.v. stays for *more veneto.* the Venetian style of the calendar: the first day of the year was March 1, following the ancient Roman Republic style; that's why September begins with *septem* Latin for seven, October with *octo* = eight ...so January and February 1606, are still 1605 m.v.

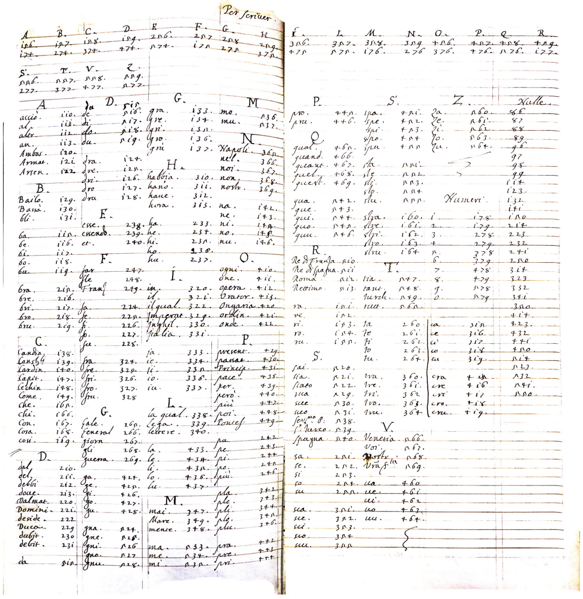


Figure 3: The 1621 cipher, original cipher sheet.*ASVe Cifre, chiavi e scontri di cifra ...busta 2, f. 15*

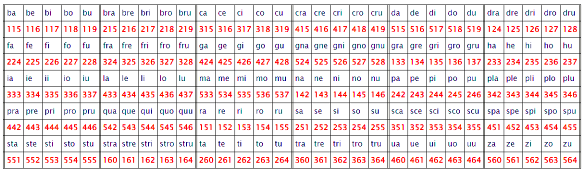
28, 1622 it took Medici four hours to encipher the message; two days later it took Lionello three hours to decipher it—a confirmation that Medici and Lionello[24] were now the two main cipher deputies.

The cipher presents some interesting changes from the Z10 cipher and from those of the last half century.

- The cipher is now, and henceforth, formed of numbers only, each letter or group of letters being encrypted with a number of three digits. The numbers are to be written continuously without separator spaces so that one cannot tell where the single cipher begins and ends, nor how many digits they consist of, two? three? four?

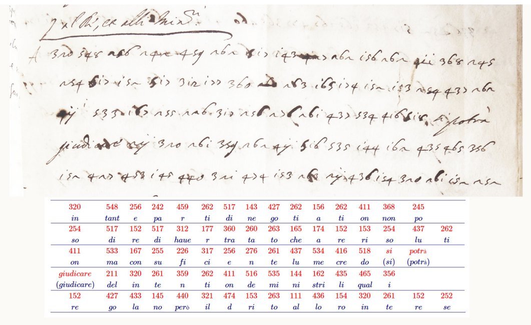
- Homophones reappear, each letter has two cipher signs; put another way, it is a cipher of the double alphabet.

- There is a large number of nulls; secretaries are advised to insert many nulls here and there between the actual ciphers; particularly at the beginning and end of the line, between the double ...



ba	be	bi	bo	bu	bra	bre	bri	bro	bru	ca	ce	ci	co	cu	cra	cre	cri	cro	cru	da	de	di	do	du	dra	dre	dri	dro	dru
115	116	117	118	119	215	216	217	218	219	315	316	317	318	319	415	416	417	418	419	515	516	517	518	519	124	125	126	127	128
fa	fe	fi	fo	fu	fra	fre	fri	fro	fru	ga	ge	gi	go	gu	gna	gne	gni	gno	gnu	gra	gre	gri	gro	gru	ha	he	hi	ho	hu
224	225	226	227	228	324	325	326	327	328	424	425	426	427	428	524	525	526	527	528	133	134	135	136	137	233	234	235	236	237
ia	ie	ii	io	iu	la	le	li	lo	lu	ma	me	mi	mo	mu	na	ne	ni	no	nu	pa	pe	pi	po	pu	pla	ple	pli	plo	plu
333	334	335	336	337	433	434	435	436	437	533	534	535	536	537	142	143	144	145	146	242	243	244	245	246	342	343	344	345	346
pra	pre	pri	pro	pru	qua	que	qui	quo	quu	ra	re	ri	ro	ru	sa	se	si	so	su	sca	sce	sci	sco	scu	spa	spe	spi	spo	spu
442	443	444	445	446	542	543	544	545	546	151	152	153	154	155	251	252	253	254	255	351	352	353	354	355	451	452	453	454	455
sta	ste	sti	sto	stu	stra	stre	stri	stro	stru	ta	te	ti	to	tu	tra	tre	tri	tro	tru	ua	ue	ui	uo	uu	za	ze	zi	zo	zu
551	552	553	554	555	160	161	162	163	164	260	261	262	263	264	360	361	362	363	364	460	461	462	463	464	560	561	562	563	564

Figure 4: The syllabary of the 1621 cipher.



320	548	256	242	459	262	517	143	427	262	156	262	411	368	245	
in	*tant*	*e*	*pa*	*r*	*ti*	*di*	*ne*	*go*	*ti*	*a*	*ti*	*on*	*non*	*po*	
254	517	152	517	312	177	360	260	263	165	174	152	153	254	437	262
so	*di*	*re*	*di*	*haue*	*r*	*tra*	*ta*	*to*	*che*	*a*	*re*	*ri*	*so*	*lu*	*ti*
411	533	167	255	226	317	256	276	261	437	534	416	518	*si*	*potrà*	
on	*ma*	*con*	*su*	*fi*	*ci*	*e*	*n*	*te*	*lu*	*me*	*cre*	*do*	*(si)*	*(potrà)*	
giudicare	211	320	261	359	262	411	516	535	144	162	435	465	356		
(giudicare)	*del*	*in*	*te*	*n*	*ti*	*on*	*de*	*mi*	*ni*	*stri*	*li*	*qual*	*i*		
152	427	433	145	440	321	474	153	263	111	436	154	320	261	152	252
re	*go*	*la*	*no*	*però*	*il*	*d*	*ri*	*to*	*al*	*lo*	*ro*	*in*	*te*	*re*	*se*

Figure 5: 1st dispatch using the new cipher, Paris July 3, 1623. *ASVe, Senate, dispacci degli ambasciatori, Francia, f.59, c.550*

- The syllabaries are partially ordered, in the sense that they are ordered by vowel as in Z10 starting from a to u, but without necessarily starting from 1, for instance.

The most significant feature is the large number of nulls, which are essential if the length of single ciphers has to be kept hidden.

The cipher is supposed to have come into use in 1622, but the earliest dispatch I have found that uses this cipher is a 23-page document dated July 3, 1623 from the Venetian ambassador to Paris, Giovanni Pesaro, encrypted only in part. In figure 5 two things are noticeable: 1) there is not even a single null; and 2) the spaces between the cipher marks are clearly visible. The previous recommendations were totally disregarded; the first dispatch using this cipher, from the ambassador to London, Alvise Vallaresso, dated August 23, 1623, also has the same problems, no nulls, spaces mostly visible, although the effort to write continuously is somewhat discernible. There is a strange difference from the cipher used in Paris; the syllables *da de di do du* are encrypted with the numbers *215 216 217 218 219*, which in Paris and in the version preserved in the Venetian archives stood instead for: bra bre bri bro bru. I did not find a sat-

[24] For some reason Lionello disappears thereafter, a plausible conjecture is that he was among the victims of the 1630 devastating plague.

isfactory explanation for this discrepancy.

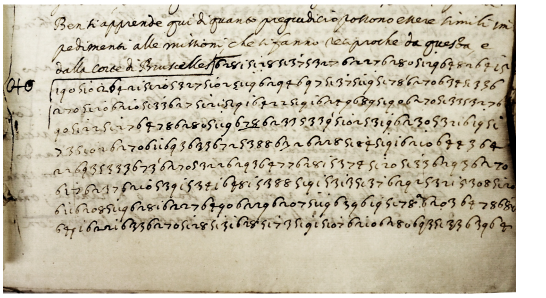
Without nulls and with spaces left visible the cipher looks no safer than Z10, there remains only the fact of having a somewhat less-orderly syllabary. In later years things improved with regard to continuous writing, which ultimately became a habit for all secretaries, who, conversely, never became accustomed to using nulls.

6 The 1624 variable-size cipher

Then, in 1624, as a result of the embassy secretary in London being robbed of many documents, including cipher sheets, the CX ordered the cipher to be changed, and therefore Medici designed a new, and a very interesting one, at that.

It is a variable-size cipher: some letters or syllables or words were encrypted with numbers of two-decimal digits, others with three, and others with four. The idea was not new; it had been proposed in his treatise by Matteo Argenti,[25] secretary to the Papal ciphers in the late 16th century. Argenti used numbers of one or two digits, relying on the acumen of the cipher secretaries for a correct deciphering; for usually, the question as to whether the next number is of two or three digits is resolved by context—usually, but not always—only one combination will produce sensible texts.

We do not know if Medici knew of Argenti's treatise—at least I have found no trace of it in the archives; he clearly prefers another solution that leaves no doubt; numerals 5 and 6 are used exclusively as the first number of a group. An original idea, yes, but one has to wonder to what extent one can fool the enemy with this trick; those 5s and 6s are already at a glance distributed in a somewhat too-regular manner that might arouse suspicion in the eye of the enemy. Not to mention that many secretaries did not understand the instructions well and kept leaving a space between one cipher group and the next, as can be seen in figure 6.



Ben si apprende qui di quanto pregiuditio possono essere simili impedimenti alle missioni, che si fanno reciproche da questa e dalla corte di Brusselles [illegible]

Figure 6: Good use of the Medici cipher. *ASVe, Senato, Dispacci degli ambasciatori in Francia, filza 74, no. 216, 3 ott 1630.*

7 Cipher A 105-115 the return of the fixed-size cipher, three digits

The 1624 cipher also had to be abandoned due to theft; Medici designed another one in collaboration with the now-elderly Pietro Amai and Antonio Marin, which was approved by the CX on March 23, 1630; there is a return to fixed-size cipher signs of three digits. There is a dual alphabet and that is two homophones per letter beginning with the A encrypted with 105 and 115. There is a syllabary sorted according to vowel by a criterion very similar to that of the deprecated Z10 cipher, only beginning with zero instead of 1: for example ba be bi bo bu are encrypted with 100 101 102 103 104, another step in the direction of simplification, at the expense of security. There is a large number of nulls; in short, it is a remake of the three-nobles cipher with different numbers. The cipher A 105–115 was used for many years even after the adoption of a new one in 1645.

8 1647, 28 February an encrypted message by the Capitano Generale da Mar

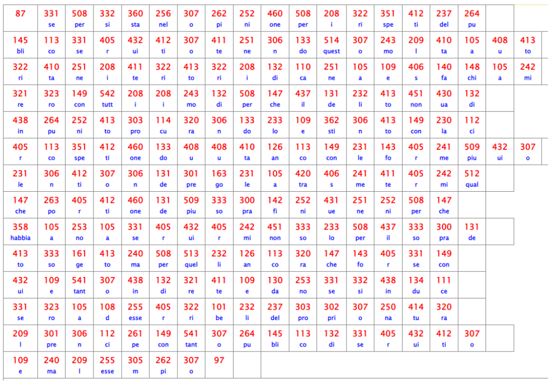
Here is an interesting example of use of the previous Medici cipher: a message encrypted only in its most delicate matter (see figure 7). It is an example of use by an admiral, the Capitano da Mar,[26] Giambattista[27] Grimani, from a galley (galera) in Porto di Scandia,[28] dur-

[25] (Argenti, 1906) p. 152, inside (Meister, 1906)

[26] *Capitano Generale da Mar* was the title of the commander in chief of the Venetian fleet. Grimani remained in office from 1646 to 1648, when he drowned with his ship in a violent sea storm while attempting to establish a Dardanelles blockade.

[27] At first reading I had interpreted the name in the signature, difficult-to-read handwriting, as *Ernesto*, but recently it clearly turned out to be *Giambattista*, in agreement with the very little historical information found on this character.

[28] It is the ancient name of a port of the island of Kythira between the Peloponnese peninsula and the



87	331 se	508 per	332 si	360 sta	256 nel	307 o	262 pi	252 ni	460 one	508 per	208 i	322 ri	351 spe	412 ti	237 del	264 pu		
145 bli	113 co	331 se	405 r	432 ui	412 ti	307 o	411 te	251 ne	306 n	133 do	514 quest	307 o	243 mo	209 l	410 ta	105 a	408 u	413 to
322 ri	410 ta	251 ne	208 i	411 te	322 ri	413 to	322 ri	208 i	132 di	110 ca	251 ne	105 a	109 e	406 s	140 fa	148 chi	105 a	242 mi
321 re	323 ro	149 con	542 tutt	208 i	208 i	243 mo	132 di	508 per	147 che	437 il	131 de	232 li	413 to	451 non	430 ua	132 di		
438 in	264 pu	252 ni	413 to	303 pro	114 cu	320 ra	306 n	133 do	233 lo	109 e	362 sti	306 n	413 to	149 con	230 la	112 ci		
405 r	113 co	351 spe	412 ti	460 one	133 do	408 u	408 u	410 ta	126 an	113 co	149 con	231 le	143 fo	405 r	241 me	509 piu	432 ui	307 o
231 le	306 n	412 ti	307 o	306 n	131 de	301 pre	163 go	231 le	105 a	420 tra	406 s	241 me	411 te	405 r	242 mi	512 qual		
147 che	263 po	405 r	412 ti	460 one	131 de	509 piu	333 so	300 pra	142 fi	252 ni	431 ue	251 ne	252 ni	508 per	147 che			
358 habbia	105 a	253 no	105 a	331 se	405 r	432 ui	405 r	242 mi	451 non	333 so	233 lo	508 per	437 il	333 so	300 pra	131 de		
413 to	333 so	161 ge	413 to	240 ma	508 per	513 quel	232 li	126 an	113 co	320 ra	147 che	143 fo	405 r	331 se	149 con			
432 ui	109 e	541 tant	307 o	438 in	132 di	321 re	411 te	109 e	130 da	253 no	331 se	332 si	438 in	134 du	111 ce			
331 se	323 ro	105 a	108 d	255 esse	405 r	322 ri	101 be	232 li	237 del	303 pro	302 pri	307 o	250 na	414 tu	320 ra			
209 l	301 pre	306 n	112 ci	261 pe	149 con	541 tant	307 o	264 pu	145 bli	113 co	132 di	331 se	405 r	432 ui	412 ti	307 o		
109 e	240 ma	209 l	255 esse	305 m	262 pi	307 o	97											

Figure 7: Grimani's cryptogram, decrypted.

ing the Siege of Candia, the long war with the Ottoman Empire over the possession of that island.

The first part of the message is in plain text and recounts that Michiel Caliergi, commander of the Canea,[29] while Grimani was visiting the nearby islands, had made himself all too familiar with the Vizier who treated him very well as a confidant. Grimani argues this behavior is treasonous (clearly, consorting with the enemy) and that it was essential to eliminate him...but in a discreet manner; he is more to-the-point in the encrypted part, presented here deciphered in English[30]

> If you confirm the opinion, for the respect of the public service, of holding so much authority in the territories of Canea and Sfachia, I will, by all means, manage it so that the offense does not go unpunished, procuring him extinguished, so with due circumspection, send me some portions of the most superfine poisons, so that I can use them not only for this subject, but for anyone in the future who may, in such an indirect and harmful way, be induced to be a rebel of his own natural prince, with such public bad service and bad example.

The State Inquisitors responded in April approving Grimani's request and loyalty; they enclosed a paper recommending three poisons: *scamonea*, poisonous if administered continuously; *cantarella*, which blocks urination; and the well-known *arsenic*. But they added that they could not procure and send them because they would have to confide the matter to many people, risking raising questions, objections and ill feelings, and Grimani certainly knew the right people to procure the poisons in Candia.

We do not know if Grimani got the poisons in Candia and if Caliergi was actually poisoned to disguise his death as natural. But, anyway, this provides a good example of when to encrypt a message.

From the cryptographic point of view, Grimani (or his secretary) does not deserve much praise; the A 105-115 cipher has a double alphabet, but here only 105 is used for A, 115 not a single time. The same for E 109 119, and other letters. In other words homophones are simply disabled. The dictionary as evident from figure 7 was never used, and so for the nulls. The rule of writing in a continuous way is well executed, but ends of lines are respected, so it is not difficult, having observed that every line has a number of digits in a multiple of 3, that the single ciphers have 3 digits. A confirmation that the military officer was less skilled than the diplomat, when writing in cipher. The reason is obviously the availability of time: the military can not spend much time encrypting messages, while the diplomat can work at a calmer pace.

9 1645 The scontro novissimo

On March 22, 1645 a new cipher with three-digit signs, was approved by the Council of Ten, under the name of scontro nouissimo (newest cipher). It was signed by Ottavian Medici and Marc'Antonio Padauin, the last to be signed by Medici. The alphabet is triplicate, so letter A is encrypted with three homophones 100, 300, 504. Despite its name, it has nothing really new.

Crete island.

[29] A port in the eastern part of the Crete island.

[30] Original 16th century Italian: *Se persista nel opinione per i rispeti del publico seruitio tenendo questo molta autorita nei teritorii di Canea e Sfachia mirerò con tutti i modi perché il delito non uadi inpunito procurandolo estinto con la circospetione douuta anco con le forme piu uiolente onde pregole a trasmetermi qualche portione de piu soprafini ueneni perche habbiaano a seruirmi non solo per il sopradeto sogeto ma per quelli ancora che forse con uie tanto indirete e danose si inducesero ad esser ribeli del proprio natural prencipe con tanto publico diseruitio e mal essempio.*

Medici retired about 1650, and in 1653 was made a nobleman, in recognition of his longtime service, and for a few years Marcantonio Padauin was at the helm of Venetian cryptography; when he died in 1653 two young men Lunardo Formenti and Ottavian Valier acquired the roles of deputy of ciphers. But one has to wait 22 years before seeing a new cipher, in 1675.

10 1675 The ghost cipher of Lunardo Formenti

In November 1674 the CX noted that 25 years had elapsed since the last change of the current cipher; they demanded that the State Inquisitors make contact with the deputies for ciphers to design a new one. Lunardo Formenti, Medici's successor, presented a new one on April 4, 1675, with an attached description in which we read:

> in several ways varied to the sign that to form a word, as for example *Bailo* may be explicated in the following four forms, that each of them in several ways referred to that same word of *Bailo*:[31]

The forms are:

`700513966 600601501802902703500 866514607839 808908607507706`

Is it possible to recover the cipher sheet? Theoretically it is impossible, if you allow for all possible enciphering of single letters, syllables, etc. But knowing the model used during those years, one can assume, with great certainty, three-digit ciphers:

700	513	966
	BAILO	

600	601	501	802	902	703	500
	B	A	I	L	O	

866	514	607	839
	BA	I	LO

808	908	607	507	706
B	A	I	L	O

And then, assuming that 500 600 700 866 966 are nulls and rather ordered alphabet and syllabaries, I found this possible, and plausible, conjecture for the alphabet:

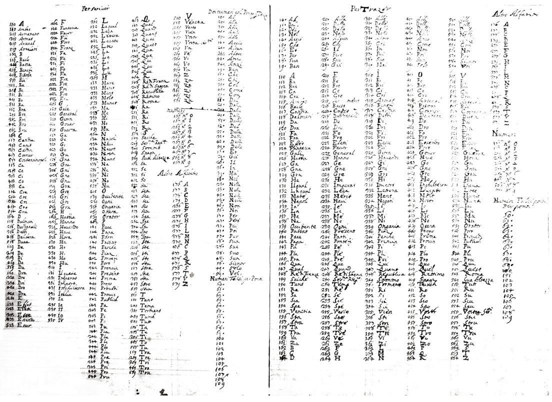


Figure 8: The 1691 cipher by Vettor Pozzo. *ASVe Cifre, chiavi e scontri di cifra ...busta 1 f.7*

A	B	C	D	E	F	G	H	I	L
501	601	701	801	901	502	602	702	802	902
908	808	708	608	508	907	807	707	697	597

M	N	O	P	Q	R	S	T	V	Z
503	603	703	803	903	504	604	704	804	904
906	806	706	606	506	905	805	705	605	505

a single cipher 513 for the word Bailo and these possible syllables.

BA	BE	BI	BO	BU	LA	LE	LI	LO	LU
514	614	714	814	914	539	639	739	839	939

But, surprisingly, I have not found a single diplomatic or military letter encrypted this way, and no cipher sheet in the huge collection of ciphers kept in the archives. And another surprise is that most dispatches by the ambassadors were encrypted using the 1621 cipher of the three noblemen!

Thus, the strange case of a ghost cipher, a new cipher rejected, and a 60 year-old cipher recycled, mark the beginning of the definitive decline of Venetian cryptography.

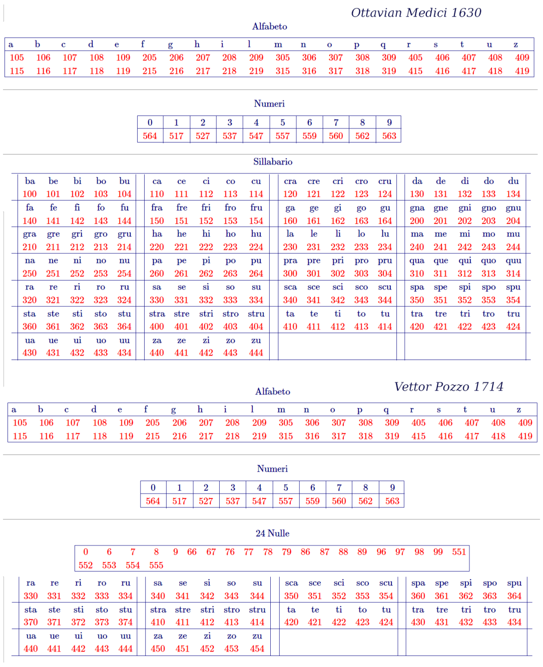
11 1691 Cipher No. 11 Vettor Pozzo

On January 16, 1691 (1690 m.v.[32] the CX adopted a new cipher by Vettor Pozzo and Constantin Nicolosi, who were the main deputies for ciphers.

The cipher is very similar to the previous one in use since 1621; only an odd variation was introduced, the use of a dot as the eleventh cipher sign after the ten digits. The dot is used only as the third sign, like **10**. for *all*,

[31] 16th century Italian: *in più modi uariate à segno che à formar una parola, come per esempio Bailo si può esplicarle nelle seguenti quattro forme, che ogn'una di esse in più modi riferisse quella stessa parola di Bailo, ciò è:*

[32] See note 4



Ottavian Medici 1630

Alfabeto

a	b	c	d	e	f	g	h	i	l	m	n	o	p	q	r	s	t	u	z
105	106	107	108	109	205	206	207	208	209	305	306	307	308	309	405	406	407	408	409
115	116	117	118	119	215	216	217	218	219	315	316	317	318	319	415	416	417	418	419

Numeri

0	1	2	3	4	5	6	7	8	9
564	517	527	537	547	557	559	560	562	563

Sillabario

ba	be	bi	bo	bu	ca	ce	ci	co	cu	cra	cre	cri	cro	cru	da	de	di	do	du
100	101	102	103	104	110	111	112	113	114	120	121	122	123	124	130	131	132	133	134
fa	fe	fi	fo	fu	fra	fre	fri	fro	fru	ga	ge	gi	go	gu	gna	gne	gni	gno	gnu
140	141	142	143	144	150	151	152	153	154	160	161	162	163	164	200	201	202	203	204
gra	gre	gri	gro	gru	ha	he	hi	ho	hu	la	le	li	lo	lu	ma	me	mi	mo	mu
210	211	212	213	214	220	221	222	223	224	230	231	232	233	234	240	241	242	243	244
na	ne	ni	no	nu	pa	pe	pi	po	pu	pra	pre	pri	pro	pru	qua	que	qui	quo	quu
250	251	252	253	254	260	261	262	263	264	300	301	302	303	304	310	311	312	313	314
ra	re	ri	ro	ru	sa	se	si	so	su	sca	sce	sci	sco	scu	spa	spe	spi	spo	spu
320	321	322	323	324	330	331	332	333	334	340	341	342	343	344	350	351	352	353	354
sta	ste	sti	sto	stu	stra	stre	stri	stro	stru	ta	te	ti	to	tu	tra	tre	tri	tro	tru
360	361	362	363	364	400	401	402	403	404	410	411	412	413	414	420	421	422	423	424
ua	ue	ui	uo	uu	za	ze	zi	zo	zu										
430	431	432	433	434	440	441	442	443	444										

Vettor Pozzo 1714

Alfabeto

a	b	c	d	e	f	g	h	i	l	m	n	o	p	q	r	s	t	u	z
105	106	107	108	109	205	206	207	208	209	305	306	307	308	309	405	406	407	408	409
115	116	117	118	119	215	216	217	218	219	315	316	317	318	319	415	416	417	418	419

Numeri

0	1	2	3	4	5	6	7	8	9
564	517	527	537	547	557	559	560	562	563

24 Nulle

0	6	7	8	9	66	67	76	77	78	79	86	87	88	89	96	97	98	99	551
552	553	554	555																

ra	re	ri	ro	ru	sa	se	si	so	su	sca	sce	sci	sco	scu	spa	spe	spi	spo	spu
330	331	332	333	334	340	341	342	343	344	350	351	352	353	354	360	361	362	363	364
sta	ste	sti	sto	stu	stra	stre	stri	stro	stru	ta	te	ti	to	tu	tra	tre	tri	tro	tru
370	371	372	373	374	410	411	412	413	414	420	421	422	423	424	430	431	432	433	434
ua	ue	ui	uo	uu	za	ze	zi	zo	zu										
440	441	442	443	444	450	451	452	453	454										

Figure 9: Part of the cipher sheet of the 1714 cipher, and of the 1630 cipher.

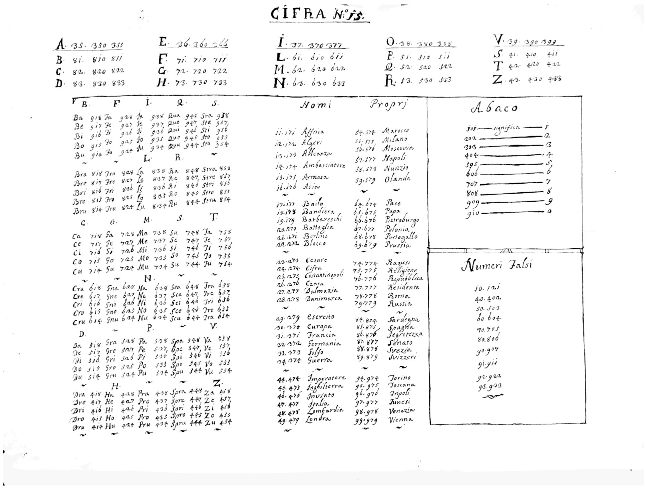
20. for *alla*. It is hard to imagine how this dot could have improved the security of the cipher; maybe one hoped to confuse the enemy? Ironically, this may have actually helped the codebreakers!

12 1714 Cipher No. 12 Vettor Pozzo

If the 1691 Pozzo cipher was very very similar to the 1622 cipher, 23 years later he did even better: cipher no. 12, approved on March 7, 1714, had been presented by the same Vettor Pozzo, purportedly as a new cipher. But a close scrutiny of the cipher unveils a simple clone of Medici's 1630 cipher, simply modified by adding 10 to every cipher, see figure 9.

13 1733-1787 Last ciphers of the Republic

A new cipher was approved as the replacement cipher in 1733, no. 13; the *primo cifrista* then was Agostino Bianchi, who designed not only no. 13, but also no. 14, to be kept by the State Inquisitors as a reserve in case something could happen to compromise no. 13. In fact, such an incident did occur 49 years later! In the meantime Agostino Bianchi had died leaving the task of ciphers to his sons Francesco and Maffio, and then to Marcantonio Buseniello. No. 14 was, already in no. 13, a cipher simplified as much as possible, reduced to a double alphabet and a syllabary, all according to very regular, and therefore, cryptographically-weak patterns.



CIFRA N.º 15.

A. 35. 330 355	E. 36. 360 366	I. 37. 370 377	O. 38. 380 388	V. 39. 390 399
B. 81. 810 811	F. 71. 710 711	L. 61. 610 611	P. 51. 510 511	S. 41. 410 411
C. 82. 820 822	G. 72. 720 722	M. 62. 620 622	Q. 52. 520 522	T. 42. 420 422
D. 83. 830 833	H. 73. 730 733	N. 63. 630 633	R. 53. 530 533	Z. 43. 430 433

B.		F.		I.		Q.		S.	
Ba	918	Fa	928	Ia	938	Qua	948	Sta	958
Be	917	Fe	927	Ie	937	Que	947	Ste	957
Bi	916	Fi	926	Ii	936	Qui	946	Sti	956
Bo	915	Fo	925	Io	935	Quo	945	Sto	955
Bu	914	Fu	924	Iu	934	Quu	944	Stu	954
				L.		R.			
Bra	818	Fra	828	La	838	Ra	848	Stra	858
Bre	817	Fre	827	Le	837	Re	847	Stre	857
Bri	816	Fri	826	Li	836	Ri	846	Stri	856
Bro	815	Fro	825	Lo	835	Ro	845	Stro	855
Bru	814	Fru	824	Lu	834	Ru	844	Stru	854
C.		G.		M.		S.		T.	
Ca	718	Ga	728	Ma	738	Sa	748	Ta	758
Ce	717	Ge	727	Me	737	Se	747	Te	757
Ci	716	Gi	726	Mi	736	Si	746	Ti	756
Co	715	Go	725	Mo	735	So	745	To	755
Cu	714	Gu	724	Mu	734	Su	744	Tu	754
				N.					
Cra	618	Gna	628	Na	638	Sca	648	Tra	658
Cre	617	Gne	627	Ne	637	Sce	647	Tre	657
Cri	616	Gni	626	Ni	636	Sci	646	Tri	656
Cro	615	Gno	625	No	635	Sco	645	Tro	655
Cru	614	Gnu	624	Nu	634	Scu	644	Tru	654
D.				P.				V.	
Da	518	Gra	528	Pa	538	Spa	548	Va	558
De	517	Gre	527	Pe	537	Spe	547	Ve	557
Di	516	Gri	526	Pi	536	Spi	546	Vi	556
Do	515	Gro	525	Po	535	Spo	545	Vo	555
Du	514	Gru	524	Pu	534	Spu	544	Vu	554
		H.						Z.	
Dra	418	Ha	428	Pra	438	Spra	448	Za	458
Dre	417	He	427	Pre	437	Spre	447	Ze	457
Dri	416	Hi	426	Pri	436	Spri	446	Zi	456
Dro	415	Ho	425	Pro	435	Spro	445	Zo	455
Dru	414	Hu	424	Pru	434	Spru	444	Zu	454

Nomi		Proprj	
11.171	Affrica	54.574	Marocco
12.172	Algeri	55.575	Milano
13.173	Alleanza	56.576	Moscovia
14.174	Ambasciatore	57.577	Napoli
15.175	Armata	58.578	Nunzio
16.176	Asia	59.579	Olanda
17.177	Bailo	64.674	Pace
18.178	Bandiera	65.675	Papa
19.179	Barbareschi	66.676	Pietroburgo
20.270	Battaglia	67.677	Polonia
21.271	Berlino	68.678	Portogallo
22.272	Blocco	69.679	Prussia
23.273	Cesare	74.774	Ragusi
24.274	Cifra	75.775	Religione
25.275	Costantinopoli	76.776	Repubblica
26.276	Czara	77.777	Residente
27.277	Dalmazia	78.778	Roma
28.278	Danimarca	79.779	Russia
29.279	Esercito	84.874	Sardegna
30.370	Europa	85.875	Spagna
31.371	Francia	86.876	Segretezza
32.372	Germania	87.877	Senato
33.373	Golfo	88.878	Svezia
34.374	Guerra	89.879	Svizzeri
44.474	Imperatore	94.974	Torino
45.475	Inghilterra	95.975	Toscana
46.476	Inviato	96.976	Tripoli
47.477	Italia	97.977	Tunesi
48.478	Lombardia	98.978	Venezia
49.479	Londra	99.979	Vienna

Abaco

101	significa	1
202		2
303		3
404		4
505		5
606		6
707		7
808		8
909		9
910		0

Numeri Falsi

10.121
40.402
50.503
60.604
70.705
80.806
90.907
91.911
92.922
93.933

Figure 10: Cipher n.15, February 21 1787. *ASVe Cifre, chiavi e scontri di cifra …Busta 3, f.78.*

The last cipher found in the archives, no.15, is dated 1787 and has a note on the back indicating the cipher sheet was received to be copied to a book on February 21, 1787, by Buseniello, and returned on February 28. It is similar to the previous, but an appreciable improvement is the reintroduction of a dictionary, although simplified, and with something new: the words of the dictionary had two homophones each, although very similar, just a dot in the place of a 7, for instance *Affrica* has two ciphers: 11., and 171, *Algeri* has 12., 172 …see figure 10.

14 Conclusions

The 1600s mark a watershed between the golden age of Venetian cryptography and the unstoppable decline that paralleled that of the last two centuries of the Republic of Venice.

Paradoxically, as hinted above, Ottavian Medici can be rated as the most successful Venetian *cifrista*: his ciphers were easier and less safe than Franceschi's or Partenio's, but met no opposition and furthermore, even after his death were still used until the end of the Republic, used in the sense of emulated as a model, imitated and sometimes simply copied and simplified and reduced.

To his credit, he attempted to reinforce the nomenclators with other expedients, such as writing ciphers in a continuous form, and an opportune use of the nulls; but after him there is only a succession of epigones.

The decline of Venetian cryptography ran parallel with the political decline of the Republic, which in the 16th century was still recognized as a major power at the European level, although that was essentially the power of its naval fleet. After the Treaties of Utrecht (1713—1714), Venice had been downsized to little more than what it is today, a destination for world tourism.

An unanswered question remains: to what extent had a similar decline occurred in other European states, from the Papacy to the Habsburg empire, from France to England and Spain? The answer is probably: it varies. As far as Papal ciphers are concerned, the recent decryption of a dispatch from the apostolic nuncio in Brussels in 1721[33] reveals a cipher described by Matteo Argenti in his treatise;[34] apparently the Papal cipher office had also experienced a period of stagnation.

But in the field of cryptanalysis, according to what F. L. Bauer in his *Decrypted Secrets*[35] and David Kahn in his *Codebreakers*[36] the great European powers had developed their own cipher bureaus known as black chambers (*cabinets noirs*), which were increasingly efficient; in Paris the Rossignol became famous, but according to Kahn the best cipher bureau in Europe was the imperial one in Vienna, the *Geheime Kabinettskanzlei*.[37]

Now, it would be of great interest a research in the Vienna archives to see if Venetian ciphers were also systematically decrypted.

Acknowledgments

Special thanks to the whole staff of the State Archives of Venice, for the help and assistance given in these years of research.

A special thanks to Richard Bosch Architect for reviewing the English text.

References

Leon Battista Alberti. 1511. De cyfris. In *ASVe, Chiavi di cifra b.41.* Manoscritto, Venezia.

Matteo Argenti. 1906. Trattatto che insegna a formar cifre di varie sorti ... In *Die Geheimschrift Im Dienste Der Papstlichen Kurie Von Ihren Anfängen Bis Zum Ende Des XVI. Jahrhunderts.* Ferdinand Schöningh, Paderborn.

Friedrich Ludwig Bauer. 1991 - 2007. *Decrypted secrets: Methods and Maxims of Cryptology.* Springer, Berlin.

Paolo Bonavoglia. 2019. The cifra delle caselle a xvi century superencrypted cipher. *Cryptologia.*

Blaise de Vigenère. 1587. *Traicté des chiffres ou secrètes manières d'escrire.* Abel L'Angelier, Paris.

Ioanna Iordanou. 2019. *Venice's secret service.* Oxford University Press, Oxford.

David Kahn. 1967 - 1996. *The codebreakers.* Scribner, New York.

George Lasry and Paolo Bonavoglia. 2022. Deciphering a short papal cipher from 1721. Uppsala. Linköping University Electronic Press.

Aloys Meister. 1902. *Die Anfänge der modernen diplomatischen Geheimschrift.* Ferdinand Schöningh, Paderborn.

Aloys Meister. 1906. *Die Geheimschrift Im Dienste Der Papstlichen Kurie Von Ihren Anfängen Bis Zum Ende Des XVI. Jahrhunderts.* Ferdinand Schöningh, Paderborn.

Luigi Pasini. 1872 - 2019. *Delle scritture in cifra usate nella Repubblica di Venezia.* Aracne, Venezia.

Giambattista [Della] Porta. 1606. *De Furtivis Literarum Notis, Vulgo de Ziferis ...* G. B. Sottile, Napoli.

Paolo Preto. 1994 - 1999. *I servizi segreti di Venezia.* EST Il Saggiatore, Milano.

Johannes Trithemius. 1507 - 1613. *Libri Polygraphiae.* Lazari Zetzneri, Argentorati (Strasbourg).

[33] (Lasry and Bonavoglia, 2022)

[34] (Argenti, 1906)

[35] (Bauer, 1991 2007) p. 71.

[36] (Kahn, 1967 1996)

[37] (Kahn, 1967 1996) p.163. They were able to open sealed parcels with a steam system, extract the letter, decrypt it, reinsert it into the envelope, seal it and forward it to the addressee, unaware of being intercepted and decrypted.

The Making of Fritz Menzer - A Secret Life

Carola Dahlke
Deutsches Museum
Munich / Germany
c.dahlke@deutsches-museum.de

Robert Jahn
Libellulafilm
Berlin / Germany
robert@libellulafilm.de

Abstract

A technical museum and a filmmaker join for a thoroughly researched documentary about Fritz Menzer, the widely unknown German inventor of cipher device 41. In seven episodes, they uncover Menzers secret life, and identify his central role for German cryptology.

1 Introduction

For some stories, an exhibition showcase is simply too small. In 2013, the Deutsches Museum in Munich, Germany, first came across a relic of a so-called cipher device 41 (“Schlüsselgerät 41”, often abbreviated as SG-41). This German cipher machine from World War II was largely unknown. Only few documents were available - and although released by the NSA since 2009, many documents were still censored line by line so that only parts could be read. Even less was known about the inventor, Fritz Menzer. From 2018 to 2022, the Deutsches Museum and filmmaker Robert Jahn therefore carried out a study, resulting in a thoroughly researched documentary series consisting of seven short episodes in which both the life of cryptologist Fritz Menzer and the background story of the cipher device 41 are presented. Since November 2022, the films are worldwide available via the app of the Deutsches Museum. The publication on the internet is planned for autumn 2023.

2 Hunting a Phantom of Cryptology

The Deutsches Museum owns two relics of the cipher device 41 in its collection. One of them was bought at an auction in 2013, the other one was bought from amateur treasure hunters in 2017.

For an exhibition on cryptology that was under construction during the years 2015-2022, extensive research was started by the Deutsches Museum to find out more about the functioning of the devices, the background to their construction, and their influence on the course of World War II.

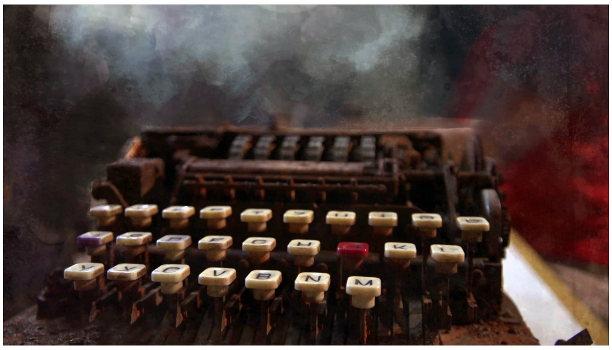


Figure 1: The Schlüsselgerät 41 of the Deutsches Museum found by treasure hunters in 2017

Since the start of these studies on the rare cipher device 41, many findings have already been published, see Dahlke (2018) and the films on Conservation studies of the Leibniz Association (2018). In addition, Kopacz and Reuvers (2021) published the mechanical implementation and algorithm of the device, Lasry (2021) published a study on cryptanalysis in 2021. The fact that the cipher device 41 was technically far ahead of its time is fascinating. The SG-41 has a very sophisticated encryption mechanism that, in terms of security, goes far beyond the well-known German machines of World War II (Lasry, 2021).

But what continued to pose great mysteries, however, was the person of the German inventor Fritz Menzer. For a long time, his name referred primarily to a phantom. For decades, almost nothing was publicly known about his - in fact numerous - inventions and his life. Strangely enough, Fritz Menzer had disappeared from all public documents from 1950 onwards (see e.g. Mowry (1983) and Boghardt (2022)), although he lived until 2005. However, as we know today, he was one of the most central figures in German cryptology during the Second World War. He in-

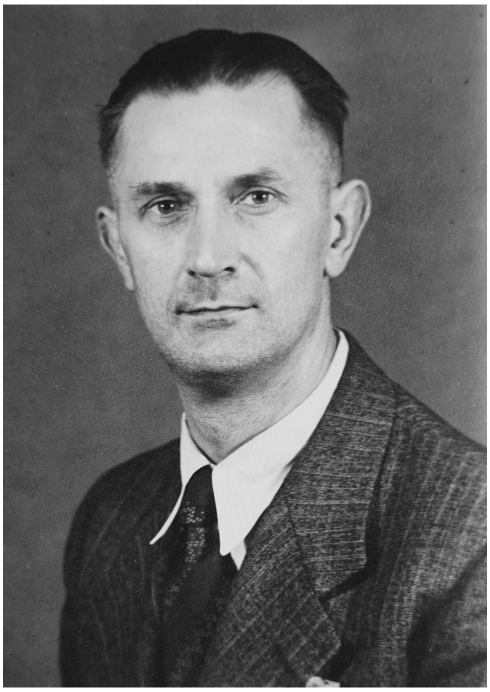


Figure 2: Fritz Menzer, around 1950, by courtesy of Gudrun Jackson

vented new encryption methods, machines and devices, he developed machines to break the Allies' encryptions, and he evaluated the security of existing systems on personel commission of Admiral Canaris.[1]

Furthermore, his inventions continued to have an influence on the development of technology for a long time after the war, not only in Germany. Traces of his inventions can be found in cipher machines used worldwide in the 1950s and 1960s (Kopacz and Reuvers, 2021).

2.1 Traces all over Europe and the USA

The research for this project resembled a criminal investigation to a large extent. When Carola Dahlke, curator of computer science and cryptology at Deutsches Museum and Robert Jahn, filmmaker from Berlin, decided in 2018 to embark on a joint search for Fritz Menzer, not one publicly known photo of Menzer existed.

And to this day, especially in American and Russian archives, many files on Fritz Menzer and his inventions are blocked or partially blacked out. Menzer's traces led to Germany, Austria and Switzerland, to Italy, England, Sweden, Russia, and the USA. First, the documents of the Target Intelligence Committee (TICOM)[2], the internal publications of the NSA[3], and documents from the British codebreakers[4] played a central role. But also the documents of the Wanderer Werke in the Saxon State Archives in Chemnitz[5], the Wehrmacht documents in the Military Archives in Freiburg[6], and the documents of the Bundesarchives Berlin[7] as well as in Koblenz[8] gave important insights. Another valuable source was the archive of the BND[9]. Many documents were only made accessible or released through this study. The researchers received support from many international colleagues. During the years of study, the information from the archives fitted together almost like a jigsaw puzzle, to give a comprehensive picture.

As the researched material turned out to become very extensive, several publications are planned. For a start, this paper focuses on the making of a documentary film for museum purposes. The pub-

[1] See Mowry (1983), p.23-24, Menzer's own memoirs in TICOM DF-174 (1949), p.18-21 and archivales of Bletchley Park, HW 73/4: Personnel of OKW/CHI, pp.3-12 (https://discovery.nationalarchives.gov.uk/details/r/C11204798) and of the BND, AZ 60331, Fragmente der Kriegstagebücher des OKW/Chi 1940-1944, pp.246-259, 272-291, 302-310 and Kothe (1998a), Kothe (1998b)

[2] TICOM Collection: see e.g. the overview volumes of WDGAS-14, the TICOM I-series at archive.org/details/ticom/: I-20, I-21, I-31, I-46, I-57, I-58, I-84, I-92, I-111, I-118, I-123, I-181, I-194, I-200, I-201, I-202, I-206, and the National Archives and Records Administration, College Park, Maryland: 5776 Fritz Menzer Contacts with American and Soviet Authorities (TICOM DF-174) and 5776A: DF-174a,b,c and 5784 List of Germans formerly connected with German Signal Intelligence activities (TICOM DF-185 Part I and Part II); RG 457: Entry P4: Records of the National Security Agency

[3] For NSA-internal publications see e.g. Mowry (1983) and the National Security Agency FOIA Historical Releases, Friedman-Documents: Folder 117 – Document ID A2436243 Report of Visit to Crypto A.G. (Hagelin) by William Friedman, 1955 and Folder 516 – Document ID A4146536 Theoretical Security of the Schlusselkasten, 1949

[4] Bletchley Park intelligence, see the National Archives, Kew, UK: The German Central Cryptographic Organisation (account based on Abwehr and SD sources, HW 73/5, 1939 Sep 01 - 1945 May 08)

[5] See the Sächsisches Staatsarchiv Chemnitz (StAC), Bestand 31030, Wanderer-Werke

[6] See the Bundesarchiv Freiburg, RW 4/777 Organisationsangelegenheiten Chi, innerer Dienstbetrieb, 1944-45 and RW 4/920 WNV/Gruppe Chiffrierstelle OKW - "Der Stand des Chiffrierwesens in der Wehrmacht", 1945

[7] See the Bundesarchiv in Berlin Personenrecherche Menzer, Fritz (Oswin), B563/22074 p 6,66,104,169,199 B563 VI/AS-BRL-1942/495 lfd. Nr. 571 and BArch MfS, Allg/P 7056/61, Aufklärung 37

[8] Bundesarchiv, Koblenz, Vorschlagsliste Nr. 164 des Bundesministers der Finanzen, Archiv Nr. 38879, lfd. Nr. 460

[9] Archiv des Bundesnachrichtendienstes, AZ 60331, Fragmente der Kriegstagebücher des OKW/Chi 1940-1944, pp. 151-323

lication of all findings and knowledge about Fritz Menzer, and about his numerous inventions in the field of cryptology is planned in follow-up studies in the near future.

2.2 Supported by Menzer's Family

One of the most decisive moments in this study was the fact that the researchers were able to make direct contact with Fritz Menzer's only living daughter in the spring of 2020, and with other family members in 2020 and 2021. This offered access not only to Fritz Menzer's life from the Second World War until his death in 2005, but also to personal belongings, photo albums, letters, sound recordings and stories about him. In the course of the study, the name Fritz Menzer to which nothing was known became a real person who held great technical talent. His adventurous life was shaped by the Nazi era, the time of liberation and the emerging cold war. But his life was also marked by a lifelong silence. It was only through the Deutsches Museum's first publications about the Schlüsselgerät 41 that Fritz Menzer's family learned about his inventions and his activities for the Wehrmacht during World War II.

3 Challenges of Museums

Modern exhibition concepts require that visitors approach the objects and contents with all their senses. Precisely because today's generations of visitors are accustomed to receiving topics comprehensively prepared on web-platforms, especially museums focused on history of science are challenged to offer a holistic experience. At the core, of course, is the physical collection, which can only be viewed in a museum. Here, however, the difficult question quickly arises of how to bring the context of objects to the visitor in an interesting, captivating, possibly even fun way. Hands-on stations, audio stories, pictures, dioramas, texts and guided tours play the main role here. But none of these media can comprehensively narrate the history and background of an object. However, research museums in particular have accumulated a profound knowledge about their collection that find little space to be told. To share this knowledge, and to walk the path of research with visitors or maybe take a small journey through time, the cinematic medium offers a wonderful opportunity. At the same time, museums should focus on exhibiting objects, not screens. Therefore, modern exhibition concepts envisage putting available additional material on objects and further facts into an app. Every visitor who has a smart-phone can download the app free of charge and dive deeper into interesting topics if desired, or discover more of the work and processes behind the museum exhibitions (e.g. the depots, workshops, the conservation and research departments).

4 Making of a Documentary

The documentaries created in this study were planned to be accessible in the museum's app. This should enable visitors to access the content directly in the exhibition or anywhere in the world at any time. The complete study, as well as the filmic realisation, was financed by the so-called Future Initiative (the modernisation project), and by the research department of the Deutsches Museum.

In the course of research, Fritz Menzer's life turned out to be very fragmentary. As well, the general knowledge about his life stayed still fragmentary, despite the intensive research. This laid the base for the filmic concept, namely to show important parts of Fritz Menzer's life and work in seven episodes. Every historic fact, event and circumstance shown in the films has been elaborately researched and substantiated by at least two independent sources.

To adopt the filmic approach to the fragmented life of Fritz Menzer the documentary was divided in seven episodes and accompanied by additional audiovisual and text-based material.

To create a complex and precise storytelling that is likewise engaging and emotional, the documentary is based on four narrative layers:

- Interviews with experts and contemporary witnesses
- Filming in sites of historical events
- Historic documents and photographs from public archives and private collections
- Animated moving images designed by the Italian artist Cosimo Miorelli

The aim was to shed light on Fritz Menzers life as well as on the development of cipher machines in the first part of the 20th century, with a very special focus on Menzer's cipher device SG-41. As the development of cipher machines and Menzer's life itself were very much connected to political conditions, it was crucial to include the main political events and figures in the documentary. Namely, the rearmament efforts of Nazi Germany in the 1930s, World War II in its various phases, German military resistance, the importance of concentration camps and forced labour for core German industries, the defeat of Nazi Germany, the emerging Cold War, and finally the reconstruction of military and intelligence capabilities in the two German states.

Each of the seven episodes focuses on a specific event in Menzers life and connects this to a historic key-event. The only exception is episode two, as this episode purely focuses on the aims to investigate the SG-41 as part of the 3D-Cipher project[10].

Menzer's life and work, as shown in the seven main episodes, illustrates the dilemma of German cryptology in the time in and around World War II. What to do with a leading-edge innovation, if it is meant to serve an evil cause? As a scientific piece of work, the documentary does not offer answers to this moral question. But it offers a wide range of newly discovered facts and additional information to the audience.

The following subsections give an overview of the seven episodes of Fritz Menzer - A Secret Life. For each episode, a brief description serves like a teaser as an introduction for the app users to generate interest.

From the Erzgebirge to the world of cryptology

Teaser: "Berlin 1935: The Wehrmacht cryptologist Fritz Menzer negotiates with Boris Hagelin. How can Hagelin's machines be improved? A concept for the future cipher device 41 matures in Menzer's head. But how does a young man from the provinces end up in such a decisive position?"

Filming Locations: Herrndorf, Zug, Brand-Erbisdorf (Saxony), Buildings of former German High Command in Berlin.

Animated content: Menzer's work on the C-35/36 Hagelin machines, the meeting of Boris Hagelin with Menzer and other Wehrmacht officials at OKW/Chi[11], the use of Enigma in the German Wehrmacht, the outbreak of World War II.

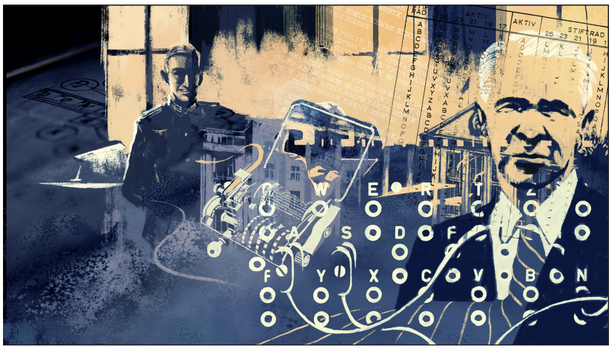


Figure 3: OKW/Chi negotiating with Boris Hagelin in 1935; designed by Cosimo Miorelli

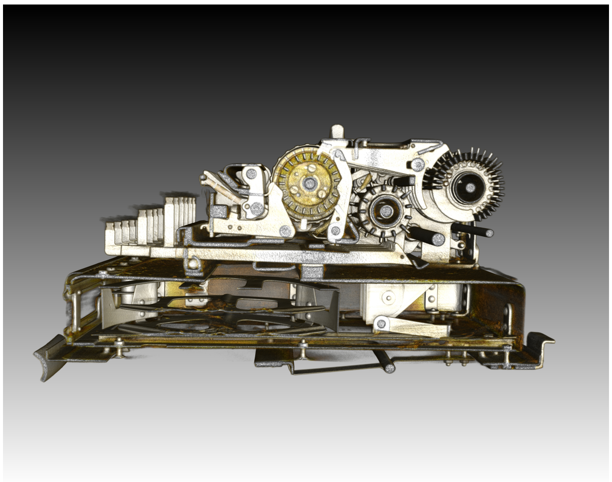


Figure 4: CT-scan of SG-41; scanned and processed at the EZRT, Fraunhofer IIS Fürth

With high-tech to the secrets of SG-41

Teaser: "For almost 80 years, the encryption mechanism of the cipher device 41 has not been fully revealed. And it was only detailed scan data from the 3D-Cipher project that unlocked SG-41's last secrets."

Filming location: Development Center X-Ray Technology EZRT, Fraunhofer IIS (Fürth, Germany)

Animated content: 3-D model of the SG-41.

Now Enigma must die

Teaser: "By the beginning of World War II, the Enigma was already obsolete technology. Fritz Menzer designs a new, secure machine to replace the Enigma. But how can this succeed?"

Filming locations: Buildings of former German High Command (Berlin), Wanderer Werke (Chemnitz), former Bunker of Wehrmacht High Command (Wünsdorf)

Animated content: Forced Labour of prisoners

[10] The 3D-Cipher project focuses on the investigation of cipher devices with computer tomography. More information can be taken from e.g. Göggerle (2022)

[11] See Svensson (2016), p. 27ff and TICOM I-31, p.6

Figure 5: Fritz Menzer at his desk at OKW/Chi; designed by Cosimo Miorelli

of war and prisoners of concentration camps at Wanderer Werke, the US-Air raid on Wanderer-Werke.

British codebreakers on Menzer's traces

Teaser: "Towards the end of World War II, the Bletchley Park codebreakers become aware of the new German cipher device. However, the British secret service had already been keeping a close eye on the inventor Fritz Menzer long before."

Filming location: Bletchley Park (UK)

Animated content: Menzer travelling to locations of the Abwehr in Europe. Menzer's connection to Admiral Canaris. The failed attempt to kill Hitler in 1944 and the assassination of Canaris. Menzer working on the SG-41.

On the run from the Allies

Teaser: "In 1945, British and American special forces search for German crypto experts before they fall into the hands of the Soviets. Target Intelligence Committee - TICOM for short - is the name of the secret project."

Filming locations: Former Bunker of Wehrmacht High Command (Wünsdorf), Werfen (Austria), Bad Aibling (Bavaria)

Animated content: the train-ride of German cipher specialists to Werfen (Austria), arrest of cipher-specialists and transport to POW camp in Bad Aibling (Bavaria). TICOM search for cipher equipment and documents, Menzer's release from POW camp.

A double agent and the secret prison

Teaser: "In 1947, the struggle between the victorious powers of the Second World War for the leading German cryptologists is in full swing. Fritz Menzer finds himself caught between the fronts of the emerging Cold War."

Figure 6: The train-ride of German cipher specialists to Austria in 1945; designed by Cosimo Miorelli

Figure 7: OKW/Chi dumps cipher equipment into the river Salzach in 1945; designed by Cosimo Miorelli

Filming location: Zschopau (Saxony), former secret prison of the Soviet Union in Dresden, surrounding of Berlin and Tempelhof Airfield (Berlin).

Animated content: Menzer in the secret prison of the Soviet Union in Dresden. Soviet officials searching for the prototype of the Schlüsselkasten in Zschopau. Menzers negotiations with the Soviets and his release. Menzers family escaping to Berlin.

New life, new secrets

"Fritz Menzer's ideas are incorporated into the new cipher machines that dominate the world market in the post-war period. However, he himself seems to have left the world of cryptology forever - or has he?"

Filming locations: Former Camp King in Oberursel (Hessen), Former buildung of Bundesschuldenverwaltung in Bad Homburg (Hessen) and Berlin-Tempelhof. Former NSA Intelligence Center at Tempelhof Airfield, former Intelligence Center at Teufelsberg (Berlin), grave of Fritz Menzer

Figure 8: Menzer being released from the army in June 1945; designed by Robert Jahn and Cosimo Miorelli

Figure 9: Menzer being Sowjet prisoner of war in 1948; designed by Cosimo Miorelli

Figure 10: Menzer gets caught between the fronts of the two Cold War powers; designed by Cosimo Miorelli

Figure 11: Menzer starts a new life; designed by Cosimo Miorelli

(Bad Homburg).

Animated content: former OKW/Chi personal working at Camp King in Oberursel (Hessen), Hagelins CX-52 and its connection to SG-41, Menzer caught in between between job offers of the US-Army, Bundeswehr and Stasi. Menzer as the founder of the punch card office of the Bundesschuldenverwaltung.

4.1 Additional Material

The medium app allows for additional material - sound recordings, interviews, specialist information and documents - with which the viewers can acquire even more knowledge for themselves. How can good storytelling be combined with the complex results of the scientific research on Fritz Menzer, and on German cryptology before, during and after World War II, and the role of cipher device SG-41? This was the challenge within this project.

The additional audiovisual and text-based material that is already included in the app will be supplemented within the next months and years. This means that on the one hand, museum visitors will be more and more able to deepen their knowledge in an interactive manner. On the other hand this process keeps the app-content up to date, as with further research, knowledge of Menzer's life and work and his role within German cryptology will widen. This aspect is crucial as there is still a wide range of German, Russian and American archival documents still classified, or in the process of declassification.

4.2 Time-travel

The aim of the study was to make the elaborate research about Fritz Menzer tangible as history. However, since the museum and the filmmaker are dealing with a largely secret world, there is naturally almost no historical visual material. Together with the Italian artist Cosimo Miorelli, a narrative level was therefore developed that creates moving images from the researched facts, historical documents and image fragments.

These images make it possible to immerse emotionally in the life and work of Fritz Menzer, and to explain better the historical backgrounds that influenced his life and actions. As well, this level of imagery makes it possible to show war themes such as bombing, imprisonment and escape without neglecting important events, but also without making them seem idealised. The film series thus create an emotional narrative that combines the personal and the political. At the same time, it is completely fact-based. Miorelli's images were combined with on-site filming at original locations and with personal interviews, where the camera was operated by Thomas Keffel.

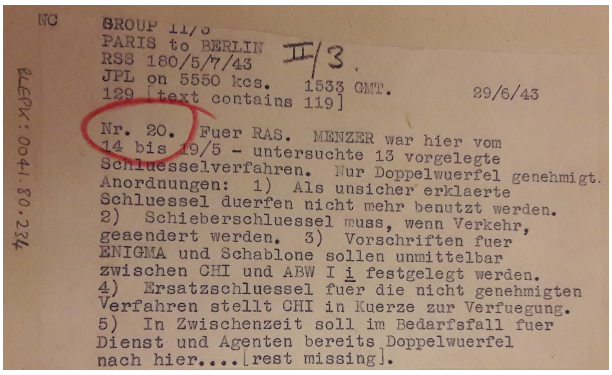
As well, the museum and the filmmaker owe the fact that Fritz Menzer's story can be told in a personal way to the trust of his relatives. They provided extensive private documents as well as pictures and sound recordings (Jackson, 2021; Langer, 2021; Kothe, 1998a; Kothe, 1998b).

4.3 Coming to terms with a difficult period

The general public still knows very little about German cryptology during World War II. Most museum visitors are familiar with the deciphering (especially the British deciphering) of the Enigma because it has become the focus of attention through famous Hollywood films. However, few have delved deeper into the background, the various machines and the people who worked with these objects. The reason for this is that the German war generation kept silent during the post-war period and took their stories to the grave. As a result, very few eyewitness accounts of the events in cryptology have survived. An invaluable find of information is thus provided by the TICOM protocols and a few photos documenting this period.

As well, in 2021, this study came across documents in the Bletchley Park archives[12] showing that the British codebreakers had recognised Fritz Menzer's central role in German cryptology as early as 1943/44, and had reconstructed his travels through Europe - facts that were completely unknown in Germany until a few years ago.

And still, although there is historical research and publication in this field (Weierud and Zabell, 2019; Rezabek, 2020; Dahlke, 2020), there is hardly any well-founded, but rather very sensational reporting on German cryptologists and German cipher machines in the public media. Apparently, the great majority of the German press is not yet in a position to deal with this difficult chapter of history scientifically.



NC

GROUP 11/3
PARIS to BERLIN II/3.
RSS 180/5/7/43
JPL on 5550 kcs. 1533 GMT.
129 [text contains 119] 29/6/43

Nr. 20. Fuer RAS. MENZER war hier vom
14 bis 19/5 - untersuchte 13 vorgelegte
Schluesselverfahren. Nur Doppelwuerfel genehmigt.
Anordnungen: 1) Als unsicher erklaerte
Schluessel duerfen nicht mehr benutzt werden.
2) Schieberschluessel muss, wenn Verkehr,
geaendert werden. 3) Vorschriften fuer
ENIGMA und Schablone sollen unmittelbar
zwischen CHI und ABW I i festgelegt werden.
4) Ersatzschluessel fuer die nicht genehmigten
Verfahren stellt CHI in Kuerze zur Verfuegung.
5) In Zwischenzeit soll im Bedarfsfall fuer
Dienst und Agenten bereits Doppelwuerfel
nach hier....[rest missing].

BLEPK:0041.80.234

Figure 12: Archive material from Bletchley Park (UK): BLEPK:0041.80.234

This proves the importance of the work of museums. Neither can curators just put their collection in a display case and brag as if these were technical German masterpieces of World War II, nor should they hide them away in the depot and not talk about them (according to Kranzberg's first law of technology, see Kranzberg (1986)). The only way to exhibit these objects is through the context of the machines, the stories about the people of the time and the reasons why something was constructed. This is the only way German museums can come to terms with history and do justice to the interest and scientific aspirations of the third and fourth generation after the war, their main visitors.

4.4 Receptions

The creation of the cipher device 41 is, along with the biography of its inventor Fritz Menzer, an interesting opportunity to illuminate the German history of the 20th century with all its facets, including all its dark sides. The fact that SG-41 came into being at all was due to the weaknesses of the Enigma. The fact that it could no longer be of decisive use in National Socialist warfare is a great stroke of luck. And the fact that it played a decisive role in the development of post-war machines illustrates the continuity that existed in German politics and business after 1945.

The film series is available worldwide via the Deutsches Museum's new app. The voiceover texts are inserted in German and as well, there are subtitles in German and English. It is a kind of experiment to put the broad context to an exhibition on cryptology into such a medium. With the download numbers of the app, good results can be seen - the app is basically downloaded about 6600

[12] HW 73/4: Personnel of OKW/CHI, https://discovery.nationalarchives.gov.uk/details/r/C11204798

times per month, and the Fritz Menzer story has been viewed 696 times since its release on the 8th November 2022 until 20th April 2023.

4.5 Outlook

Surprisingly, many of Menzer's colleagues from the Third Reich's various cipher bureaus continued to work for and with the victorious powers of World War II shortly after the end of the war. This opens up numerous new interesting stories for further investigations in the future.

For the time being, the films have only been released for the museum's app, as the Deutsches Museum and the filmmaker hope to attract a TV station to finance a proper documentary. As well, they have already received numerous requests to rent the films. The films have also already been accepted at film festivals. In the near future, a complete publication of all episodes on the internet is planned, not only with German voiceover, but also with an English version.

Acknowledgement

The authors would like to especially thank Gudrun und David Jackson as well as Andreas Langer, Karin Kothe, Carsten Beier, Manfred Kopp, Dermot Turing, Klaus Kopacz, Nils Reims, Wolfgang Zimmermann, Martin Bitschnau, Claus Taaks and Rudolf Staritz for their great support. As well, many thanks go to Jana Walter, Matthias Göggerle and to Helmuth Trischler. In addition, we thank three anonymous reviewers for their valuable comments and suggestions.

References

Thomas Boghardt. 2022. *Covert Legions, U.S. Army Intelligence in Germany, 1944-1949*. United States Army, Center of Military History.

Carola Dahlke. 2018. What we know about cipher device Schlüsselgerät 41 so far. *Proceedings of the 1st International Conference on Historical Cryptology*, pages 109–111.

Carola Dahlke. 2020. The Auxiliary Devices of OKW/Chi. *Proceedings of the 3rd International Conference on Historical Cryptology*, pages 60–69.

Matthias Goeggerle. 2022. Opening Black Boxes: 3D-CT digitalisation of historical cipher machines. In J. Bowen, editor, *Proceedings of EVA London 2022*, pages 20–22.

Gudrun Jackson. 2021. Photographs of Fritz Menzer. Privatbesitz Familie Jackson.

Klaus Kopacz and Paul Reuvers. 2021. *Schlüsselgerät 41: Technical aspects of the German WWII Hitlermühle*. Cryptomuseum, NL, Eindhoven.

Karin Kothe. 1998a. Audio-Recording of phone call with Fritz Menzer. Privatbesitz Familie Kothe.

Karin Kothe. 1998b. Notebook of conversation with Fritz Menzer. Privatbesitz Familie Kothe.

Melvin Kranzberg. 1986. Technology and History: "Kranzberg's Laws". *Technology and Culture*, 27(3):544–560.

Andreas Langer. 2021. Soldbuch Fritz Menzer, 1940-45. Privatbesitz Familie Langer.

George Lasry. 2021. Modern cryptanalysis of 'Schlüsselgerät 41´. *Proceedings of the 4th International Conference on Historical Cryptology*, pages 101–110.

Fritz Menzer. 1949. 5776 Fritz Menzer Contacts with American and Soviet Authorities (TICOM DF-174). The National Archives and Records Administration RG 457: Records of the National Security Agency, College Park, Mayland.

David Mowry. 1983. Regierungs-Oberinspektor Fritz Menzer: Cryptographic Inventor Extraordinaire. *Cryptologic Quarterly Articles*, 2(3–4):21–36.

Deutsches Museum and Libellula-Film. 2018. Cipher machine 41. In *Preserving for the Future*. Leibniz Association.

Randy Rezabek. 2020. TICOM and the Search for OKW/Chi. *Cryptologia*, 37(2):139–153.

Sixten Svensson. 2016. *Boris-Projektet*. Vaktel Förlag.

Frode Weierud and Sandy Zabell. 2019. German mathematicians and cryptology in WWII. *Cryptologia*, 44(2):97–171.

A “Mirror for All Traitors”. Captured Ciphertexts from a Portuguese Spy in Dutch Brazil (1646)

Jörgen Dinnissen
Historian,
The Netherlands
dinnissen.jorgen@gmail.com

Hugo Araújo
Post Doc Researcher,
Federal University of Santa Maria,
Brazil
hugoaffa@hotmail.com

Abstract

A deciphering report found in the National Archives at The Hague presents an intriguing story. A Portuguese spy inside the walls of Recife gathered information about the Dutch defences and wrote it in encrypted letters addressed to the Portuguese rebels that besieged the heart of the Dutch West India Company (WIC) administration in Brazil. The encrypted letters were delivered to the Dutch authorities, who summoned a Jewish cryptanalyst to read them. The report of Abraham de Pina contains a detailed description of the process he used to decipher these letters and presents the complete content of all four ciphertexts. In this paper, we will reconstruct the events of this case and analyze the design of the nomenclature cipher used by the Portuguese rebels. We also will present the flow of information of these intercepted letters within the WIC in Brazil and between them and their company superiors, the Gentlemen XIX, in the Netherlands.

1 Introduction

A group of Portuguese collaborators in Brazil rebelled against the Dutch in 1645, starting a war to reclaim the territories occupied by the WIC (Dutch West Indies Company) since 1630.[1] The conflict lingered until 1654 when Dutch forces capitulated. In the early years of the revolt, the Portuguese held Recife and Mauritsstad under siege. This paper focuses on events that occurred during the siege when WIC forces struggled with the lack of supplies and support from the Netherlands. In this context, Portuguese spies in Recife used ciphers and signal communication to inform the rebel army about the Dutch situation.[2]

On May 8 1646, Antonio Bugalho (“a mulato from Angola”) delivered a little box with hidden encrypted letters to the High Council of Recife in Dutch Brazil. He was ordered to deliver this box to the Portuguese rebels by João Vieira d’Alagoa,[3] one of the last Portuguese who remained in Recife pretending loyalty to the Dutch. Vieira’s decision to spy on the Dutch can be attributed to his debts to the company and his involvement in exploring Brazilwood[4] as a contractor for the WIC, which provided opportunities for him to establish connections and gather information from Dutch officials.[5] There are elements to believe he turned against the Dutch at least since 1644 when other rebels visited him and helped to design the cipher he used. Bugalho's betrayal leads to the imprisonment of João Vieira d'Alagoa. A search of Vieira's house revealed more ciphertexts and notes in Portuguese, proving that he was responsible for the espionage and secret communication. On May 29 1646, Vieira was found guilty of high treason by the Dutch based

[1] Araújo (2022) p.2-7.

[2] As Comissoli (2021: 7) indicates, "The Iberian Atlantic witnessed many espionage actions, although mentions of this are non-systematic and most reports were secondhand, narratives in which other people mentioned spies. Reports written by spies are rare. Similarly, identifying their names is difficult, since the need for discretion led them not to sign their messages."

[3] “d'Alagoa” is not a surname for João Vieira. It is probably a reference to where he lived in Pernambuco, which used to differentiate him from others since João Vieira is a common name in Portuguese.

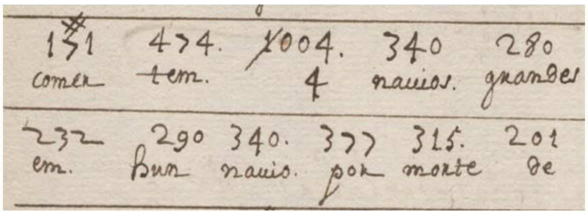
[4] Brazilwood is a timber tree used to make red dye.

[5] Hoge Raad, 1644.

on three pieces of evidence: (1) the decipherment of his ciphertexts, (2) the testimonies of Francisco Ribeiro (another Portuguese who still lived among the Dutch) and Antonio Bugalho against him, and (3) his own confessions. For these actions, he was rigorously punished "as a mirror for all traitors" (Kort Discours Rebellye, 1647: on 30th of May). Dutch authorities confiscated all his goods and properties; he was then publicly executed. They displayed his head on a stake, then quartered and hung his body on half gibbets.

The ciphertext from De Pina is a five-page manuscript found in the WIC documents at the Dutch National Archives, in The Hague, the Netherlands. The description made by the archivist[6] reads: “Statement by Abraham de Pina, in which he gives the key to the number- and secret scripture that members of the Alto Segredo Concilio and the Concilio da Justicia are using in correspondence with him, and letters deciphered with the aid of it. May 1646.”

In this paper, we will show that the archivist of the National Archives made a mistake. De Pina was the cryptanalyst, and João Vieira d'Alagoa was the spy that used this ciphertext to communicate with the enemy. Understanding the different roles of these two characters helps us to comprehend more about the use of cryptography in this particular context.



171	474.	1004.	340	280	
comer	~~tem.~~	4	navios.	grandes	
232	290	340.	377	315.	201
em.	hum	navio.	por	morte	de

Figure 1: Eleven ciphercode nomenclature elements and their plaintext-words from De Pina (1646).

The paper is organized as follows: In Section 2, a reconstruction of the cipher that De Pina received is presented alongside an analysis of his corresponding notes on the rules governing it. Through this comparative approach, we aim to assess De Pina’s aptitude in cryptanalysis. Section 3 shows the key players and the complete trail of the spy letters within the Dutch administrative process. Section 4 examines the conviction of João Vieira d’Alagoa for sending letters to the enemy and investigates whether he was falsely accused. It analyzes the plausibility of this case by contextualizing it and scrutinizing the espionage report in light of information about Dutch defenses in Brazil. Finally, Section 5 concludes this paper.

2 De Pina's Cryptanalysis in Dutch Brazil

Abraham de Pina, also known as Aarão de Pina or Aarão Sarfati (his Jewish name), was a merchant of Iberian descent who arrived in Dutch Brazil in 1636.[7] Despite historical evidence suggesting that his correct name is Aarão Sarfati, we will use Abraham de Pina for consistency since it is the name presented in our source.

Not long after the arrest of João Vieira d'Alagoa, Dutch officials asked Abraham de Pina to decipher the ciphertext that Vieira tried to send to the Portuguese rebels. For several days employees of the WIC tried in vain to decipher the four letters, having the two written pages with the Portuguese alphabet at their disposal. On the other hand, De Pina managed to read its contents using his knowledge of cryptanalysis: “by a certain count table or alphabet what each number means” (Hoge Raad, 1646a).

The ciphertexts and other evidence found in different sources (court records and printed accounts) provide clues of what kind of ciphers were used 377 years ago. According to David Kahn, the nomenclature cipher was the predominant cryptographic system during the early modern period. This system “usually had a separate cipher alphabet with homophones and a nomenclature list of names, words, and syllables. This list, originally just of names, gave the system its name: nomenclature” (Kahn, 1996, xvii).

[6] Based on a wrong interpretation of Appendix 4, lines 1-5.

[7] In Recife, he acted as a rabbi and sometimes worked for the Dutch as a contractor, making shirts for the soldiers, extracting brazilwood and as translator. De Pina received four enslaved negros as payment for his deciphering work. With the fall of Dutch Brazil in 1654, he returned to the Netherlands, where he died in 1670. Mello (1989) p.389-390.

2.1 Sheets "Written With the Portuguese Alphabet"

We found evidence that the Portuguese spies used the nomenclature cipher in Brazil. The book of criminal punishment of Vieira (1646) registers that four sheets were uncovered at the detainee's residence within a cabinet. Two of these were written "full with numerical characters" and the other two were "written full with the Portuguese alphabet". The pages with the Portuguese alphabet must have been the nomenclature used by the WIC employees and De Pina because the Portuguese and Dutch alphabets are identical Latin alphabets. Since the nomenclature is a list of words, the pages found with Portuguese words in alphabetical order (*A*, *Ao*, *As*, *Até*, and so on) matches the description.

2.2 Design of the Cipher

Initially, we analyze the distribution of cipher codes and plaintext. The pattern behind De Pina's output in the report to the High Council and the Councils of Justice in Brazil is presented in Table 1. For the complete design, see Appendix 1. Figure 1 and Table 2 depict the actual distribution of plaintexts over cipher codes.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1 to 50	A	A	A	A	A	6th day	A	A	A	A							
51 to 100																	
101 to 150	A	A	A	A	A	6th day	A	A	A	A	A	A	A				
151 to 200		B		B			B	B	B	B	B	B		B			C
201 to 250	D		D	D	D	D	D		D	D							
251 to 300			E	E			F	F				F	F		F	F	

Table 1: Part of the pattern of the cipher.

Ciphercode	Plaintext	English
171#	comer	eat
171	comer	eat
201	de	of
474	tem	has/ have
1004	4	4

Table 2: Example of De Pina's actual distribution of plaintexts over codes.

Plaintext letters are randomly coded with the first letter range below 504, wherein words beginning with *A* fall between 1 to 139, *B* from 145 to 164, *C* from 167 to 198, and so on. However, exceptions lie in words starting with *M* as they range from 309-330 and 495-503. The order of the second letter in each word appears to be random and does not adhere to either an ascending or descending alphabetic sequence. Some plaintext words (14 times) and plaintext numbers (3 times) have two cipher codes, thus, it's homophonic. For example, the plaintext word *Angola* has the codes 8 and 108. The plaintext number *300* has the codes 300 and 1300.

The plaintext numbers are represented in range above 1,000. There are only two exceptions. First, plaintext number *300* falls into the range of the plaintext words (1 to 503). Secondly, the sixth day of the week, *sexta-feira* in Portuguese (English: Friday) has code 6, which is in the first letter range of *A*.

In addition to that, certain cipher codes incorporate the symbol # (in 16 of 472 codes). Also, it is clear that all four cipher letters that De Pina deciphered used the identical nomenclature.

Probably, De Pina did not have the complete nomenclature at his disposal. As we see in Appendix 4, line 119-120 there are four cipher codes that he did not manage to decipher.

2.3 De Pina's Explanation

	Code without #	Code with #	Code 201	Numbers	Plaintext
Ciphercode	352	171#	201	1020	balanca
Rule	minus 1	minus 0	minus 0	minus 1,000	itself
Plaintext-code	351	171	201		
Plaintext	hollandezes	comer	de	20	balanca
Translation English	Dutch	eat	of	20	scale

Table 3: Rules and examples De Pina mentions.

De Pina provided an explanation for the rules he utilized to decipher. In his report, he gives details about the cipher rules (Appendix 4, lines 20-28): "It is warned that the author to write his cipher almost always uses one less than the one he points out, because 474 is 473 and 352 is 351, as I will soon show, and only a few rare times he uses right number and use this sign # and especially the number 201, which he always uses right to it when he wants to say (de). (…) when it says 474, that as I have said, one less is 473, it means TEM and number 352, 1 less means HOLLANDEZES." To summarize, De Pina explains the following rules applied to read the cipher (see Table 3):

- A code without the symbol # means code number minus 1.

- A code with the symbol # means code number minus 0.
- The code 201 is minus 0 and means “de”.
- The words “not [listed] in the alphabet” are written without a code.

De Pina’s report states that the cipher was created by the author along with two Portuguese individuals who had visited Vieira a year prior (Appendix 4, lines 36-39).

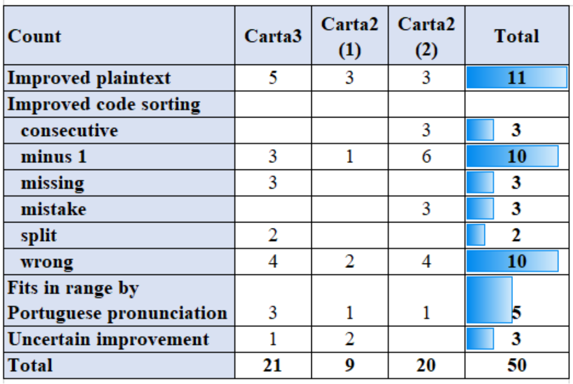
2.4 De Pina's Skill in Cryptanalysis

According to the rules explained by De Pina in the report (see Subsection 2.3) and the distribution of plaintexts over codes (see Subsection 2.2), we conclude that he made some mistakes in the report, in approximately 11% (50 out of 472) of the 472 codes analyzed.[8] These are the mistakes of De Pina:

- Assigned the wrong plaintexts to codes. For example, code 154 for plaintext *hun* (1 time) should read *bastimento* (1 time) in the range of letter B.
- Assigned the wrong codes to plaintexts with a difference of 1. For example, code 377 had plaintext *podemos* (1 time). We also have code 377 for *por* (4 times) and code 378 for *podemos* (1 time). It should have read 378 *podemos*. We identified this as a code error minus 1.
- Assigned the wrong code to a plaintext. Code 258 had plaintext *fora* (1 time). We also have 258 for *forte* (7 times). Based on frequency analysis, the plaintext *forte* should have been assigned to another unknown code in the range of letter *F*.
- Used words that fit in another range by Portuguese pronunciation.[9] Code 349 with plaintext *hollanda* fits in range of letter *O* (and not of letter *H*) because in Portuguese *H* is a silent letter.
- The valuation of a mistake is uncertain in the case of code 1105, which produces plaintext 15 (1 time). It remains unclear whether the error lies with the code or plaintext since a correct reading would yield either 105 or 1015. The discrepancy bears relevance to military strength as it pertains to determining whether there are either 105 soldiers in a fort or just 15.

Most of his 50 mistakes (see Table 4) were minor mistakes. Only in five cases did he pick the wrong plaintext that yielded another reading: *lhe* (English: *you*) instead of the correct *he* (English: *is*); *com* (English: *with*) instead of *tem* (English: *has/ have*); *hun* (English: *one*) instead of *bastimento* (English: *supply*); *e* (English: *and*) instead of *paraiba* (English: *paraiba*). The last mistake with *e*, he made twice.

It is important to note that De Pina wrote the report using the rule “one less” in the plaintext cases. In the report, he indicates: “when it says 474 one less is 473 it means *tem*”; however, the report uses code 474 for the plaintext *tem*. Another example: “when it says 352 one less is 351 it means *hollandezes*”; meanwhile, in the report, code 352 is used for plaintext *hollandezes*. In other words, the report presents the cipher code from the original letters and its reconstructed plaintext.



Count	Carta3	Carta2 (1)	Carta2 (2)	Total
Improved plaintext	5	3	3	11
Improved code sorting				
consecutive			3	3
minus 1	3	1	6	10
missing	3			3
mistake			3	3
split	2			2
wrong	4	2	4	10
Fits in range by Portuguese pronunciation	3	1	1	5
Uncertain improvement	1	2		3
Total	21	9	20	50

Table 4: Mistakes made in letters (Portuguese: cartas).

De Pina accomplished what the WIC staff couldn’t, even with access to the key. It appears that he had previous expertise in cryptanalysis before his arrival in Dutch Brazil.[10] What specific abilities and knowledge were demonstrated in his report?

De Pina was knowledgeable about nomenclatures, which are ciphers that use a list of words assigned to codes. He also discovered an additional security layer known as “super

[8] In the Decode Database at *record 1861*, you can find the complete document with all original and reconstructed codes, plaintexts, and analytics.

[9] Maybe this is not a mistake. In the 17th century there was no rigorous grammar style. *H* has a mute sound in Portuguese, it is not spelled in most words.

[10] Dinnissen and Araújo (2022) and Jütte (2015).

encryption”, where he applied specific rules such as subtracting 1 from the cipher code without symbol # and subtracting 0 from codes with symbol #. Furthermore, he used frequency analysis to decipher certain words like “de” by assigning the code 201 minus 0.

Our analysis evidentiate De Pina’s expertise in cryptography, especially during the short span of time he dedicated to decipher the ciphertext. Besides that, this episode sheds light on the pivotal role played by espionage and cryptology in colonial disputes across the Atlantic.

3 Flow of Information

In this section we demonstrate the flow of information within the administrative process of the WIC. Appendix 2 depicts the exact names mentioned and the information concerning ciphertexts in the sources.

3.1 The Events of May 1646 in Dutch Brazil

On May 8 1646, Antonio Bugalho (referred to as a “mulato”[11]) brought a little box with some letters written with numerical characters from João Vieira d’Alagoa to the High Council in Recife. That same day the Dutch arrested João d’Alagoa and Francisco Ribeiro. These two characters were part of the few Portuguese who still lived among the Dutch after the insurrection of 1645 (Kort Discours Rebellye, 1647).

A week later, on May 15th, Antonio Bugalho requested a third of João Vieira d’Alagoa’s confiscated possessions. Since he betrayed the latter, Bugalho stated that he could not return to his fatherland (Angola) due to fear of Portuguese revenge. Van Walbeeck acknowledged Bugalho as a poor young man who understood the concept of reporting, which granted the right to receive one-third of the offender's possessions (Bugalho, 1646).

When the High Council registered these events on May 16th, they did not mention Bugalho's name. They described the actions of a Portuguese defector from Angola who delivered to them a little box with encrypted (Dutch: *gecijferde*) parchment and some papers from João Vieira d’Alagoa. After being imprisoned, Vieira was tortured and denied being the owner of the letters. The register of the High Council indicates the fiding of similar encrypted papers, but it does not mention the location of the discovery. This source also did not mention De Pina's name and register that “a certain person from the Jewish nation” found the decipherment using “a certain count table or alphabet”. Later, the High Council summoned this person to explain his method. The Jewish cryptanalyst told the High Council that the deciphered letters contained instructions about how the enemy could attack and invade Recife (Hoge Raad, 1646a).

In May 1646, without an exact date, Abraham de Pina wrote a report with the deciphered four ciphertexts written to the enemy between April and the beginning of May. This report does not mention the name of the sender. De Pina indicates it throughout the text as “the author”. However, De Pina indicates that the invention of the nomenclature and additional steps were made by the author in his house, together with Brás Afonso and Manoel João, both described as “from the other side”, i.e. Portuguese rebels. The encrypted letters contained valuable military intelligence, including details on the number of ships and their weapons, fort locations and troop maintenance. The content also provided information on tactics for attacking and communicating specific details through signals regarding food supply, health status, ship transit schedules to destinations like Holland, Guinea and Angola. For a complete transcription and translation of De Pina (1646), we refer to Appendix 4.

On May 29 1646, João Vieira d’Alagoa was convicted for high treason because he corresponded with the enemy. Dutch officials found at the house of Vieira more encrypted letters in the same handwriting. Francisco Ribeiro testified that he saw Vieira cut pieces from a book on which he wrote the ciphertext. These cut pieces fit the indicated book found in Vieira's house. Faced with the evidence, Vieira confessed that the

[11] “Mulato” is a derogatory word in Portuguese and Dutch for describing people of mixed race. Mello (1985) p.222.

letters were indeed his property. See Appendix 3 for the complete transcription and translation of Vieira (1646).

In their daily minutes, the High Council wrote on May 28 1646 (sic) about João Vieira d'Alagoa conviction for sending letters in numerical characters to the enemy. The council showed certainty that he wrote these letters (Hoge Raad, 1646a).

The High Council wrote to the Gentlemen XIX on June 4 1646[12], a letter in favour of Antonio Bugalho, who came from Angola with the yacht *Heemstee*. They paid Bugalho 75 guilders for handing over the box with the letters to them and not to the enemy. On his request, Bugalho went to the fatherland, i.e. the Netherlands. The High Council asked the Gentlemen XIX to give Bugalho an "important and pleasant work" since his favour was meritorious to their state (Hoge Raad, 1646c).

At last, the High Council wrote to the Gentlemen XIX in their periodic report on June 21 1646. They indicated that under letter F was a copy of the deciphered (sic[13]) advice (Dutch: *ontciferde advijsen*), informing that João Vieira d'Alagoa intended to send secret messages to the enemy and for that he received a conviction on May 29th (Hoge Raad, 1646b).

3.2 List of Documents

In the Letterbook (1646) dated approximately June 21 1646, contains a list of relevant documents, ordered alphabetically, for this case:

- Letter F. Copy of an encrypted letter (Dutch: *geciferde brief*) written by João Vieira d'Alagoa.
- Letter L. Periodic report from the High Council.
- Letter V. & deciphered letter (Dutch: *ontcijfferde brief*).
- Letter W. Extract from the criminal verdict against (Dutch: *tegens*) Joan Fer(nan)do Viera.
- Letter Y. Antonio Bugalho mulato.
- Letter Z. Extract from the criminal verdict about (Dutch*: over*) Joan Fer(nan)do Viera.

In Subsection 3.4, we'll explain more about the reconstruction of this letterbook.

3.3 Flow of Information Between Brazil and The Netherlands

The package with letters gathered around June 21 1646, arrived in the Netherlands approximately six weeks later, probably in the first or second week of August. We could not establish if the Gentlemen XIX used this information or if it changed their policy in Dutch Brazil since there are no preserved (secret) minutes of this period.

However, we identify that on October 23 1646, the Gentlemen XIX (1646) sent orders to the High Council Brazil, replying to their report of June 21st. There is no mentioning of João Vieira d'Alagoa's conviction or his ciphertexts. Not even a word about Antonio Bugalho and Franciso Ribeiro. There is a complete silence about the information received about the Portuguese. This information most likely had no direct impact on their strategy in Dutch Brazil. At that time, they had other concerns, like sending troops and supplies to break the siege of Recife and establishing a blockade of Salvador (the capital of Portuguese Brazil) to diverge the Portuguese attention away from Pernambuco.

The siege almost led to the capitulation of the Dutch. As Araújo (2022: 11-12) explains, the capitulation "was only prevented by the arrival in August 1646 of a WIC fleet bringing supplies and military reinforcements. (...) After alleviating the hardships caused by the siege of Recife, the Dutch authorities decided to go on the offensive", launching a naval blockade on the city of Salvador.

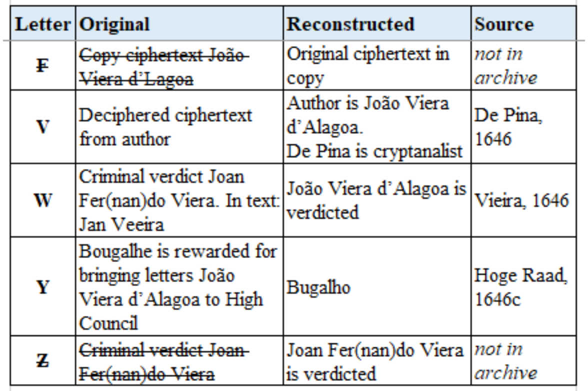
3.4 Jan Veeira and João Vieira d'Alagoa are One and the Same

It cannot be directly confirmed that Jan Veeira and João Vieira d'Alagoa, who was convicted on May

[12] The date reads: 1646-1-4. This information must be an error of the clerk. The arrest of Vieira was on 1646-5-8 and this letter went with the package around 1646-6-21 to the Netherlands. Month must read: June.

[13] In Hoge Raad (1646b) it reads "deciphered". In Letterbook (1646) it reads "encrypted". In Dutch *ontciferde* versus *geciferde*. They are opposites!

29 1646, are the same person. However, by analyzing fragmented evidence from various sources such as the Kort Discours Rebellye (1647), Bugalho (1646), Hoge Raad (1646a, 1646b, 1646c) and Fonseca (1646), we can conclude that they refer to one individual. It should be noted that João Vieira is a common name among the Portuguese which could lead to confusion. Therefore it is important not to confuse João Vieira d'Alagoa with rebel leader João Fernandes Vieira[14] solely based on their shared names.[15] Additional details regarding this case can be found in Appendix 2 where each source's contribuition is explained more extensively.



Letter	Original	Reconstructed	Source
~~F~~	~~Copy ciphertext João Viera d'Lagoa~~	Original ciphertext in copy	*not in archive*
V	Deciphered ciphertext from author	Author is João Viera d'Alagoa. De Pina is cryptanalist	De Pina, 1646
W	Criminal verdict Joan Fer(nan)do Viera. In text: Jan Veeira	João Viera d'Alagoa is verdicted	Vieira, 1646
Y	Bougalhe is rewarded for bringing letters João Viera d'Alagoa to High Council	Bugalho	Hoge Raad, 1646c
~~Z~~	~~Criminal verdict Joan Fer(nan)do Viera~~	Joan Fer(nan)do Viera is verdicted	*not in archive*

Table 5: Information Letterbook (1646) reconstructed.

Isaac Aboab da Fonseca, a prominent leader of the Jewish community of Recife, witnessed the events of the execution of João Vieira d'Alagoa, mentioning his conviction in the poem "*Zekher asiti leniflaot El*" (I have set a memorial to God's miracles). Fonseca (1646) describes the events that followed the insurrection of 1645. As he wrote, the "hardships weakened these people [of Recife], for the conspiracy threatened from within and without. Traitor bastards and black Mamelukes revealed secrets to the enemy to capture Recife, but *the Council decreed the death penalty for one of them: João Vieira de Alagoas.*"

A comparison between the information of letters W and Z in the Letterbook (1646) reveals that Joan Fer(nan)do Viera was sentenced twice, possibly due to a clerk's error, registering the same person under different surnames. The correct surname for João Vieira d'Alagoa should have been used instead of Joan Fer(nan)do Vieira in letter W. However, it is impossible to determine conclusively as letter Z could not be located at the National Archives (The Hague). See Table 5 for futher details.

Unfortunately, the letter F, described as a copy of the ciphertext used by João Vieira d'Alagoa, could not be located within the National Archives.

We believe that letters W and Z are both about João Vieira d'Alagoa, because the High Council wrote on June 21 1646, mentioning that they would send records of the confession and conviction to the Gentlemen XIX. It suggests that there were two separate documents (Hoge Raad, 1646b).

4 The Bigger Picture

In this section, we tackle the inquiries regarding the plausibility of this case by providing context and scrutinizing the espionage report in light of existing information pertaining to the state of Dutch defenses in Brazil.

4.1 Was João Vieira d'Alagoa Framed?

Upon reading Vieira's (1646) case, it appears that the Councils of Justice had legitimate grounds to prosecute João Vieira d'Alagoa in May 1646, as they presented substantial evidence against him. Nonetheless, one cannot completely dismiss the possibility that he was falsely accused or framed. The Dutch pamphlet Brasyls Schuyt-Praetjen (1649) argues that justice in Dutch Brazil was blind and incapable of perception or action. The passage discusses a Dutch practice of falsely incriminating Portuguese residents. A black slave (Dutch: *negeros*) was coerced into delivering a fabricated letter, supposedly written by his master's counterfeited handwriting (Dutch: *konterfeyte sijn handt*), which would harm to the Dutch government. Under "the promise to set him free or some other encouragement", the slave, after being seized by the military, would testify

[14] João Fernandes Vieira was a sugar mill owner who rebelled against the Dutch, leading other rebels in the insurrection of 1645.

[15] Wiesebron (2005) incorrectly indicates that Jan Viera, João Vieira, João Vieeira, and Joan Fernandes Vieira are all referring to João Fernandes Vieira d'Alagoa.

against his master allowing the Dutch officials to seize the goods of the accused traitor.

It is unlikely that the Dutch framed João Vieira d’Alagoa. The complexity of their ciphertext (see letter F in Subsection 3.2) and the report on deciphering suggest authenticity, supported by mistakes in De Pina's report due to an incomplete key. Although it would have required bribery and false confessions for the Dutch to frame him, this hypothesis lacks sufficient evidence.

4.2 Found Ciphertexts are Rare

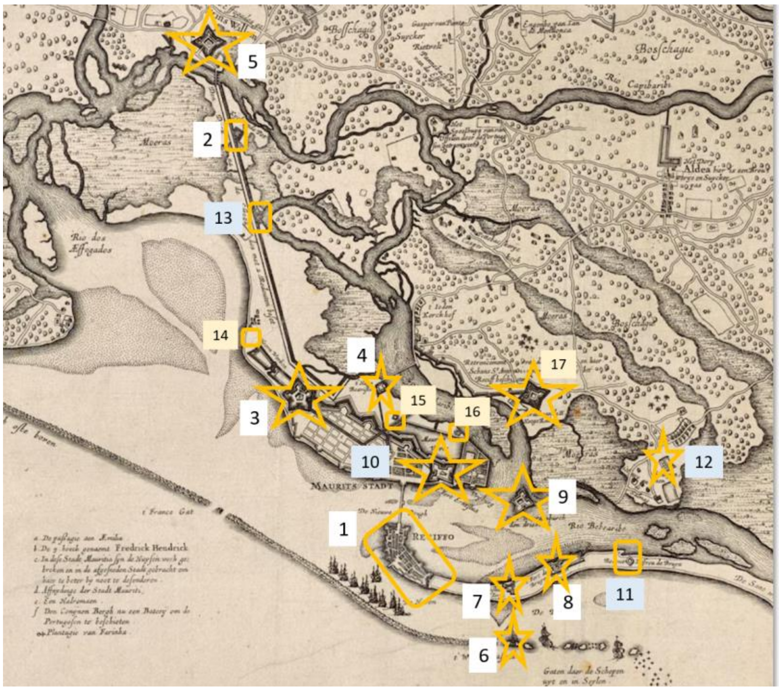
Dinnissen and Araújo (2022) researched the use of ciphertexts during the 17th-century war in Brazil by Portuguese forces. While they discovered some evidence of this practice being employed, it was noted that such occurrences were rare. On the other hand, until now, there is no evidence of use of cryptography by authorities in Dutch Brazil.

The New West India Company (1675-1792) used, more than 120 years later, ciphers. In 1710, in a reply to a letter from February 21 of the same year, the Gentlemen X [16] gave orders to the Council in Guyana, instructing them that encrypted letters (Dutch: *cijfer letteren*) should use the old form instead of the new, advising to be used with great care to avoid mistakes. Otherwise, it could not be properly decrypted (Dutch: *ontcijffert werden*).

4.3 Number of Soldiers in Recife According to Spy Report

In this section, we made a visual aid to understand the information presented in the spy report of João Vieira d’Alagoa (De Pina, 1646). We utilized the map of Recife and Mauritsstad to locate fortifications and their respective garrisons. The ciphertext indicates soldiers' race, distinguishing between Europeans, blacks, and indigenous peoples, as well as their assigned fortification (Appendix 4, line 66-82). See Figure 2 and Table 6.

On the map in Figure 2, we identified numbers 1 to 9 as fortifications that the spy mentions with the number of soldiers. The numbers 10 to 13 are fortifications mentioned by the spy without the number of soldiers. Numbers 14 to 17 are

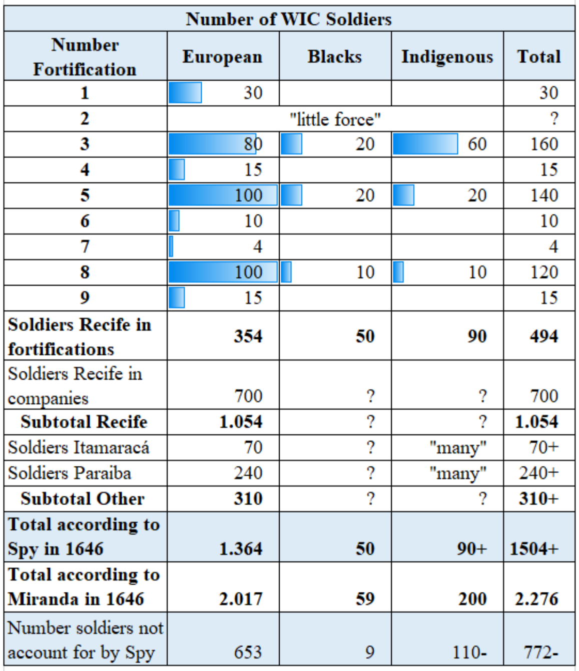


Nr	Portuguese (English)	Dutch
1	“Forte da Cerca” (Fence)	Fence ***of Recife***
2	“Força Taboada” (Timber)	***Redoute Kyk in de Pot***
3	“Os Cinco Cantos” (Five Corners)	Fort Frederik Hendrik
4	“Boa Vista”	Fort Boa Vista
5	“Afogados”	Fort Prins Willem
6	“Forte de Mar” (Sea Fort)	Water casteel
7	“Forte Velho” (Old Fort)	Landt casteel
8	“Forte Bruno”	Fort de Bruyn
9	“Forte Força” (Force)	***Fort Waerdenburgh***
10	“Claustro” (Cloister)	Fort Ernestus
11	“Força Joffr de Bruyn”	Redoute Juffrouw de Bruyn
12	“Salinas”	Fort de Soutpanne
13	“Gato”	Redoute Kat
14	(Stone redout outside 3)	?Fort van Goch?
15	(Second of four redouts landside of Mauritsstad)	Redoute without name
16	(Third of four redouts landside of Mauritsstad)	Redoute without name
17	Forte Antonio Novo	Fort Altena

Figure 2: Dutch fortifications (symbol for forts * and symbol for redouts ■) with their Dutch and Portuguese names in Recife, Brazil. Map is from Cornelis Goliath, ‘Olinda, Maurits-Stadt ende ’t Reciffo’ (1648), engraving published by Claes Jansz. Visscher, Scheepvaartmuseum, inv.nr A.3143 (03).

[16] The company superiors of the Old WIC are the Gentlemen XIX (read: nineteen). The superiors of the New WIC are the Gentlemen X (read: ten).

fortifications not mentioned by the spy in the ciphertexts.[17] These fortifications have different names in Dutch and Portuguese, and sometimes they have more than one name. We identified fortifications through their descriptions, like "forte da Cerca" (1) related to the "fence of Recife". In the case of "força Taboada" (2), we identify it as the redoute "Kyk in de Pot" because its description was a wooden battery surrounded by palisades (*taboada* means *wooden plank* in Portuguese) and its role for defending the fortified dike. We identify "forte Força" (9) as "fort Waerdenburg" since it was a strong fortification, in a key position for the defence of Recife.



Number of WIC Soldiers				
Number Fortification	**European**	**Blacks**	**Indigenous**	**Total**
1	30			30
2	"little force"			?
3	80	20	60	160
4	15			15
5	100	20	20	140
6	10			10
7	4			4
8	100	10	10	120
9	15			15
Soldiers Recife in fortifications	**354**	**50**	**90**	**494**
Soldiers Recife in companies	700	?	?	700
Subtotal Recife	**1.054**	?	?	**1.054**
Soldiers Itamaracá	70	?	"many"	70+
Soldiers Paraiba	240	?	"many"	240+
Subtotal Other	**310**	?	?	**310+**
Total according to Spy in 1646	**1.364**	**50**	**90+**	**1504+**
Total according to Miranda in 1646	**2.017**	**59**	**200**	**2.276**
Number soldiers not account for by Spy	653	9	110-	772-

Table 6: Number of soldiers in Recife according to the spy his report around April 1646.

The "stone redoute outside Fort Frederik Hendrik" (14) is described in other sources of 1646, but its date of construction is unknown. The date of construction of "fort Altena" (17) is uncertain; it was abandoned by the Portuguese and occupied by the Dutch in April 1648; if it existed in April 1646, it was a Portuguese fortification.

It is unclear why the spy doesn't mention the "second and third of four redoutes on the landside of Mauritsstad" (15 and 16). They were built in 1631 and remain visible on later maps.

One possible explanation for the lack of information could be related to the changes in the landscape of Recife, mostly because "After 1645 the city was reorganized so that it could be better defended. Houses were demolished and new fortifications were built, which unfortunately have not been depicted in maps." (Hulsman, 2015: 34). Another possible explanation is that the spy did not have access to the number of soldiers in these redoutes.

By examining additional contemporary sources, we can gauge the reliability of the data provided. The estimated population of Dutch Brazil during this time is 12,703 individuals encompassing all genders and ethnicities including Europeans, blacks, and indigenous peoples.[18] All European soldiers of the WIC, not specified by captaincy, reached a total of 2,017; the lowest number since the beginning of Dutch Brazil in 1630.[19] These numbers are followed by 200 indigenous stationed between Recife and Itamaracá and 59 black soldiers of the company.

The report by the spy notes 354 WIC soldiers, 50 blacks, and 90 indigenous individuals present in the fortifications (refer to Table 6). Moreover, he also includes a count of soldiers involved in operations outside Recife who are not included as garrison troops. The spy mentions in line 163-165: "The six companies in Recife have 360 soldiers. The three companies of Santo Antônio have 120. The Governor's company, Huyter's Company, Claes' company, and Kil's company all have 160. The company has 60 soldiers." Putting together these numbers, we have a total of 700 European soldiers outside the fortifications in Recife. In total there are 1,054 European soldiers in the fortifications (354) and companies (700). That is 52% of the number (2,017) mentioned by Miranda above. If we consider that Recife was the capital of Dutch Brazil this number seems plausible and accurate.

[17] Dating and details based on Hulsman (2015: 27-37) and Miranda (2011: 65).

[18] In Letterbook (1646b) from March 1646 this list (Hoge Raad, 1646d), without a date, is mentioned under reference 'qq'.

[19] Miranda (2011: 38) presents the numbers of WIC soldiers in Brazil between 1630-1654 by year.

These elements corroborate that the spy managed to gather sensitive information about the weakest spots of the Dutch. Besides that, this data serves as a picture of the time endured by the Dutch during the siege.

5 Conclusions

Our investigation revealed that Abraham de Pina was the skilled cryptanalyst who deciphered the four letters written by João Vieira d'Alagoa. In his report to the High Council, he disclosed both his deciphering of the original letters' cipher codes and his efforts towards reconstructing their plaintext. Despite having access to an incomplete nomenclature, De Pina accomplished a remarkable feat in deciphering all of the letters within only several days. Nonetheless, our analysis shows that there were four unresolved ciphercodes within "ciphertext 1".

The collaboration between spy João Vieira d'Alagoa and Portuguese rebels Brás Afonso and Manoel João resulted in the creation of a complex nomenclature cipher that utilized super encryption for added secrecy. This case serves as direct evidence of the use of nomenclatures by the Portuguese in Brazil during the 17th century.

In conclusion, despite De Pina's report detailing a concerning situation of espionage and information gathering for the rebels, the Gentlemen XIX chose not to take any action in response. This decision was reflected in their subsequent orders which indicated that this particular case did not alter their strategy for Brazil.

Acknowledgements

We would like to thank Tiemen Bosma, Suzanne Litrel, Nils Kopal, Adriano Comissoli, and the three anonymous reviewers for their assistance in the preparation of this article.

References

Adriano Comissoli. 2021. 'Spies and Espionage in the Iberian Atlantic'. *Oxford Research Encyclopedia of Latin American History*. New York: Oxford University Press, 2021.

Brasyls Schuyt-Praetjen. 1649. *Ghehouden tusschen een Officier, een Domine, en een Coopman, noopende den Staet van Brasyl:*: ffB2r-v.

Bruno Romero Ferreira Miranda. 2011. *Gente de guerra: Origem, cotidiano e resistência dos soldados do exército da Companhia das Índias Ocidentais no Brasil (1630-1654)*. Ph.D dissertation (Leiden, Netherlands, 2011).

Bugalho, 1646. 1646-5-15. NL-HaNA, OWIC, 1.05.01.01, inv.nr. *62-115*: unfoiled.

Daniel Jütte. 2015. *The Age of Secrecy. Jews, Christians, and the Economy of Secrets, 1400-1800*: 8-11, 26-27, 56-60.

David Kahn. 1996. Revised and updated edition 1967. *The Codebreakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet*. New York: Scribner.

Decode Database. Record 1861. https://de-crypt.org/decrypt-web/RecordsView/1861

Den Haag, Koninklijke Bibliotheek: 76 A 16. Rapport van H. Hamel, A. van Bullestraten en P. Jansen Bas over de toestand in Brazilië. 1646.

Fonseca, 1646. *Zekher asiti leniflaot El.* Translated from Hebrew to Spanish in: Günter Böhm, 1992. *Los sefardíes en los dominios holandeses de América del Sur y del Caribe,1630-1750*. Frankfurt/M: Vervuert, 1992, p. 55.

Gentlemen X, 1710. 1710-9-5. NL-HaNA, WIC, 1.05.01.02, *inventarisnummer 2, 1708 nov. 5 - 1710 okt. 4*: scan 11 (unfoiled), f176v.

Gentlemen XIX, 1646. 1646-10-23. NL-HaNA, OWIC, 1.05.01.01, *kopieboeken van uitgaande stukken, 10 1646 juli 21 - 1657 okt. 10*: ff21-23.

Hoge Raad, 1644. 1644-1-21. NL-HaNA, OWIC, 1.05.01.01, inv.nr. 70: unfoiled, scan 435.

Hoge Raad, 1646a. 1646-5-16, 1646-5-19, and 1646-5-28. NL-HaNA, OWIC, 1.05.01.01, inv.nr. *71*: unfoiled, scans 420, 421, 426 and 437.

Hoge Raad, 1646b. 1646-6-21. NL-HaNA, OWIC, 1.05.01.01, inv.nr. 62-49: unfoiled, scans 1, 10, and 14.

Hoge Raad, 1646c. 1646-1-4. NL-HaNA, OWIC, 1.05.01.01, inv.nr. 62-42: unfoiled.

Hoge Raad, 1646d. 1646-?-?. NL-HaNA, OWIC, 1.05.01.01, inv.nr. 61-51: unfoiled.

Hugo Araújo. 2022. 'The Insurrection of Pernambuco and the Surrender of the Dutch in Brazil (1645–1654)'. *Oxford Research Encyclopedia of Latin American History*. New York: Oxford University Press 2022.

Jörgen Dinnissen and Hugo Araújo. 2022. 'Prey to a Privateer. Two Portuguese Ciphertexts from 1649'. *Proceedings of the 5th International Conference on Historical Cryptology* (HistoCrypt 2022): 50-71.

José Antônio Gonsalves de Mello. 1985. *Fontes para a História do Brasil Holandês.* Vol. 2: A administração da conquista. (Recife: MinC, 1985)

José Antônio Gonsalves de Mello. 1989. *Gente da nação: cristãos-novos e judeus em Pernambuco, 1542–1654* (Recife: Massangana, 1989).

Kort Discours Rebellye. 1647. *Journael ofte kort Discours nopende de Rebellye ende verradelijcke Desseynen der Portugesen, alhier in Brasil voorgenomen, 't welck in Junio 1645 is ontdeckt*.

Letterbook, 1646. 1646-6-21 (around). NL-HaNA, OWIC, 1.05.01.01, inv.nr. *62-45*: unfoiled.

Letterbook, 1646b. 1646-3-?. NL-HaNA, OWIC, 1.05.01.01, inv.nr. *61-56*: unfoiled.

Lodewijk Hulsman. 2015. *Colonial fortifications in Brazil preliminary inventory part 1, Historical research in the Netherlands* (Amsterdam: 2015).

Marianne Wiesebron (ed). 2005. *Brazilië in de Nederlandse Archieven 1624-1654*: 586.

Pina de, 1646. 1646-5-??. NL-HaNA, OWIC, 1.05.01.01, inv.nr. *62-44*: unfoiled.

Vieira, 1646. 1645-05-29. NL-HaNA, OWIC, 1.05.01.01, inv.nr. *62-43*: unfoiled.

APPENDIX 1. Distribution of plaintext-words (first letter) and plaintext-numbers over cipher-codes

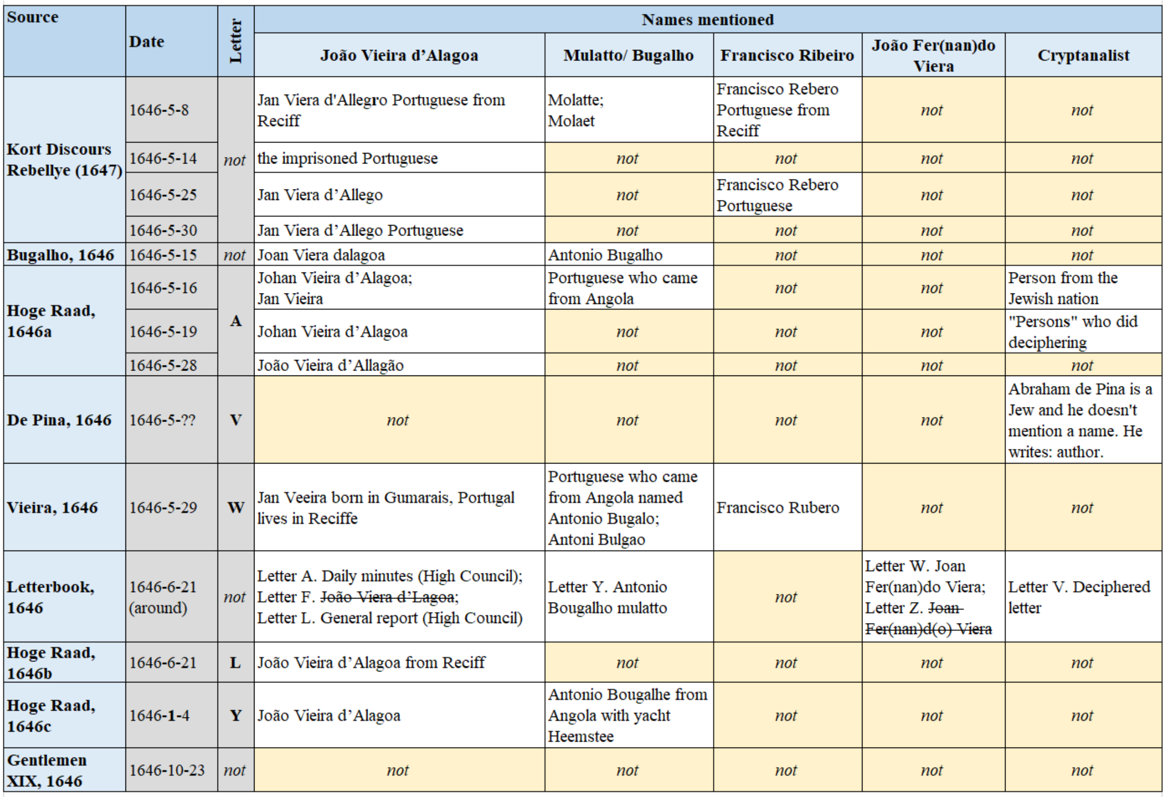
Distribution of plaintext letters and plaintext numbers over ciphercodes

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
1 to 50	A	A	A	A	A	6th day	A	A	A	A																																								
51 to 100																																																		
101 to 150	A	A	A	A	A	6th day	A	A	A	A	A	A	A									A		A															A						B		B		B	B
151 to 200		B		B			B	B	B	B	B	B		B			C		C	C	C	C		C					C						C				C	C			C	C			C	C		
201 to 250	D		D	D	D	D	D		D	D									D									D	D			E	E	E	E	E					E		E	E			E	E	E	
251 to 300			E	E			F	F				F	F		F	F			F	F				F				G		G				G		G	G	H	H	H	H	H	H	I	I				I	300
301 to 350	J	J		J	L			L	M			M	M	M	M		M			M		M								M						N	N	N	N	N	N	N	N		N		N	O	O	O
351 to 400	O	O	P				P		P	P						P	P	P		P					P		P	P	P						P	P	P		P		P	P			P	P		P		
401 to 450				Q				Q							R	R		R				S				S					S								S				S	S			S			S
451 to 500					S				T	T	T			T						T			T	T		T					T			V	V			V	V		V			V	M					
501 to 550		M	M																																															
551 to 600																																																		
601 to 650																																																		
651 to 700																																																		
701 to 750																																																		
751 to 800																																																		
801 to 850																																																		
851 to 900																																																		
901 to 950																																																		
951 to 1000																																																		
1001 to 1050	1	2		4	5	6		8	9	10					15					20					25					30																				
1051 to 1100										60										70										80																				100
1101 to 1150					15/ 105																																													
1151 to 1200																																																		
1201 to 1250																																								240										
1251 to 1300																																																		300

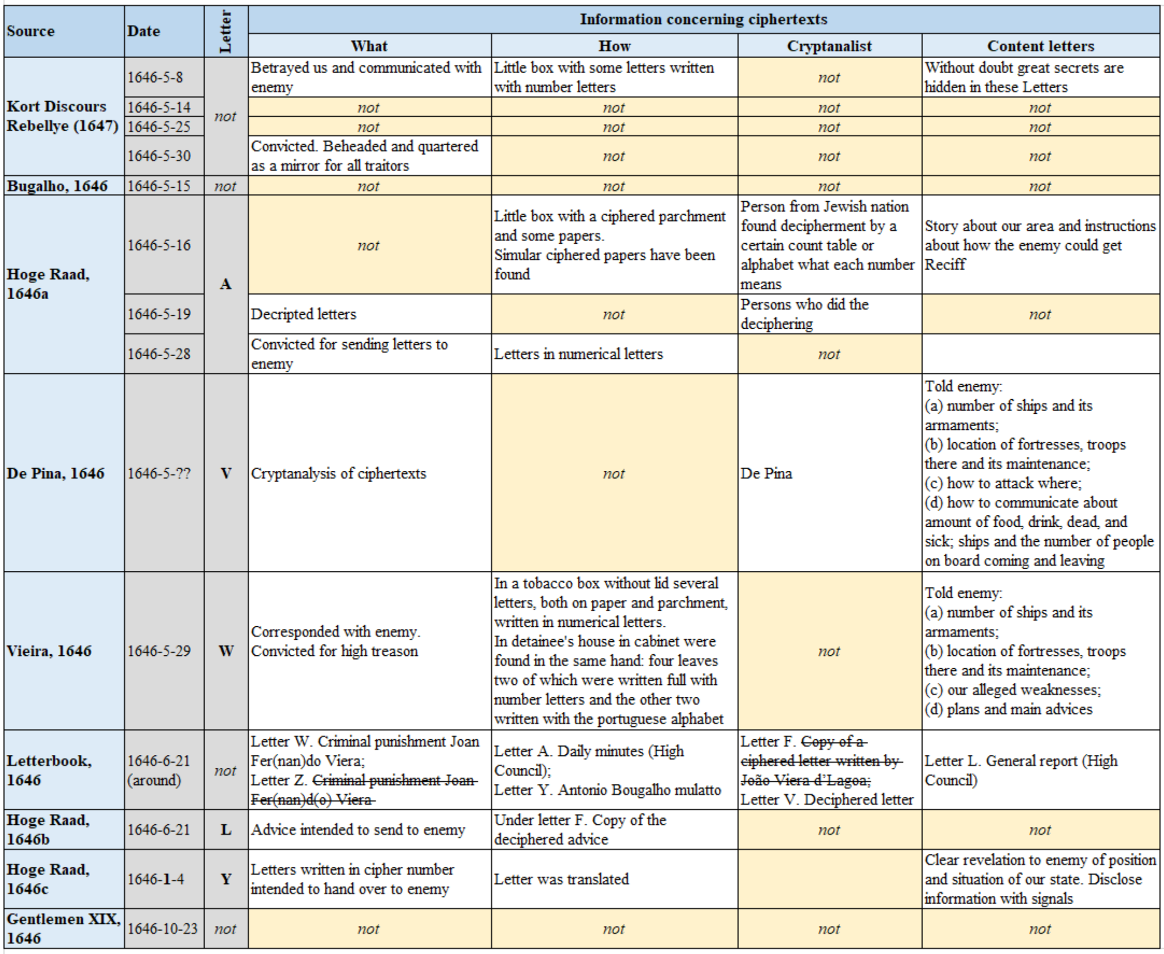
Index	A	Codes with first letter plaintext-word	B	Reconstructed codes with first letter plaintext-word	8	Numbers	

APPENDIX 2. Summary sources. Names mentioned and information concerning ciphertexts

Letters F and Z in Letterbook (1646) have been crossed out because they are mentioned but were not found in the National Archives, The Hague.



Source	Date	Letter	Names mentioned				
			João Vieira d'Alagoa	**Mulatto/ Bugalho**	**Francisco Ribeiro**	**João Fer(nan)do Viera**	**Cryptanalist**
Kort Discours Rebellye (1647)	1646-5-8	*not*	Jan Viera d'Allegro Portuguese from Reciff	Molatte; Molaet	Francisco Rebero Portuguese from Reciff	*not*	*not*
	1646-5-14		the imprisoned Portuguese	*not*	*not*	*not*	*not*
	1646-5-25		Jan Viera d'Allego	*not*	Francisco Rebero Portuguese	*not*	*not*
	1646-5-30		Jan Viera d'Allego Portuguese	*not*	*not*	*not*	*not*
Bugalho, 1646	1646-5-15	*not*	Joan Viera dalagoa	Antonio Bugalho	*not*	*not*	*not*
Hoge Raad, 1646a	1646-5-16	**A**	Johan Vieira d'Alagoa; Jan Vieira	Portuguese who came from Angola	*not*	*not*	Person from the Jewish nation
	1646-5-19		Johan Vieira d'Alagoa	*not*	*not*	*not*	"Persons" who did deciphering
	1646-5-28		João Vieira d'Allagão	*not*	*not*	*not*	*not*
De Pina, 1646	1646-5-??	**V**	*not*	*not*	*not*	*not*	Abraham de Pina is a Jew and he doesn't mention a name. He writes: author.
Vieira, 1646	1646-5-29	**W**	Jan Veeira born in Gumarais, Portugal lives in Reciffe	Portuguese who came from Angola named Antonio Bugalo; Antoni Bulgao	Francisco Rubero	*not*	*not*
Letterbook, 1646	1646-6-21 (around)	*not*	Letter A. Daily minutes (High Council); Letter F. ~~João Viera d'Lagoa~~; Letter L. General report (High Council)	Letter Y. Antonio Bougalho mulatto	*not*	Letter W. Joan Fer(nan)do Viera; Letter Z. ~~Joan Fer(nan)d(o) Viera~~	Letter V. Deciphered letter
Hoge Raad, 1646b	1646-6-21	**L**	João Vieira d'Alagoa from Reciff	*not*	*not*	*not*	*not*
Hoge Raad, 1646c	1646-**1**-4	**Y**	João Vieira d'Alagoa	Antonio Bougalhe from Angola with yacht Heemstee	*not*	*not*	*not*
Gentlemen XIX, 1646	1646-10-23	*not*	*not*	*not*	*not*	*not*	*not*



Source	Date	Letter	Information concerning ciphertexts			
			What	**How**	**Cryptanalist**	**Content letters**
Kort Discours Rebellye (1647)	1646-5-8	*not*	Betrayed us and communicated with enemy	Little box with some letters written with number letters	*not*	Without doubt great secrets are hidden in these Letters
	1646-5-14		*not*	*not*	*not*	*not*
	1646-5-25		*not*	*not*	*not*	*not*
	1646-5-30		Convicted. Beheaded and quartered as a mirror for all traitors	*not*	*not*	*not*
Bugalho, 1646	1646-5-15	*not*	*not*	*not*	*not*	*not*
Hoge Raad, 1646a	1646-5-16	**A**	*not*	Little box with a ciphered parchment and some papers. Simular ciphered papers have been found	Person from Jewish nation found decipherment by a certain count table or alphabet what each number means	Story about our area and instructions about how the enemy could get Reciff
	1646-5-19		Decripted letters	*not*	Persons who did the deciphering	*not*
	1646-5-28		Convicted for sending letters to enemy	Letters in numerical letters	*not*	
De Pina, 1646	1646-5-??	**V**	Cryptanalysis of ciphertexts	*not*	De Pina	Told enemy: (a) number of ships and its armaments; (b) location of fortresses, troops there and its maintenance; (c) how to attack where; (d) how to communicate about amount of food, drink, dead, and sick; ships and the number of people on board coming and leaving
Vieira, 1646	1646-5-29	**W**	Corresponded with enemy. Convicted for high treason	In a tobacco box without lid several letters, both on paper and parchment, written in numerical letters. In detainee's house in cabinet were found in the same hand: four leaves two of which were written full with number letters and the other two written with the portuguese alphabet	*not*	Told enemy: (a) number of ships and its armaments; (b) location of fortresses, troops there and its maintenance; (c) our alleged weaknesses; (d) plans and main advices
Letterbook, 1646	1646-6-21 (around)	*not*	Letter W. Criminal punishment Joan Fer(nan)do Viera; Letter Z. ~~Criminal punishment Joan Fer(nan)d(o) Viera~~	Letter A. Daily minutes (High Council); Letter Y. Antonio Bougalho mulatto	Letter F. ~~Copy of a ciphered letter written by João Viera d'Lagoa;~~ Letter V. Deciphered letter	Letter L. General report (High Council)
Hoge Raad, 1646b	1646-6-21	**L**	Advice intended to send to enemy	Under letter F. Copy of the deciphered advice	*not*	*not*
Hoge Raad, 1646c	1646-**1**-4	**Y**	Letters written in cipher number intended to hand over to enemy	Letter was translated		Clear revelation to enemy of position and situation of our state. Disclose information with signals
Gentlemen XIX, 1646	1646-10-23	*not*	*not*	*not*	*not*	*not*

APPENDIX 3. Vieira (1646) transcription and translation

TRANSCRIPTION

(62_43 scan 1)
(ìn margin: a round stamp in purple with text "RIJKSARCHIEF 'S GRAVENHAGE")[20] / 29 mei 1646 (later annotation in lead pencil) / 43 (later annotation in lead pencil)

W[21]

Extract uijt het Criminele Sententieboeck / vande raaden van Justitie

Alsoo Jan Veeira geboortigh van gumarais in poortugael / omtrent out 40 Jaren, tegenwoordigh gedetineerde, niet / tegenstaende de generale rebellije vande portugeesen / vergunt was in ruste & vreede alhier op 't Reciff te / verblijven, & sijne woninge & goederen te behouden, echter / misbruijckende de voors(eijde) gunste ende faveur, & / vergetende sijnen Schuldigen plight & eet van getrouwigh(eijt) / bestaen heeft gedurende, dese troebelen met onse vijanden / correspondentie te houden, aen deselve bedeckter wijse, te / Schrijven & haer alsoo alle de gelegentheijt van onsen / staet bekent te maacken sulcx hij getraght heeft te doen, / gevende aen seecker portugees gekomen van Angola / genaemt antonio Bulgalo (naer dat hij hem alrede / tot het overloopen getracht hadde te induceeren) seecker / Tabacx doosken sonder Scheedel bedeckt, met een houtten / bodemtgen, & eerst met hars o(ver)loopen & daer op met / metridaet[22] bedeckt daer inne & waernevens hij veeira / hadde gedaen verscheijde brieven, soo op papier als / francijn[23] met cijffer letters geschreeven, bij de welcke / hij aenden gouverneur vanden Vijant Het getal van / onse Scheepen, de Monture van dien, de gelegentheijt / van onse fortten, de besettinge van deselffde als / onderhout, mitgaders alle nootlijckheeden[24] (ver)meijnde / bekent te maacken, & met diverse teijckenen uijt eenige / Hooghten onse desseijnen & voornemen te adviseeren / blijkende tselve evident & klaerlijck bij de voornoemde / geintercipieerde brieven, & nogh eerst bij seeker stucxken, / francijn, hebbende omtrent de lenckte van een vinger / bij de Heeren Commissarisen tot het inventariseeren van den / gedetineerdens goederen gecommitteert Sijnde in / blanco geschreeven **^{g}w : ^{r}s : ^{r}90** [25] Met deselve Hant & letteren / als seecker pampierken in't voors(chreven) doosken bevonden / Twelck met deselve woorden beschreeven was, & eenigh harpuis[26] / sijnde als tgeen, waermeede de brieven int v(oorseijde) doosken / voor te werden waren gepreserveert[27], tsamen in des / gedetineerdens huijse gevonden, Nogh bij seecker pampierken / aldaer bij de Vernoemde, Heeren Commissarisen In seecker / Schiftoor[28] off kasken Met laetgens bevonden hebbende / aen de eene sijde even & deselve maniere van doorgeschrapt / & met een & deselve Hant gemaeckte sijffer letters /

Verto[29]

(62_43 scan 2)
Als int gecijfferde & aen de andere de eijgen hant / vant portugees AlphaBet als inde andere brieff / als int v(oorschreven) doosken bevonden sijn, Mitsgaders uijt seecker / kleijn & smal gecijffert francijntgen doorgesteecken / & geknoopt aen het andere stuck gecijffert parckement / twelck Francisco Rubero gedetineerde inden Rade verclaert / heeft gesien te hebben dat Jan Veeira hetselve van seecker / kleijn boeckgen met een mes

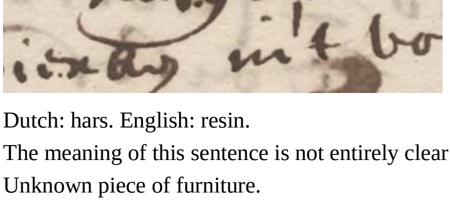
[20] Stamp of the State Archive in The Hague, the Netherlands. Until 1913 this was the name of the current National Archive (Dutch: Nationaal Archief) in the Netherlands.

[21] Letter W refers to list of documents (numbered from A-Z) destined to the WIC, chamber Zeeland (Letterbook, 1646). Transcription: extract wt de Crimineele sententie tegens joan fer(nan)do viera. Translation: extract from the criminal verdict against Joan Fer(nan)do Viera.

[22] English: mithridate or mithridaticum. Generally all-purpose antidote.

[23] Dutch: fransijn. Parchment imported from France. Processed animal skin, used for writing.

[24] Dutch: noodzakelijkheden. English: necessities.

[25] 

[26] Dutch: hars. English: resin.

[27] The meaning of this sentence is not entirely clear. It is probably the fabric or cloth under which the letters were hidden.

[28] Unknown piece of furniture.

[29] Latin: verto. English: I turn. Read: turn page over.

affgesneeden Heeft & / nogh specialijck uijt seecker boecxken, bij de voors(chreven) Heeren / Commissarisen op den 28en deeser ten selven Huijse in seecker / kisgen gevonden sijnde, geintituleert[30] Regras da Compahia / de Jesu[31] van de welcke de gedetineerde verclaerde nogh / een goede partije ten sijnen huijse te hebben aen welcx boecxken / nogh een kleijn stuckxken francijn aen den rugge was gebleven / waer aen de twee eersten gementioneerde stucxkens / parcement gevoeght sijnde de lenghte & breedte in bant / & naeijtsel[32], & alle andere omstandigheeden aen gebleecken / heeft het selffde boecxken te sijn, Waer van hij / gedetineerde de selffde ten deele heeft affgescheurt / ende affgesneeden, & nogh uijt seecker boecxken geintituleert / Primaira examen gene(ra)l quese ad e propone a / todo los que pediere ser admittidas en la de Comp(anhia) / de Jesu[33] gebonden in octavo met swart leer o(ver)trocken, met / de Hant geschreeven sijnde met purper coleur op de suede / geverft (dogh verblickt) & meede ten selven Huijse int / voornoemde Casken gevonden uijt het welcke vier bladeren / waren gescheurt sijnde twee vol sijffer letters / & de andere twee met 't portugees Alphabet volschreven / & int Tabacx doosken bij Jan Veeira aen antoni / bulgao gegeven gevonden, welck boeck neffens / de twee voorseijde brieven aende gedetineerde vertoont / sijnde bekende tselve sijn eigen goet te sijn, & deselve / papieren uijt 't vernoemde boeck gescheurt te Hebben / Wt alle welcke ongetwijffelde & onwederlegge- / lijcke inditien mitsgaders uijt de Verclaringhe van / Francisco Rubero & des gedetineerden gequali-/ ficeerden Confessie inden rade gedaen klaerlijck / gebleecken des gedetineerdens Verradelijcke & / trouloose Minees[34], met dewelcke hij voorgehadt / heeft deesen onsen Staet aende Rebelleuse & / Meijnnedige portugeesen te ontdecken & bekent te / maaken, & ons alle met Vrouwen & kinderen / te stellen in een generael Bloetbadt, sijnde tselve van / seer Schadelijcke, ende pernitieuse[35] gevolge, die in een / lant van pollitije, & daer men gewone Justitie /

(62_43 scan 3)
te administreeren, niet mogen geleeden off / getollereert werden, maer andere ten exempel op het / rigoreuste gestraft, Soo ist dat den rade / van Justitie naer gehoorden eijsch vanden ad(vocaa)t fiscael / de voornoemde klare & onwederleggelijcke inditien / mitsgaders, twelck meer ter materie dienende was / & haer Ed(ele) hadden konnen off mogen moveeren o(ver)wogen / hebbende, doende Reght uijtten naem & van weeghen / de Ho: Mog: Heeren Staaten Generael der / vereenighde Nederlanden Sijn Voocht den Heere / prince van Oragnen, & de generale geoctroijeerde / Westind(ische) Comp(agnie) den voornoemden gedetineerde / verklaert te hebben gelijck sij hem v(er)klaaren bij deesen / Begaen te hebben Crimen Lese Maiestatis[36] / & condemneert hem gebracht te werden ter plaetse / daer men gewoon is criminele Justitie te doen / & aldaer met den Swaerde ter doot geexecuteert / het Hooft gestelt te werden op een staeck het doode / Lighaem gevierendeelt, & ijder vierendeel gehangen / te worden aen halve galgen aende naeste plaetsen / vanden vijant & geconfisqueert alle sijne goederen. / Aldus gedaen & gearresteert inden rade desen 29en / Maij 1646. & gepronuntieert den 30en daer aen / volgenden was onderteeckent B: van groenesteijn

TRANSLATION FROM DUTCH INTO ENGLISH

(62_43 scan 1)
(stamp) / 29 may 1646 (in lead pencil) / 43 (in lead pencil)

W

Extract from the book of Criminal Punishment / of the Councils of Justice

Although Jan Veeira, born from Gumarãis in Portugal, / around 40 years old, nowadays prisoner, not / withstanding the general rebellion of the Portuguese, / was allowed to dwell here on the Recife in rest and peace, / to keep his house and property, he however / abusing the aforementioned benefit and favor, and / forgetting his due duty and oath of loyalty, / had the audacity, during this revolt to correspond with our / enemies, to write to them in a disguised / manner and thus to make known to them / the whole condition of our state, trying to do this: / He gave

[30] Dutch: getiteld. English: titled.

[31] Book 'Regras Da Companhia De Jesu'. Written by the Jesuits. https://books.google.nl/books?id=pqoQD1JTSUkC&hl=nl&pg=PA1#v=onepage&q&f=false

[32] Dutch: de boekband en het naaisel (genaaide gedeelte). English: binding and sewing of the book.

[33] A bound manuscript from the Jesuits. Translation of title: 'First exam which is proposed to all who asked to be admitted by the Company of Jesus'.

[34] Dutch: manieren. English: manners. Not in dictionary, but also found in a pamphlet 'De Quade Minees en Practiken Van seeckeren George Carew, Ondeckt en de open geleght, Tot Onderrichtingh en Waerschouw aen Nederlandt' (Middelburg 1675).

[35] Dutch: gevaarlijke of verderfelijke gevolgen. English: dangerous.

[36] High treason. Latin: Crimen laesae maiestatis. French: Lèse-majesté. Meaning "offence to the majesty", is an offence against the dignity of a state (or its reigning head).

to a certain Portuguese who came from Angola, / named Antonio Bugalo (after trying in advance / to persuade him into defecting) certain / little tobacco box without lid, with a little wooden / bottom, and first doused with resin and thereafter with / mithridate covered, in which and whereby he Veeira / has put several letters, both on paper and / parchment written in numerical letters, in which / he told the governor of the enemy the number of / our ships, their armaments, the location / of our fortresses, with its troops there for its / maintenance, furthermore meaning to disclose all / necessities, and to communicate with several symbols to some / extent our plans and intention. / This turns out to be evident and clear with the aforementioned / intercepted letters, & still first with a certain little piece of / parchment, about the length of a finger, (found) / by the Gentlemen Commissioners with commission to make an inventory of the / goods of the detainee, which was / written in blank **^{g}w : ^{r}s : ^{r}90** in the same hand and letters / as certain little paper in the aforementioned little box, / which was inscribed with the same words, and some resin, / like that, with which the letters in the aforementioned little box / were to be preserved, found together / in the detainee's house. The same with a certain piece of paper / there by the aforementioned Gentlemen Commissioners located in a certain / ‘schiftoor’ or little chest with drawers, which had / at the one side equally and the same way of strikethrough / & with one and the same hand made numerical letters /

Turn page

(62_43 scan 2)
as in the ciphertext and on the other side his own handwriting / with the Portuguese alphabet as in the other letter / as found in the aforementioned little box, also on certain / small and narrow little parchment with some ciphertext pierced / and knotted to the other piece of parchment with ciphertext, / of which Francisco Rubero, detainee, had declared in the council / that he had seen Jan Veeira cutting it of from a certain / small book with a knife and / still especially from certain booklet, by the aforementioned Gentlemen Commissioners on the 28th of this month in the same house / found in some small box, with the title “Regras da Compahia / de Jesu”. Of this (copy) the detainee declares to still / have a good stock at his home. This booklet / still had a small piece of parchment on its back / to which the first two mentioned pieces of / parchment added sum up the length and width in of the / binding and sewing of the book, and with all circumstances showing / to be the same book, from which he, / detainee, declared having torn off / and cut off parts, and also from a certain book titled / “Primaira examen general quese ad e propone a / todo los que pediere ser admittidas en la de Companhia / de Jesu“, bound in octavo covered with black leather, / handwritten, painted with a purple color on the suede / (but faded) and also found in the same house in the / aforementioned little chest from which four leaves / were torn, two of which were fully written with numerical letters / and the other two fully written with the Portuguese alphabet / and found in the tobacco box that Jan Veeira / gave to Antoni Bulgao. Which book besides / the two aforementioned letters shown unto the detainee, / he confessed to be his own property, and to have / torn those papers from the aforementioned book./ From all those unquestionable and irrefutable/ clues together from the statement of / Francisco Rubero and the detainee's qualified / confession in the council had turned out obviously / the detainee’s treacherous and / faithless undermining, by which he had in / mind to disclose and reveal our state to the rebellious and / perjured Portuguese and to / expose us all, including women and children, / to a general massacre. This is the / very harmful and dangerous consequence, that in a / country of police, and where common justice is /

(62_43 scan 3)
administered, should not be suffered or / tolerated, but to others as an example in the most / rigorous way should be punished. Therfore it is that the Council / of Justice after hearing the demand of the Attorney Fiscal / the aforementioned clear and irrefutable clues / together with, which was serving to substantiate more / and could or should having the Honorable Gentlemen[37] moved or / considered, to do justice on behalf of and because of / the High and Mighty Lords of the States General of / United Netherlands, his Guardian the Lord / prince of Orange, and the General Chartered / West Indian Company, have declared to the aforementioned detainee / as they declare him hereby to have / committed High Treason / and condemn him to be taken to the place / where it is customary to punish criminals / and to be executed there with the sword, / the head to be put on a stake, the dead / body quartered, and each quarter must be hung / on half gibbets near the places / of the enemy and all his goods to be confiscated./ Thus done and confirmed in the council this 29th / May 1646, and ruled on 30th following. / Was signed B. van Groenesteijn.

[37] Literally: her Honorable (i.e. from the Councils of Justice).

APPENDIX 4. De Pina (1646) translation and transcription

Translation into English and below it the *transcription*. The lines with the ciphercode and plaintext are treated separately code by code: *code (transcription)*, *plaintext (transcription)*, plaintext normalised (if any), translation into English, code reconstructed (if any), plaintext reconstructed (if any), translation reconstructed (if any). In footnotes the reconstructed code or plaintext will be justified.

[62-44 scan 1]

L1 V[38] / mei 1646[39] / 44
V / may 1646 / 44

L2 Declaration of these alphabetical letters, and ciphers, which before
Declaração destas cartas alfabetas e cifras que diante dos

L3 the men of the high and secret Council and the men of the Council
Homens do Alto e Secreto Conselho e dos homens do Conselho

L4 of Justice by me Abraham de Pina, were declared
da Justiça por mim Abraham de Pina, foram declaradas

L5 in May 1646.
em Maio de 1646

L6 Firstly the alphabet of words so various are like an index of the
Primeiramente o Alfabeto de palavras tão várias são como um índex das

L7 words that whoever uses them in their letters whereby each one of them is
palavras que quem usar nas suas Cartas por onde cada uma dela é

L8 required to have a number which is as follows: the first word is A
necessário ter um número o qual é o seguinte: a primeira palavra é "A",

L9 and must have number 1. The second one is number 2[40], AS three. ATÉ 4, AVENDO Five,
há de ter n°1. A segunda de n° 2, "as" três, "até" 4. "Havendo" cinco

L10 the SEXTA FEIRA 6, until the word 10 ASIMA that will have the number 10.
a "sexta-feira" 6, até a palavra 10 "acima" que terá o n°10.

L11 Apart from this it is necessary that the same words
Fora isto é necessário que as mesmas palavras

L12 from the first A start numbering 101 and go on
desde "A" primeira "A" se comece a numerar 101 e vá

L13 until the end of the whole alphabet
seguindo até o fim de todo alfabeto com o

L14 with the number followed in this way
número seguido desta maneira

L15 so that when you get to Br.ª A.º M.ᵉ J.° it will
com que quando chegares a Br.ª A.º M.ᵉ J.° virá a

L16 touch and it is n° 495 and it will be clear.
tocar e é n°495 e ficará claro.

L17 From understanding last paper n°4 whose words
De entender último papel n° 4 cujas palavras

L18 and numbers follow this alphabet directly
e números seguem a este alfabero diretamente

L19 up to the manufacture of all this key.
até a chave de tudo esta se fabricam.

[38] Letter *V* refers to letter in Letterbook (1646).

[39] 'mei 1646' and '44' are both later annotations written in lead pencil.

[40] Here De Pina forgot to indicate that the Word to number two is AO.

C1	C2		C3
n° 101	*A…*	The	*1*
n° 102	*ao…*	To	*2*
n° 103	*as…*	The	*3*
n° 104	*até...*	Until	*4*
n° 105	*avendo…*	Having	*5*
n° 106	*sexta feira*	Friday	*6*
n° 107	*Alcatifa*	Carpet	*7*
n° 108	*angola..*	Angola	*8*
n° 109	*algunes…*	Some	*9*
n° 110	*acima...*	Above	*10*
n° 111	*a mesmas*	The same	
n° 112	*Águas*	Waters	
n° 113	*a tal*	Thus/ the such	
	segue até o fim acima	continues to the end above	

L20 It is warned that the author to write his cipher almost always uses one
Advirta-se que o Autor para escrever sua cifra usa quase sempre uma
L21 less than the one he points out, because 474 is 473 and 352 is 351, as I will soon
menos da que aponta, porque 474 são 473, e 352 é 351 como logo
L22 show, and only a few rare times he uses right number and use
mostrarei e só algumas raras vezes usa ao justo, se assim lhe assi-
L23 this sign # and especially the number 201, which he always uses
ná-lo este sinal # e principalmente o n° 201 que sempre usa
L24 right to it when he wants to say (de).
ao justo, que quer dizer (de).

L25 Made of the Alphabet an index numbered all the words until the
Feito do alfabete um índex numerado de todas as palavras até o
L26 end there is that when it says 474 (that as I have said one
cabo, achar-se-á que quando diz 474 (que como tenho dito um
L27 less is 473 it means TEM and number 352 1 less means
menos, é 473 que dizer "tem". E numero 352, 1 menos que dizer
L28 HOLLANDEZES, and thus I will begin to decipher the last letter n°3 for
Holandeses, e assim começarei a decifrar a última carta n° 3 por
L29 being brief with the ciphers for better intelligence of the others.
ser breve com as cifras para melhor inteligência das outras

[62-44 scan 2]

L30 It is also necessary to warn that when using a thousand and so many that the thousand will be
Também é necessário advertir que quando usa mil e tantos, que o mil se há
L31 blurred. This is the attention of saying you will find by counting as well as it would be
de borrar, esta é atenção de dizer achereis contando bem muito como se
L32 said that you remember to blur the thousand and so when you find n° 1020 blurring
dissera que se lembrem de borrar o mil e assim quando se achar n° 1020, borrando
L33 the thousand so 1020 becomes 20 and when 1002 comes it is 2, which purpose is
o mil assim ~~1~~020 ficam 20 e quando vier 1002 são 2, o qual o propósito
L34 the same above if it is two or if it is the word n° 2 and in the units
mesmo em cima se são dous ou se é a palavras n° 2 e nas unidades
L35 it almost always uses the right.
quase sempre usa ao justo.

L36 The sign Br.^{a}. A.^{o}. M.^{el}. 10 means Bras Afonso and Manoel João,

A firma Br.$^{a.}$ A.$^{o.}$ M.$^{el.}$ 10 quer dizer Bras Afonso e Manoel João

L37 who are on the other side, and he tells them to declare the letters
os quais estão da outra banda e a lhes manda que declarem as cartas
L38 because, as they were here in his house a year ago,
porque como houvera um ano que estiveram aqui em sua casa, juntos
L39 they communicated this invention of a cipher together.
aí comunicaram esta invenção de cifras.

L40 When he lacks words that are not in the alphabet, like MEÇA,
Quando lhe falta palavra que não está no alfabeto, como "meça",
L41 ABRIL, BALANÇA, COMPANHEIRO and others, he puts the same
"abril", "balança", "companheiro" e, outras assim põe a mesma
L42 word clearly in place of the cipher.
palavra claramente em lugar de cifra.

[**Ciphertext 3**]
L43 Declaration of the letter n° 3 which is the last one from the beginning of May
Declaração da carta n° 3 que é a última do princípio de maio.

L44 The Dutch have six large ships of more than 20 pieces, one of

Code	*474*	*352*	*1006*	*340*	*280*	*201*	*495*	*201*	*1020*	*387*	*1001*	*201*
Plaintext	*tem*	*hollandesses*	*6*	*navios*	*grandes*	*de*	*mais*	*de*	*20*	*pesas*	*hua*	*de*
Plaintext normalised		holandeses								peças	uma	
Translation	has/ have	Dutch	6	ships	big/ large	of	more	of	20	pieces	one	of
Code rec.												
Plaintext rec.		ollandesses[41]										
Translation rec.												

L45 10 pieces, 5 of 6 to 8 pieces, 4 of bronze of all few people.

Code	*1010*	*387*	*1005*	*201*	*1006*	*4*	*1008*	*387*	*1004*	*164*	*201*	*473*	*385*	*278*
Plaintext	*10*	*pesas*	*5*	*de*	*6*	*ate*	*8*	*pesas*	*4*	*brouie*	*de*	*todas*	*pouca*	*gente*
Plaintext normalised		peças				até		peças		bronze				
Translation	10	pieces	5	of	6	until	8	pieces	4	Brass	of	all	few	people
Code rec.														
Plaintext rec.														
Translation rec.														

L46 Recife with the other parts are as told. Fort Taboada [2. Kyk in de Pot] has more fortification.

Code	*415*	*170*	*3*	*201*	*495*	*353*	*253*	*172*	*289*	*207*	*258*	*476*	*474*	*495*	*269*
Plaintext	*recifo*	*com*	*as*	*de*	*mas*	*partes*	*este*	*como*	*lhe*	*ditto*	*forte*	*taboada*	*com*	*mais*	*fortificacion*
Plaintext normalised	Recife				mais					dito		tabuada			
Translation	Recife	with	the	of	more	parts	this	how	you	said	fort	board / timber / plank	with	more	fortification
Code rec.															
Plaintext rec.									he[42]				tem[43]		
Translation rec.									is				has/ have		

L47 A boat came from Angola in 25 days of travel. It gives news that they have little to

Code	*488*	*201*	*8290*	*{empty}*	*154*	*{empty}*	*{empty}*	*1025*	*206*	*201*	*491*	*203*
Plaintext	*veio*	*de*	*angola*	*hu*	*hun*	*barco*	*em*	*25*	*dias*	*de*	*viasem*	*da*

[41] Fits in range of letter O in Portuguese pronunciation.

[42] Plaintext error. Code was '289 lhe'. Should read '289 [h]é'.

[43] Plaintext error based on frequnecy. Code was '474 com' (count 1). Should read '474 tem' (count 15).

Plaintext normalised				um	um						viajem	
Translation	came	of	Angola	one	one	boat	in	25	days	of	voyage	of
Code rec.			8[44]	290[45]		160[46]	232[47]					
Plaintext rec.					bastimento[48]							
Translation rec.					supply							

Code	*342*	*404*	*288*	*367*	*404*
Plaintext	*noticia*	*q[ue]*	*ha*	*pouco*	*q[ue]*
Plaintext normalised			há		
Translation	news	what / which / that	has/ have	little	what / which / that
Code rec.					
Plaintext rec.					
Translation rec.					

L48 eat. There are 4 big ships, 2 patachos walking on the coast. The governor is imprisoned

Code	*171*	*474*	*1004*	*340*	*280*	*1002*	*398*	*404*
Plaintext	*comer#*	*tem*	*4*	*navios*	*grandes*	*2*	*pataxos*	*q[ue]*
Plaintext normalised							patacho	
Translation	eat	has/have	4	ships	big/large	2	patacho	what / which / that
Code rec.								
Plaintext rec.								
Translation rec.								

Code	*falta*	*345*	*174*	*235*	*370/ 360*	*287*
Plaintext	*andao*	*na*	*costa#*	*esta*	*preso*	*gouvernador*
Plaintext normalised	andam			está		governador
Translation	walk / go to	in	coast	is	arrested	governor
Code rec.						
Plaintext rec.						
Translation rec.						

Attention - begin!

The complete ‘APPENDIX 4. De Pina (1646) translation and transcription’ can be found at record 1861 in the Decode Database: https://de-crypt.org/decrypt-web/RecordsView/1861.

Attention - end!

[44] Code error. Code was '8290 angola'. Should read '8 angola' and '290 hu'.

[45] Code error. Code was '8290 angola'. Should read '8 angola' and '290 hu'.

[46] Code error. Code was '{missing code} barco' (count 1). Should read '160 barco' (count 1).

[47] Code error. Code was '{missing code} em' (count 1). Should read '232 em' (count 5).

[48] Plaintext error. Code was '154 hun' (count 1). Should read '154 bastimento' (count 1) in range of letter B.

A WW2 device for breaking the M-209 encryption machine

Magnus Ekhall
Independent Scholar
magnus.ekhall@gmail.com

Klaus Schmeh
Independent Scholar
klaus@schmeh.org

Abstract

According to an eyewitness report by German engineer Reinold Weber published in 2004, a German cryptanalysis unit broke the U.S. cipher machine M-209 in the Second World War. For this purpose, the specialists involved built an electromechanical machine ("Weber machine"), which included binary logic and bore some resemblance with the Turing Bombe. In previously unpublished documents contained in the TICOM reports, information can be found about a cryptanalysis device that is probably identical with the Weber machine. Based on the said sources, this paper describes what is known about this deciphering technology.

1 Introduction

Contrary to the Americans and the British, who concentrated their crypto activities in Arlington Hall and Bletchley Park respectively, the Germans didn't have a centralized cipher authority in the Second World War. Instead, there were about a dozen crypto units in Nazi Germany, operated by different military and civilian authorities without much interchange (Schmeh 2022). This failure in bundling cryptologic forces is today considered a major failure, which contributed to the German World War II cryptology not being as successful as its Allied counterparts.

From a historian's point of view, the fragmentation of the German WW2 crypto efforts makes this topic difficult to research, as the sources are spread to many different locations. The most important source on this topic so far is the information gathered by the TICOM (Target Intelligence Committee), an Allied project aiming to find and seize German intelligence assets, particularly in the field of cryptology and signals intelligence, starting in 1944 (Rezabek 2017). Parts of this material were declassified starting in 2010.

Even before the TICOM files became available to historians, it had been known that German cryptanalysts in World War II had achieved a number of notable successes. Among other things, they broke the U.S. tactical encryption machine M-209, which was designed by the well-known Swedish crypto entrepreneur Boris Hagelin (Leiberich, 1996).

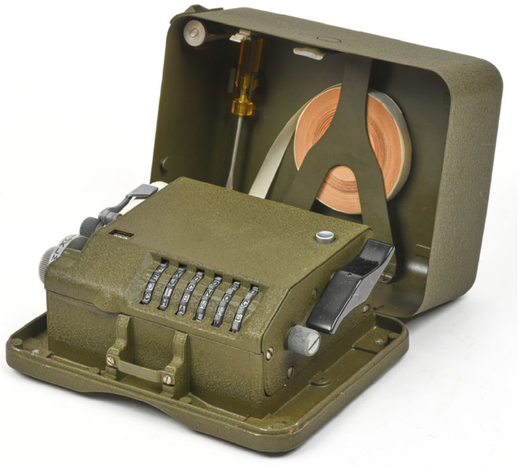


Figure 1. The M-209 was used by the U.S. military for encrypting tactical communication during the Second World War. Source: Cryptomuseum

The M-209 can be regarded as a device that produces a key-dependent pseudo-random sequence (Reuvers 2022b). For encryption, the plaintext is added to this sequence; for decryption, the sequence is subtracted from the ciphertext. The M-209 is a variant of Hagelin's C-38 cipher machine, and it works similarly as the other Hagelin C machines, such as the C-36 and the C-52 (Reuvers 2022a). It uses a key consisting of three parts:

- *Cipher wheel settings*: The M-209 includes six letter wheels that can be set like a combination lock. The setting of these wheels was used as a short-time key, which was changed frequently.
- *Pin settings*: Each cipher wheel comprises a pin for each letter. Each of these pins can be activated or deactivated. Changing this part of the key is more laborious than setting the cipher wheels.
- *Lug positions*: An M-209 includes a drum, on which lugs can be positioned. This part of the key was typically used as a long-term key.

For the cipher wheel settings to be changed, the case of an M-209 device needs to be open, while the machine as such can be closed (War Department 1944). The pin settings and the lug positions, on the other hand, require the machine itself to be opened, too. For this reason, the cipher wheel settings can be regarded as an "external setting", while the pin settings and the lug positions represent the "internal settings".

2 The keying procedure

The M-209 was employed with a keying protocol that ensured that an enemy cryptanalyst had only access to a small amount of ciphertext encrypted with the same key. (War Department 1944) and (Barret 1943) provide information about this procedure.

Another description was created by German cryptologist Alfred Pokorn from OKH/Chi (TICOM 1945b). OKH stands for "Oberkommando des Heeres" ("Army High Command"), while "Chi" is an abbreviation of "Chiffrierabteilung" ("Cipher Department"). Alfred Pokorn might be identical with the author of the books "Pfadfinder-Handbuch" (Pokorn 1950) and "Apatschen-Indianer" (Pokorn 1960). His description is presented in the following:

"When ciphering, the pins and lugs had to be set according to key list and date – 'internal setting', the wheels were turned to any chance position – 'external setting' – and the letters then showing on the wheels were used for the 3rd to 8th letter of the external indicator of the message to be sent. Then any letter of the alphabet was printed several times, this clear letter being used as the first 2 letters of the external indicator. The first letters of the cipher text produced by this printing, were then [set] on the wheels. As only one of the wheels bears all the letters of the alphabet, usually more than 6 letters had to be printed to provide letters suitable for all the wheels. After that, the ciphering could begin."

This means that enciphering a message started with configuring the M-209 according to the daily key by setting the pins on the wheels and the lugs on the drum ("internal setting"), as indicated on some list. Then the sender randomly selected a starting position for the six wheels, say ABCDEF. He then repeatedly enciphered a random letter, say X, and received an output, say GHIJKL ("external setting"). The operator then set the wheels of the M-209 to GHIJKL and started to encipher the actual message. The message indicator in this example would be: XXABC DEFYZ, where YZ designates the cipher key list used.

As Pokorn mentions, the internal setting was changed according to a key list and the date. Most likely the internal setting was changed daily and constitutes a common "daily key" which is shared between all M-209 operators on a given network and period of time. The purpose of the keying procedure is to allow different external settings to be used for each message that is sent. The message indicator, which is transmitted in the clear, can easily be used by the receiving cipher clerk to get the correct external setting for the message in question thus allowing for deciphering of the message. But if you do not know the internal setting, the message indicator is useless.

It is possible for a cryptanalyst to manage to find an internal setting and external setting that correctly deciphers an M-209 message but at the same time not knowing the corresponding letters of the cipher wheels. In this case only this particular message can be deciphered since the absence of the true external setting prevents using the message indicator of other messages. When this situation occurs, it is said that the cryptanalyst has produced a "relative key or setting". If however the true external setting has been found, thus allowing the deciphering of all messages from that network and date, it is referred to as the "absolute key or setting".

3 Reinold Weber's report

In September 2004, the German online magazine Telepolis published an article telling the story of Reinold Weber (1920-2021), who worked as a codebreaker in the Second World War (Schmeh, 2004). In the following, the main facts from this publication are provided. As a caveat, it must be taken into account that Weber's report, which was recorded six decades after the events described, might contain errors. One of the reviewers of this paper has made the authors aware of additional sources that can be used to verify Weber's account. While this task is not within the scope of this work, it is well possible that Weber exaggerates his role in the codebreaking work he was involved in.

According to his report, Weber started his service in Louveciennes near Paris, France, working for a codebreaking unit he remembers as being named FNASt 5. A reviewer told us that Weber probably remembered wrong, and that he instead meant NAASt 5.

According to Weber, his first cryptanalytic success was the breaking of a U.S. code referred to as "TELWA code" by the Germans. This system was later identified as the War Department Telegraph Code, also known as SIGARM (Schmeh 2015). Weber's success (which was rated as exaggerated by a reviewer) earned him recognition among his codebreaking colleagues and led to his promotion to a subunit that took care of more advanced enemy ciphers. When Weber joined this group, his comrades were already able to decipher messages encrypted with the M-209 cipher machine in some cases. The breaking was a laborious and time-consuming process based on manual work only.

4 M-209 cryptanalysis

We don't know how Weber and his co-workers broke M-209 messages. However, one of the authors of this work discovered several descriptions of M-209 cryptanalysis methods in the TICOM files (TICOM undated, TICOM 1948/2), one of which is the aforementioned report by Alfred Pokorn (TICOM 1945b).

As Pokorn writes, it was necessary to first find out how the pins on the wheels and the lugs were arranged on that particular day, i.e. the internal setting had to be determined. For this purpose, one first needed to attack messages in depth, which depended on operator errors. The messages in depth could be broken linguistically, and with the clear-text available one could figure out the pin and lug settings.

At this point one knew exactly, for this particular message, what sequence of active pins had been used by the six wheels of the M-209 for every pair of ciphertext and clear-text letters. However, one did not know where in this pin sequence the letters printed on the M-209 wheels were: where in the sequence is position A, B and so on? In other words, the external setting was still unknown.

To find the external setting, a number of candidate settings were tried, working with the various pin sequences and the indicators of the messages involved and trying to get a consistent situation where the use of the message indicators sets the M-209 wheels in the right place.

For more information about the breaking of the M-209, check (Miller 1950) and (Barret 1943).

5 The Weber machine

In April 1944, Weber had, according to his report, the idea of constructing a machine that would facilitate the M-209 deciphering. This device was to consist, on the one hand, of four Bakelite rollers with slots into which punched sheet metal templates could be inserted to reproduce the relative setting. On the other hand, a relay circuit was planned with a plate above it on which flashlight bulbs marked with letters could be plugged in. The multiple switchable relays were to be soldered to each other and to the electric bulbs inside a box.

As Weber reports, he received permission to build such a machine and traveled to Berlin to ask the company Hollerith, which was later to become a part of IBM, for support. However, his inquiry was rejected, partially because Weber was not allowed to talk about the real purpose of this device.

After the U.S. Army had entered France on D-Day in June 1944, the NAASt 5 was relocated to Germany. It took its home in a former cigar factory in Krofdorf am Gleiberg north of Frankfurt. In Krofdorf there was a precision engineering company called Dönges (today a part of Schunk Phono Systems), which had stocks of silver steel and brass as well as various machining equipment. Weber saw an opportunity to use them to now realize his machine. His supervisor agreed and allowed him to work with a colleague three days a week to build his deciphering device. Although neither of them had any experience with processing metal, the two succeeded in producing the four rollers, each with 26 slots, as well as punched sheet metal plates. In addition, a considerable number of cable connections had to be soldered.

The two decipherers were able to procure the necessary relays, each of which had to be able to establish from one to 256 connections. So, they finally created a machine consisting of two boxes: one the size of a desk, which contained the relays and the four rotating rollers, and another box with edges 80, 80 and 40 centimeters long. The latter box contained 26 by 16 bulb sockets that could be used to replicate the letters of the relative setting. By the end of August 1944, the Weber machine was operable.

According to Weber, his machine needed about seven hours to determine an absolute setting. Without machine aid, this task had lasted about a week when three people worked on it.

At the beginning of 1945, when the U.S. Army approached German territory, the NAASt 5 was relocated again, this time to Salzburg, Austria. To Weber's great surprise, his deciphering machine had also found its way there. However, the unit lacked the radio technology to intercept Allied radio messages, and so the device now proved useless. His superior therefore ordered the machine to be destroyed. With pickaxe, hatchet, hammer and steel saw, Weber then scrapped the device, the construction of which had occupied him for several months.

Until recently, the information provided in Weber's report was everything that was known about the Weber machine. No second source existed. There was no drawing or photograph of the machine.

6 The DF 114 device

In 2022, one of the authors of this paper discovered a document titled "German Cryptanalytic Device for Solution of M-209 Traffic" in the TICOM file (TICOM 1948a). This document, which we will refer to as DF 114, had been declassified in October 2010, six years after the publication of the Weber report. It describes a device used for "machine treatment of AM-1 compromised texts in depth of 5". AM-1 was the name of the M-209 used by the Germans. We'll refer to this machine as DF 114 device.

It is important to note that the TICOM reports, just like the Weber report, need to be approached with care and suspicion. Much of the information provided depends on interrogations conducted months or years after the events occurred and certainly contain errors and inaccuracies. Also, TICOM reports written by the British and US TICOM members have been found to contain factual errors, most likely due to a lack of a complete understanding of the situation when the reports were written. The reports were often based on prisoner-of-war interrogation reports, which often contained errors and incomplete information.

The DF 114 document mentions that it is a translation of a German document catalogued as TICOM 2785 item 19. To our regret, we don't have access to this source.

According to DF 114 the device consisted of three major parts: A "skip box", a distributor and a switching device. Apart from that, there were several auxiliary parts such as a lamp panel, plug-board and various power related electrical components.

The "skip" (German "Sprung") referred to in the skip box refers to how the M-209 enciphers and deciphers a letter: the clear text letter is advanced a number of steps, or skips, in a reversed alphabet in order to produce the resulting letter. This wording is also used for example in (TICOM 1948b).

The skip box contains 120 electrical switches which are spring loaded. Five cylinders are mounted above the array of switches. The cylinders can be fitted with sheet metal lugs and it is these lugs that eventually press the electrical switches when the cylinders are rotated. Each lug has six tabs that can be present or cut away. This represents the state of the six cipher wheels of the M-209 at a given position and can be directly associated with a fixed number of "skips".

The distributor is an electro-mechanical device which steps through the different combinations produced by the skip box and distributes each combination, in turn, to the switching device.

The switching device contains six test circuits, one for each wheel of the M-209. This is a circuit implemented with relays, connected as a binary tree, and lamp panels with flashlight bulbs. The purpose of the switching device is to indicate if all six test circuits conduct: a possible solution has then been found.

DF 114 has rather precise figures related to the electrical characteristics of the device. Voltages and currents are listed with up to two decimal points. Maximum current is described to be 14.31 Ampere and the maximum power consumption 448 Watt.

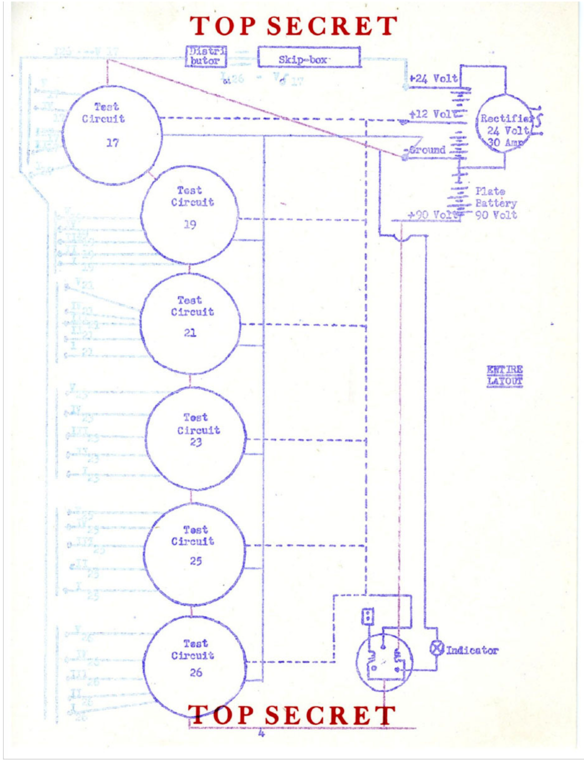


Figure 2. The document DF 114 describes an M-209 breaking machine that is probably identical with the Weber machine.

7 Are the machines identical?

Neither the Weber report nor DF 114 provide an exact description of the devices they cover. Nevertheless, it is clear that both devices served the same purpose and worked in a similar way. In addition, Weber's unit ended up in the Salzburg area, while the first page of DF 114 mentions that the document in question was found buried not far from Salzburg. This suggests that the two machines are identical.

In his report, Weber explains that his machine was used to speed up the process of finding the absolute key given a few relative keys from the same day. This might be related to what is mentioned in (TICOM undated). In other words, the problem to be solved seems to be to find the daily key given a number of message keys. In the following, we assume that the two devices are identical or at least different implementations of the same concept. According to Weber's report, his machine was destroyed by himself in early 1945, before the enemy reached the Salzburg area. This explains why the TICOM commission apparently only found a description of the Weber machine, but not a specimen of the device itself. Of course, it is also possible that several specimens of this machine existed. Perhaps, the design was copied inside the NAASt 5, even without Weber's knowledge.

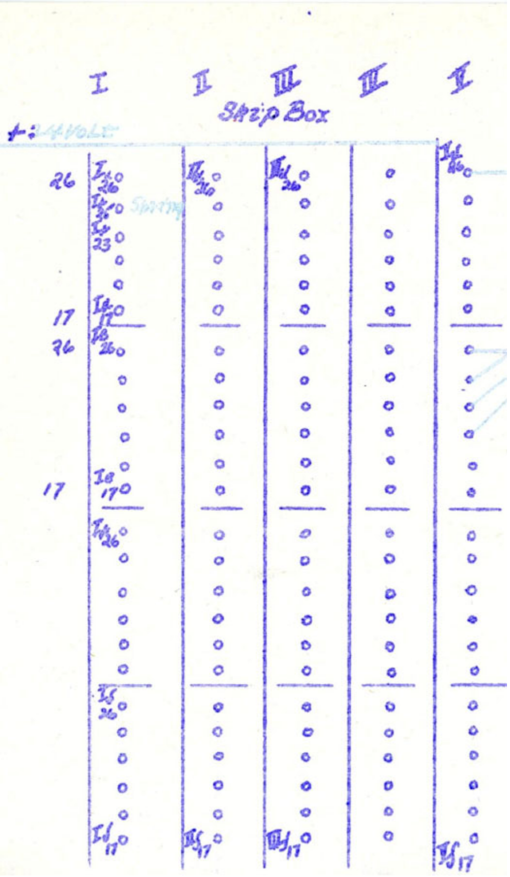


Figure 3. Excerpt from DF 114 showing a schematic view of the skip box. Note the set of six circles representing the six wheels of the M-209. The smallest wheel had 17 letters and the largest 26 which is seen in the annotation of the schematic.

Weber describes his machine as consisting of two boxes. DF 114 does not describe the physical properties of the complete machine but mentions that it consists of three major functional parts. This is not necessarily a contradiction since more than one functional part could be housed in one physical box. In fact, Weber describes one box containing both the relays and the rollers which matches the “skip box” and “switching device” described in DF 114.

A difference worth pointing out is that Weber mentions four rotating rollers but DF 114 counts to five cylinders. Both DF 114 and Weber describe the cylinders having 26 slots where metal lugs could be inserted. This is a detail that is identically described by both sources. The use of light bulbs and relays are also consistent with both sources.

All in all, the similarities between the devices described by Weber and DF 114 are convincing. Both functionally and technically the similarities are many. The few differences that are present are not surprising, considering that the reports about this machine are certainly not error-free.

8 Comparison with the Turing Bombe

The Weber machine is an electromechanical device, just like the Turing Bombe (Turing 2014). It also contains some components that are used in the same way. The Turing Bombe was an electromechanical machine which was used as a tool to help break the German Enigma cipher. The Turing Bombe iterates through a part of the key-space, namely the different Enigma wheel starting positions. For each such position a test is made in order to see if that particular starting position can be ruled out given the current message and crib. If a starting position can not be ruled out, then this position together with some additional output is manually processed further using other types of machines. The Turing Bombe does not output deciphered text but is rather a tool used as part of the deciphering process.

This is similar to the description of Weber's device which was used as a tool to produce the absolute setting given that you already have produced the relative settings. DF 114 further specifies that the device used up to five messages in depth to perform its work.

According to DF 114, the device contains a number of test circuits connected in series, one for each of the six wheels of the M-209. Input from the supposed relative settings are sequentially input through these test circuits and if the test current is able to pass through all six test circuits then a possible solution has been found and a light bulb is switched on. This is similar to how the Turing Bombe uses a test circuit through its drums in order to find a possible solution.

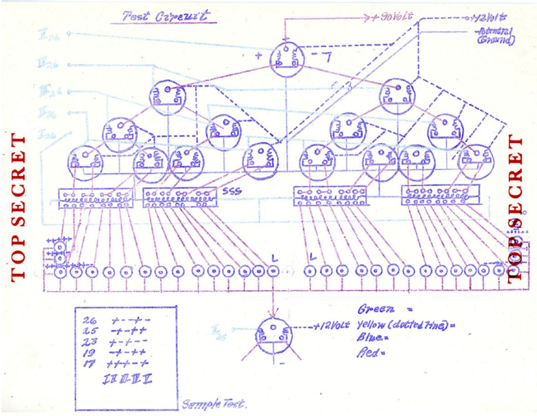


Figure 4. Excerpt from DF 114 showing a test circuit. The test circuit consists of relays connected to form a binary tree into which a test current is injected.

DF 114 describes the use of an electro-mechanical component used to distribute an electric signal sequentially to different outputs. It consists of a rotating arm with carbon contacts bridging two different conductive surfaces. As the arm rotates the pair of conductive surfaces are changed, thus connecting the input to different outputs. A similar solution is used in the US Navy Bombe to generate electrical signals that are synchronized with the mechanical parts of the Bombe (Navy Department 1946).

9 Conclusion

The Weber machine can be named in one breath with World War 2 codebreaking devices such as the Cyclometer, the Bomba, the Turing Bombe, the Desch Bombe, Heath Robinson, Colossus, and the Nightingale. Contrary to the systems mentioned, the Weber machine was constructed by the Germans (Dahlke 2020). The aim of this paper is to provide additional information about this system.

To the regret of the authors, it is still not known how the Weber machine worked. Apparently, the authors of the documents used in this work did not understand this device, either. To close this gap, unfortunately, Reinold Weber can't be asked any more – he died in 2021. Additional sources might exist, and it seems possible to retrieve more information from the TICOM documents referenced in this work, including the figures, which are often difficult to understand. The authors welcome any advice.

Apart from this, a possible way to conclude this device's functioning is to write a computer simulation of it and use real M-209 messages in depth to recover some relative settings that can be used to test the simulation and hence the device. However, writing such a simulation is clearly a difficult task. It would have to be done in stages, trying out various theories about the exact working of the machine.

Acknowledgments

We would like to thank the reviewers of this paper, particularly "Reviewer 1", who provided excellent comments and additional information. We also would like to thank Paul Reuvers and Marc Simons from the Crypto Museum (cryptomuseum.com) for permitting us to use their M-209 photograph, as well as Marek Grajek for shepherding this publication.

References

T. R. W. Burton Miller et al. 1950. Special Conference on M-209 Security. https://www.nsa.gov/Portals/75/documents/news-features/declassified-documents/friedman-documents/patent-equipment/FOLDER_371/41755249079440.pdf.

Carola Dahlke. 2020. The Auxiliary Devices of OKW/Chi. HistoCrypt 2020, Proceedings of the 3rd International Conference on Historical Cryptology.

Otto Leiberich. 1999. Vom diplomatischen Code zur Falltürfunktion. Hundert Jahre Kryptographie in Deutschland. *Spektrum der Wissenschaft*, 6, 26-34.

Navy Department. 1946 *Technical and Theoretical Report of N-530 Bombe*. Navy Department, Washington D.C.

Alfred Pokorn. 1950. *Pfadfinder-Handbuch*. Pfad-Verlag, Salzburg.

Alfred Pokorn. 1964. *Apatschen-Indianer*. Oldenbourg, Munich.

Paul Reuvers and Marc Simons. 2022a. Crypto AG. https://www.cryptomuseum.com/crypto/hagelin.

Paul Reuvers and Marc Simons. 2022b. M-209. https://www.cryptomuseum.com/crypto/hagelin/m209.

Randy Rezabek. 2017. *TICOM: the Hunt for Hitler's Codebreakers*. independently published, Rochester, NY.

Klaus Schmeh. 2015. *Wie ein Rätsel der Kryptologie-Geschichte nach 70 Jahren gelöst wurde*. https://scienceblogs.de/klausis-krypto-kolumne/2015/07/04/wie-ein-raetsel-der-kryptologie-geschichte-nach-70-jahren-geloest-wurde/.

Klaus Schmeh. 2004. *Als Codeknacker im zweiten Weltkrieg*. https://www.heise.de/tp/features/Als-deutscher-Code-Knacker-im-Zweiten-Weltkrieg-3436447.html.

Klaus Schmeh. 2022. *Codeknacker gegen Codemacher*. Springer, Heidelberg. 362-364.

TICOM 1945a. *Consolidated Report on Information Obtained from PW Erdmann, Grübler, Hempel, Karrenberg, Schmitz, Suschowk*. C.S.D.E.I.C. (U.K.), S.I.R. 1717.

TICOM 1945b. *Report by Alfred Pokorn, of OKH/CHI, on M-209*. TICOM Document 2785.

J. C. Barret. 1943. *Signal Operation Instructions*. 84th Infantry Division, 34-45.

TICOM. 1946. *Volume 4 Signal Intelligence Service of the Army High Command*. WDGAS-14, 20-41.

TICOM. 1948a. *German Cryptanalytic Device for Solution of M-209 Traffic*. TICOM Document 2785, DF 114. NARA, NAID: 23889821. https://catalog.archives.gov/id/23889821.

TICOM. 1948b. *Report on The Solution of Messages in Depth of The American Cipher Device M-209*. TICOM Document 2794, DF 120. NARA, NAID: 23889823. https://catalog.archives.gov/id/23889823.

TICOM. Undated. *Determination of the Absolute Setting of the AM-1 (M-209) by Using Two Messages with Different Indicators*. TICOM Document 2795, DF 105. NARA, NAID: 26466553. https://catalog.archives.gov/id/26466553.

Christos Triantafyllopoulos. 2017. *Christos military and intelligence corner*. http://chris-intel-corner.blogspot.com/2017/02/the-compromise-of-croat-enigma-k-cipher.html.

Dermot Turing. 2014. *Demystifying the Bombe*. The History Press, Stroud.

War Department. 1944. *Converter M-209, M-209-A, M-209-B (cipher)*. U.S. War Department.

Mysteries of P.C. Cadix and its evacuation in 1942/43

Marek Grajek
Freelance researcher
mjg@interia.eu

Abstract

From 1940 to 1942 a group of Polish codebreakers, authors of the initial Enigma breakthrough, had been working at the P.C. Cadix, a secret codebreaking center situated in the unoccupied France. Their work in this period, and in particular the circumstances of their evacuation after the German/Italian occupation of Vichy France, are shrouded in mystery. This paper represents an attempt to put information known so far into a broader context, revealing in the process possible distortions and omissions in the accounts of the participants of the events.

1 What did we know so far?

In June 1940 several officers of the pre-war French intelligence services, gathered in the Bon Encontre seminar near Agen, had decided to continue their fight against Germany in spite of their country's defeat. Major Gustave Bertrand, pre-war commanding officer of Section D of the French Service de Renseignement, was in their number. By manipulating the truth, he managed to retain his control over two groups of foreign codebreakers, Poles and Spaniards, constituting the entire assets of his service. Having temporarily secured their members in the French Africa, he started work on reorganizing his service. In November 1940 a country house near Uzès in the non-occupied zone of France, was ready to accept new residents. This is how the clandestine codebreaking center known as P.C. Cadix started its operation.

Clandestine organizations usually preserve a minimal documentation of their activities. P.C. Cadix was no exception. Even declassified in 2015 parts of Bertrand's private archive provide surprisingly limited information on its operations. In this situation any reconstruction of the P.C. Cadix' story must be based on the available personal recollections of its former staff members. Bertrand had published his memories in 1973, unveiling for the first time the information about breaking the Enigma cipher. Marian Rejewski had deposited his memories with the Institute of Military History in 1967, but they were published for the first time only in 2011 (Rejewski 2011). After reaching Great Britain in 1943 Wiktor Michałowski submitted his report from the activities of P.C. Cadix, or rather Ekspozytura 300, as the center was known to the headquarters of the Polish intelligence service in London. Lt Colonel Gwido Langer submitted his report only after his return from the internment camp, in 1945.

Reconstruction of real activities of P.C. Cadix based on those sources is difficult, if possible, at all. P.C. Cadix represented a slightly schizophrenic organization, staking several layers of external and internal conspiracy one upon the other. Its existence and activity were obviously hidden from the Germans, as contradictory to the terms of the armistice treaty. They were kept secret, to the possible extent, also from the Vichy government and the HQ of its army, both pretending to respect the terms of the armistice. The purpose of center's work and its position within the structure of the clandestine French secret services were kept hidden from the teams of Polish and Spanish codebreakers. Both teams were working, and to some extent living, in the strict isolation from each other. Finally, Polish team was working in the conditions of double subordination. It was administratively subordinated to Bertrand, and through him to the French intelligence service, but operationally directly to London HQ of the Polish intelligence – fact well hidden from Bertrand and the French hosts. In the conditions described it is obvious that no member of the team could learn and report a credible story of the P.C. Cadix. Moreover, reports and memories of the team members representing different groups do not sum up to a consistent picture, and include obvious gaps and contradictions.

What did we know then about center's activities, or rather what we believed to know? P.C. Cadix represented a part of Travaux Ruraux, created by Paul Paillole immediately after the collapse of France, on 1st July 1940 (Paillolle, undated). Travaux Ruraux represented a branch of the secret services of the Vichy France masqueraded as a commercial organization specializing in the agricultural works. As a successor to the former codebreaking section of the French military intelligence service, P.C. Cadix was focusing on the signals intelligence and the codebreaking, and consisted of three teams: 'L'équipe D' – group of Spanish republicans dealing mostly with Italian and Spanish ciphers, 'L'équipe Z' – team of Poles breaking the German and Soviet ciphers; finally the French component, responsible for purely administrative functions. Maj. Gustave Bertrand acted as center's commanding officer, with Lt Col Gwido Langer commanding L'équipe Z, and Antonio Camazón L'équipe D. Both intelligence services were aware that the possible imprisonment of any team member by the Germans was a mortal threat not only directly to him, but above all to the secret of breaking the Enigma code.

The real command structure of P.C. Cadix was complicated by the fact that, unbeknown to Bertrand, L'équipe Z was operationally subordinated directly to the London HQ of Polish intelligence service. In that role it was known as "Ekspozytura 300", and its existence seemed to serve a double purpose. First, it permitted keeping the pre-war team of the Cipher Bureau, for political reasons alienated within the structures of the service, far from London. Second, its presence in the southern France offered a comfortable relay service for the messages exchanged with the highly efficient Polish intelligence network in North Africa, "Rygor", too distant to communicate directly with London.

Requirements of the "Rygor" network were probably one of the reasons for the creation of P.C. Cadix's branch station in Algiers, known as "Kouba". Maj. Maksymilian Ciężki, deputy commander of L'équipe Z, was spending in 1941-1942 most of his time at "Kouba", combining the functions of its commander with that of "Rygor's" signals officer. We may only guess that the presence of the Polish codebreakers in Algiers had one more, and very important reason. Marian Rejewski and his colleagues attempted to continue their work on Enigma ciphers, but were suffering from the lack of the cipher material. French interception service, *Groupement des contrôles radioélectriques (GCR)*, delivering cipher material to P.C. Cadix, did not intercept Enigma traffic due to the technical problems, and, as we shall learn later, different priorities of its work. Poles could continue their work on Enigma only taking care of interception themselves. Shifting their wireless sets to Algiers, closer to the battlefields of Libya and Egypt, they could hope to provide their codebreakers with the current cipher material. We will probably never know how Różycki and Zygalski (who were usually manning the Algiers outpost) managed to break Enigma keys during 1941 and 1942, without access to equipment used at that time at Bletchley Park. All we know from Bertrand's later book is that they managed to read around 4.000 Enigma messages, including several of high importance. Poles were able to break the Enigma keys used by the German forces fighting in the eastern front. Some of them represented the earliest information about the atrocities committed by the Einsatzgruppen following the Wehrmacht in Russia. End of their adventure with Enigma was marked on 9 January 1942by the catastrophe of the passenger liner *Lamoricière*, taking the lives of three Polish codebreakers, including Jerzy Różycki, and two copies of Enigma machine returning with them to France.

While Różycki was working in Algiers on Enigma messages, his colleagues in P.C. Cadix were attacking other German ciphers. In his memories Rejewski describes two ciphers being solved by the Polish team. Antoni Palluth, nominally the team's engineer, and not the codebreaker, focused his attention on the first one, representing a columnar transposition with blanks. For the reasons we shall cover later, his solutions must have been extremely precious for the French hosts. This cipher was in use by the German agents working in the unoccupied part of France. Rejewski describes one of Palluth's solutions leading to the discovery of agents' secret meeting in a hotel at Marseille. Their summary arrest was facilitated by the fact, that all of them appeared at the meeting place carrying their radio sets in identical suitcases.

The other cipher represented a standard Playfair, so its breaking was a rather routine job, but at some time its solution proved to be of a critical importance for entire team of P.C. Cadix. One of the most spectacular achievements of the

French service was tapping the telegraph lines used to coordinate the network of German radio monitoring stations. Those stations used to exchange via teletype information on the underground transmitters, their working frequencies and fixed positions. Using the cables believed to be under German control, network stations did not bother to use a more secure cipher. Tapping the cables by the French underground, and breaking the intercepted messages by the Polish codebreakers must have saved a number of Allied wireless operators, whose sets were mentioned as *aushebereif* (ready to be captured). In one of the most spectacular twists of the history, in September and October 1942 messages broken by the Polish codebreakers started to contain references to their own radio station. Finally, in late October two German cars with visible direction-finding antennas appeared in the close vicinity of the villa hosting P.C. Cadix. Officers commanding the center were informed about coming Allied landings in North Africa and realized the danger of German occupation of Vichy France; it was obvious that the existence and activities of P.C. Cadix are coming to an end.

Events of next few months represent one of the most mysterious periods in the Enigma history. The effects of actions undertaken by the participants of the events diverge from their declared intentions to an extent, that suggests caution in their interpretation. Before we compare these declarations with several new sources, external to the participants of the events, let us recall the story in the version resulting from the accounts of their direct actors.

Communicating directly with London SIS HQ, Maj. Bertrand had been pre-warned about the danger to P.C. Cadix resulting from the planned Allied operations in North Africa. London was to signal the imminent start of the operation with a pre-agreed message: "the harvest is bountiful". The SIS and Polish intelligence headquarters assumed that this signal would immediately cause the evacuation of a group of Polish codebreakers out of the zone where they would be in danger of falling into German hands. British intelligence offered help in the evacuation aboard a submarine that could pick up the Poles from the French coast. The evacuation of General Giroud aboard a British submarine on November 4 confirms the possibility of such an operation. Polish intelligence networks in France also had secure evacuation routes that could be used to evacuate the team to Spain and then via Gibraltar to the UK. Both services were aware that the possible capture of any team member by the Germans represented a deadly threat not only directly to him, but above all to the secret of breaking the Enigma cipher.

In October 1942 Bertrand and Langer decided to pay a visit together to the head of Polish intelligence network “F”, to investigate its potential role in the evacuation of the Polish team. It is difficult to understand the need for this visit. More or less at the same time Bertrand had decided to use means controlled by his own service to evacuate to Algiers the Spanish codebreakers, members of “l'équipe D”. Their evacuation must have been smooth, as they were soon able to continue their work in North Africa, this time under American control. When asked by Langer why not to use the same route to evacuate the Polish team, Bertrand considered this option too risky, as North Africa, in his view, would soon turn into a battlefield.

During Bertrand’s absence at P.C. Cadix, an expected message from London was received, but not read, having been enciphered using Bertrand’s private key. It was only after his return that he deciphered the text informing that the “harvest is already very bountiful”. Finally, Bertrand's dilemmas were resolved by the previously described visit of German cars with direction finding equipment. Over the next few hours crucial equipment was hidden in previously prepared caches in the walls, center’s archive was secured in Bertrand’s mother’s house at Grasse, the villa had been abandoned and the entire team of P.C Cadix went into hiding, preparing for the evacuation from the danger zone. On a personal level, the same danger applied to all members of the organization within which the center operated. In his memoirs Bertrand (1973) suggests that he expected his superiors to organize the evacuation of the team. The time was ripe for a decision, and the Germans provided a few days for its implementation. “Fall Anton”, German and Italian occupation of Vichy France, started only in the evening of 10 November, three days after Allied landings in North Africa. Therefore, Bertrand suggests to be disappointed after his superiors departed for Africa from the air base at Istres, leaving behind him and entire team. His disappointment was all the greater because the team of P.C. Cadix could reach the airfield in about an hour, and there were still plenty of empty seats on board the flights to Africa.

Immediately after the German/Italian occupation of Vichy France Bertrand transferred Èquipe Z to Italian zone, considering it safer. At the same time, he divided the entire team into smaller groups, which made it easier to find hiding places and reduced the risk of exposure. A side effect, however, was making the Polish team completely dependent on Bertrand's concepts and actions. And these changed frequently. For some time, the plan was to evacuate the Poles on board of a submarine, which was to pick them up from one of the secluded bays on the Côte d'Azur. This plan failed, reportedly due to the planned embarkation point being manned by Italian troops. Next, the possibility of evacuation via Switzerland was analyzed. However, this would entail the inevitable internment of Poles, which did not generate enthusiasm either on the part of those concerned or the British, with whom Bertrand consulted this option. As the days and weeks passed, the crisis of trust between Bertrand and the officers commanding the Polish team deepened. The Poles accused Bertrand of sabotaging the evacuation of their team and thereby exposing the secret of Enigma. Bertrand responded accusing Poles of alcohol abuse and rejecting his subsequent concepts without justification. Ultimately, both sides agreed that the only viable, albeit risky, option was to cross the green border to Spain, from where Poles would be evacuated to Great Britain via Red Cross channels. From the distance of time, it is hard to tell whether the news that Bertrand would not be accompanying them on the road, caused more anxiety or relief. To supervise the operation Bertrand designated his deputy from P.C. Cadix, certain Captain Louis (whose talents, as later recollections show, he did not highly value).

Poles, still divided into smaller groups, were moved near the Spanish border, to territory controlled by German troops. There, they waited for the possibility of crossing the border in the hideouts in Toulouse, Perpignan and Narbonne. Two groups, consisting mainly of the codebreakers, managed to cross the border. In particular, Rejewski and Zygalski made it to Spain on the night of January 29/30, 1943. And although on the way they were thoroughly robbed by the guide, and on the Spanish side they were arrested by the gendarmerie, they were safe. The fate of group led by Maj. Wiktor Michałowski was similar.

The fate of the officers, commanders of the Polish team, was different. A group including Col. Gwido Langer and Maj. Maksymilian Cieszki, on the first attempt to cross the border, was stopped by the French gendarmerie, and the officers were imprisoned in the Perpignan citadel. The French unofficially expressed regret over the incident, implying that they had not received the customary in this situation warning about the need to turn a blind eye to a group of travelers. The sympathetic French gendarme also noted that certain Monsieur Gomez, the smuggler whom the French handlers had hired for the task, was suspected of collaborating with the Germans. Finally, after paying the necessary bribes, the Poles were released and in March they were able to make another attempt to cross the border.

In the circumstances described above it was natural for the officers, once they landed again in the border zone, to try to avoid contact with Monsieur Gomez, who botched the previous attempt. Following the suggestion of a French gendarme, they agreed terms of service with another smuggler group. However, on the eve of the planned border crossing, a local representative of Monsieur Gomez found them and forced them back into cooperation. Gomez must have sensed a lack of trust from his clients. To calm them down, he took out a bill, ordered Langer to sign it, then tore it halfway, keeping one half, and handing the other to the officer instructing him to give half of the note to the guide only after safely reaching Spain; the guide would be paid only after presenting the correct half of the banknote. On the night of March 10-11 the escapees had covered only a few kilometers when they were surrounded by a group of German soldiers on motorcycles. All were detained, except for the guide, who was allowed to leave unhindered by the Germans.

2 Greater picture in the context

Functioning, organization, and activities of P.C. Cadix were so far analyzed mostly from the point of view of inter-allied cooperation in the codebreaking. This point of view does not permit to explain and clarify ambiguities or even contradictions in the available sources. However, there exists an alternative point of view that has not been sufficiently exploited by historians of the subject so far. P.C. Cadix represented a part of the secret services of Vichy France, whose functioning was, and is, the subject somewhat overlooked by historians. Their activities, however, left important source materials that permit to place the activities of Polish codebreakers in the appropriate context and find answers to questions that have caused problems in previous attempts to analyze the subject.

This paper represents an attempt to place the known facts in a new context and to offer the reader the resulting conclusions. It is focused on two particular questions; P.C. Cadix's scope of activity and the circumstances of codebreakers' evacuation after German/Italian occupation of Vichy France.

2.1 P.C. Cadix - scope of activity

The mainstream of Enigma history research treats the operation of the P.C. Cadix as a natural continuation of the activities of the P.C. Bruno, under slightly more complicated conditions. In this version, Polish codebreakers were to continue breaking Enigma, with significantly lower efficiency than before, due to the lack of technical equipment. In fact, breaking Enigma ciphers was only a marginal part of their activity during this period. The lack of technical equipment was, of course, an important factor, but it was the completely different priorities set for the center by the French principals that were decisive.

P.C. Cadix represented a part of the structure of the French organization Travaux Ruraux (TR), one of the secret services of Vichy France disguised as a commercial company. TR was created on 1 July 1940, by Capt. Paul Paillole, before the defeat of France the adjutant of the head of the French military counterintelligence, Lt Col. Guy Schlesser (Paillole, undated). In accordance with the previous experience of its founder, TR was an organization of a clearly counter-intelligence character. In this role, it enjoyed a significant level of autonomy. Bureau des Menées anti-Nationales (BMA), the structure, which was supposed to act as the supervisor of Vichy France's special services, was not established until August 25, 1940. From the outbreak of war to the defeat of France, the 5ème Bureau (Col. Louis Rivet) provided some level of coordination of the intelligence activities of the services. After the defeat of France, this structure was not recreated, so the counterintelligence, as well as the intelligence services of the army, navy and air force, regained almost complete independence[1].

As a result of the described changes, the codebreakers' team, previously subordinated to the command of the French intelligence service, was transferred entirely to the domain of counterintelligence. Enigma messages the codebreakers had been previously providing were of marginal importance to their current supervisors. TR's main task was to identify and neutralize agents of foreign powers operating in territories subordinated to the Vichy authorities. Neither the agents, nor their handlers were enciphering their messages with Enigma, so the main asset of the Polish team lost its significance in the eyes of their present French superiors.

Continuation of work on Enigma was also difficult due to the lack of cipher material. Before the defeat of France, the signals intelligence services of France and Great Britain kept on exchanging intercepted messages. The defeat of France disorganized her intercept service and cut off the codebreakers from British sources. From Rejewski's memoirs (1967), we know that Bertrand tried to keep Poles busy by delivering a package of intercepted Swiss messages, encrypted with a commercial Enigma (which the Poles easily broke). In the meantime, the French attempted to rebuild their own intercept service, which was not a simple task given the limitations of the Armistice Treaty. Groupement des Contrôles Radioélectriques (GCR) started its operation on 10 August 1940 at the Château des Cours in Hauterive. GCR, organization created and led by Gabriel Romon, formally represented a part of the French Post Office, less formally providing valuable information to Vichy secret services. Its early activities are poorly documented; the earliest GCR documents preserved in French archives date back to 1942 (GCR). Fortunately, its history has been largely reconstructed by the son of its founder, François Romon, and described in his book (2017). Romon suggests that '*German Enigma messages deciphered by Gustave Bertrand at P.C. Cadix mostly came from GCR's intercepts*'. Cooperation between GCR and P.C. Cadix was greatly facilitated by the fact, that one of the GCR's

[1] Second Bureau de l'Armee d'Armistice, again under Rivet, served as an administrative rather that operational hub for the services.

interception centers was located in the vicinity of Uzès, at Bouillargues. However, Romon's assumption is not confirmed in Rejewski's memoirs. He mentioned that the Polish team had to dedicate two out of four radio sets and several operators only to intercept Enigma messages. Romon's claim may refer to the work of Polish codebreakers in Algiers, where they had at their disposal only one radio set and had to rely on the French intercepts. On the other hand, Romon himself points out that '*GCR's intercepts concerned mainly the messages of German agents in non-occupied zone. Content of some messages, among the others, permitted to discover and annihilate their entire network, comprising 10 agents and 6 radio sets*'. Romon refers clearly to an episode described also by Rejewski, when the messages broken by Antoni Palluth (who specialized in double transposition cipher used by the agents) permitted to surprise a group of agents during their meeting in a hotel in Marseille. Paillole estimated that between October 1940 and November 1942 TR was able to eliminate over 1.000 enemy agents (Paillole, *Chronologie*). We will never know how many of them had been identified by the messages broken by P.C. Cadix. Moreover, Paillole's definition of the "enemy agent" had been a bit fuzzy… Anyway, it is clear that TR's priority was not intercepting and breaking Enigma, but the hand ciphers used by the agents working in Vichy France.

The second area of activity of GCR, and consequently Polish codebreakers, was the interception and decrypting of German teletype communications. The operation, which resulted in the acquisition of the cipher material, referred to within the GCR as Source K, began in October 1941. In March 1942, the French rented a house located on the cable route connecting Paris with Metz. On April 18, they simulated a cable failure, and during the "repair" tapped the lines used by the Germans between Paris and Berlin. With one exception, no data is available on how this source of information was used. This only exception are messages exchanged between German signals intelligence stations cooperating in the direction finding of Allied underground transmitters in the occupied countries. Since the messages were exchanged over a cable line that the Germans considered safe, they were secured by a relatively low-level cipher, i.e. Playfair. In this form, they were transferred to P.C. Cadix, where Polish codebreakers kept on breaking them without much difficulty. In his memoirs, Rejewski describes the satisfaction of being able to warn underground station operators against exposure. He did not record his own reaction, however, when the broken messages referenced the P.C. Cadix's own station, indicating impending danger. Later on the tap was discovered by the German security service, and the line was secured using machine cipher. Sadly, some members of the GCR were imprisoned and executed, including Gabriel Romon.

Summing up, breaking by P.C. Cadix of about 4,000 Enigma messages was an astonishing success of a Polish codebreakers' team, who did not have at their disposal any of the equipment used at the same time at Bletchley Park. However, it is equally obvious that breaking Enigma keys was merely a sideline of the team's activities. P.C. Cadix functioned as part of the French counterintelligence, whose priority was to decipher messages, on the one hand, enabling the identification of foreign agents in Vichy France, and on the other, protecting its own structures against identification by the enemy.

2.2 P.C. Cadix – mystery of evacuation

The evacuation of the Polish team from the south of France as a result of its occupation by the German and Italian forces represents one of the most mysterious episodes in the history of Enigma. Considering the facts presented in the previous section, during the period of work at P.C. Cadix, Polish codebreakers only marginally dealt with problems in which they had the greatest experience. The value of their work in the south of France to the Allied cause was highly debatable. The occupation of the south of the country by the Axis forces made it not so much worthless as dangerous. Their French superiors were fully aware of the impending threat: '*For some time Cdt Paillole knew the landing points* [of Allied troops in North Africa] *and the approximate date of the operation. He knew also that the German response presupposes a complete occupation of Vichy France*' (Paillole, *Résume*). Falling into German hands of any of the team members threatened to expose the secret of Enigma breaking. The safe evacuation of the group from the areas under German control became a priority for the special services of Great Britain, Poland and, with some hesitation, France.

During 1942, cooperation between London and P.C. Cadix was not so much continued as simulated only. When recommending the continued cooperation with Bertrand in March 1941, Wilfred Dunderdale, British liaison officer with the French underground, suggested to '*offer* […] *harmless stuff,* [while] *exploiting every opportunity of obtaining information*' (Jeffery, 2011). When founding the BMA, Col. Rivet strictly forbade his subordinate services any contacts with the Allies, making an exception for three officers only, including Bertrand (Paillole, *Résume,* p 14). Paillole still in 1942 represented an opinion that although Germany was '*enemy number one*', Britain was '*enemy number two*' (Kitson, 2008). In line with this mutual distrust, at the turn of 1941 and 1942 the British handed over to P.C. Cadix several keys to the Enigma cipher, taking care of the security of their own operation; they only shared several keys captured in North Africa, excluding those broken cryptanalytically (Borowiak, Grajek, 2022; note for CSS, April 9, 1942, TNA, HW 65/7).

The British made efforts to urgently evacuate the Polish team from France. However, there was little they could do but make the French aware of the importance of the matter and offer help. Quoting Paillole: '*on the evening of November 5, Colonel Rivet, head of the special services, called me urgently to inform me of the content of a dispatch from the IS, signaling the imminence of an Allied landing in North Africa. The dispatch also asked to expressly withdraw to Algiers the precious Polish decryption personnel of the PC "Cadix"*' (Paillole, 1975). On November 6, Rivet convened the final meeting of the body known as the Little Chancellery, composed of the heads of Vichy's secret services. He wanted to obtain information about the intentions of individual services in the face of the upcoming Vichy occupation. Paillole requested the navy's assistance in evacuating the TR's archives, weighing some 40 tons, waiting on the quay in Toulon. Chief of Naval Intelligence, Capt. Sanson, after consulting his superiors, conveyed their answer stating that '*no, there is no reason why those archives should fall into the hands of the British any more than into those of the Germans.*' (Paillole, *Résume...*). The climate of those days was clearly not conducive to cooperation with former (and future) allies.

The complexity of the situation is compounded by the fact that an Allied operation caught the Vichy special services in the process of reorganization. The Germans, having a good understanding of the actions of the French, forced the Vichy government to dissolve the BMA and its subordinate agencies as early as April 1942. The French delayed the implementation of their orders, but the takeover of the reins of government by the pro-German Laval precluded ignoring them for a long time. In August 1942, the BMA was disbanded and replaced by the Service de Sécurité Militaire (SSM), under the command of ... Paillol. At the end of September, the new SSM chief organized a secret meeting of the commanders of intelligence services of the army, navy and air force. The subjects discussed included plans for action in the event of the German occupation of Vichy. It was established that in such a case TR structures would stay in the occupied territories and continue its counterintelligence activities.

Despite this conclusion, on November 9, just after the first news about the Allies' landings, '*Bonnefous is sent to Istres to find out about possible departures. He returns on 10th at 10 p.m.: the disorder and congestion on the airfield are such that a quick connection to and from Algiers is unimaginable.*' (Paillole, *Résume*). The same chaos and congestion did not hinder the planned departure to Algiers of the entire team of the air force intelligence service, with its commander, Colonel Ronin. It was to this episode that Bertrand referred to in his memoirs, when he expressed his disappointment about the order of his superiors, that did not allow him to place a Polish team on board as well. Bertrand's unidentified superior was to decide that only officers, presumably French, were to be evacuated (*les officiers d'abord*). In fact, the evacuation covered a slightly wider group: '*Colonel Ronin succeeded, when the time came, in transferring Air Intelligence en bloc by flying it from Istres to Algiers. With him, incidentally, he took other members of the Special Services who were particularly sought after by the Germans*' (Stead 1959).

Bertrand was undoubtedly in a predicament. On the one hand, he was aware of the need to evacuate the codebreakers and knew about the request of the British. On the other, the order to stay in France along with other TR structures concerned not only him personally, but also the team he led. The chaos accompanying the evacuation of the Vichy services to Africa probably meant that there was no one to take care of giving Bertrand instructions to make an exception for the Poles. On a strictly personal level, Bertrand must have also been in trouble.

After the Spanish team had been evacuated to Africa, the Poles were his *raison d'être* in the structures of the French special services. To let them evacuate meant to lose his position in the only environment that mattered to him. Anyone who knew Bertrand, either directly or through what he did or said, could not doubt his decision: he let events take their own course.

After the opportunity to evacuate the team to Africa by air before and after the Allied landings in Africa had been wasted; as the period of poor surveillance of the coast just after the occupation had been missed, the situation became significantly more complicated. The heads of the Polish team pressed for evacuation, and its members were waiting, dispersed in the Italian occupation zone. Bertrand signaled the impending evacuation aboard the submarine several times, but each time the operation was canceled at the last minute.

It is interesting to learn why the evacuation by submarine, vigorously rejected in the period preceding German occupation of Vichy France, was accepted now? The answer seems to be very simple; this time it was supposed to be a French, and not British submarine. The descriptions of the evacuation plans in Bertrand's and Rejewski's memoirs correspond exactly to the fragments of Paillole's report, presenting the operations he was planning in Algiers with the participation of the submarine *Casabianca*, one of the French ships that managed to leave Toulon. *Casabianca* under command of Capitaine Jean L'Herminier was to deliver men and equipment for the French underground to the landing site between Cannes and Nice, and pick up evacuees from there. However, the original plans had to be delayed and modified: '*two TR officers reconnoitered the point selected: it was occupied by Italian troops and the whole coast between Cannes and Nice was bristling with defense-works and guard posts*' (Stead 1959).

It will be understandable for the reader to ask how Paillole could plan these operations in Algiers, since we said goodbye to him in France, where he declared to remain in the occupied country? Quoting his own report: '*on 17* [November] *in the morning, departure for the Pyrenees. While the details of the clandestine border crossing are worked out by the local TR post, Paillole goes to Toulouse*'. And then: '*on the night of November 28 to 29, 1942 the border is crossed at Puigcerda in the company of Villeneuve and Poniatowski*' (Paillole, *Résume*). From (Stead, 1959) we learn more details of his escape: '*travelers dined at an inn at La-Tour-de-Carol (...), they slipped into a friend's house 100 yards from the frontier. A guide was waiting for them and he led them without undue incident to Puigcerda*'. This evacuation route was managed by Ramonatxo brothers from Perpignan, and its functioning was presented in the book written by one of them (Ramonatxo, 1955). From Puigcerda Paillole got to Barcelona, Gibraltar, and further on through London to Algiers, to secure his position in the power struggle between the supporters of General de Gaulle and General Giroud.

In his later book Paillol clearly refers to the evacuation of the Polish team and the route he had used himself: '*Poles, left to their own devices, resorted, in circumstances unknown to me, to a network of Pyrenean smugglers controlled by the enemy. It is a serious fault of our house which had absolutely safe channels of passage to Spain'* (Paillole, 1959). This statement does not fully reflect the facts. The details of the evacuation described in Rejewski's memoirs fully correspond to the realities of the route covered also by Paillol. As far as we know, the team under the command of Wiktor Michałowski followed the same trail. Due to Antoni Palluth's death in a concentration camp, we will never know the circumstances of his, and his companions' capture. The controversy mainly concerns the circumstances of the capture of the group including both officers commanding the team of Polish codebreakers.

According to Bertrand's (and Paillole's) account, after the fiasco of the first attempt, the Poles rejected the guides provided by TR and were betrayed by members of the other smuggler group they had chosen on their own. Langer and Ciężki's version has been reported in the last sentences of section 1 above. Which version is closer to the truth is decided by a detail in Bertrand's later book and Langer's post-war memories. Bertrand tried to put the blame for the failure of the evacuation on Langer, arguing that the Poles had rejected the services of smugglers provided by his organization. At the same time, both in his book and in a post-war conversation with Langer, he confirmed having paid the smugglers' price after they delivered the other half of the bill signed by Langer (confiscated by Germans after the officer was captured). This lie, rather unprofessional for an intelligence service officer, confirms that Langer and his companions were led into the ambush by a guide hired by Bertrand, or by someone supervising their evacuation on behalf of TR.

In the described circumstances, the natural question is whether the failed evacuation of a significant part of the Polish team was caused (according to Paillole's assessment) by tragic negligence on the part of the French services, or was it a conscious settlement of scores with a partner who was no longer useful? '*Thou shalt not kill, but need'st not strive officiously to keep alive*'[2]. Ciężki's aggressive attitude towards his French fellow prisoners during his internment at the Eisenberg castle and the unambiguous note in Langer's post-war letter to his wife (at the author's disposal) clearly confirm that they both pointed, not surprisingly, to the latter possibility. The fact that the capture of Langer and Ciężki did not lead to the disclosure of the Enigma secret, the Allies owe only to the tactics of both officers, adopted during the interrogations in March 1944 (Grajek 2017).

Somewhat paradoxically, if Bertrand intentionally or negligently handed over the heads of the Polish team into the German hands, he had the right to assume that he was not exposing the Enigma secret itself. When in an earlier exchange with London he insisted that the current cipher keys be sent to P.C. Cadix, Denniston instructed SIS to reply that this was impossible, as BP was currently unable to recover them. As a consequence, Bertrand may have believed that he was only jeopardizing his disliked partners, but not a secret vital to the Allied cause.

3 Appendix

The description of the difficult relationship between Bertrand and the group of Polish codebreakers is a good opportunity to draw attention to its aspect with long-term consequences. It is natural to ask why haven't been Rejewski and Zygalski invited to join Bletchley Park after their arrival to London in mid-1943? According to one of the possible explanations the secret of Enigma could not be entrusted to people arriving from territory controlled by the enemy. Another explanation was that the British counter-intelligence believed Bertrand's organization to be infiltrated or even inspired by the Germans. The sum of information originating from two independent and unrelated sources might indicate that such speculations were not completely unfounded.

In his memoirs, Bertrand describes his contacts with an unidentified representative of the German embassy in occupied Paris, who was willing to share some useful information with the French underground. He referred to his interlocutor as "Max". In his memoirs, Oskar Reile, head of the Abwehr's post in Paris, mentions an attempt to infiltrate the French underground through two Germans publicly demonstrating their disillusionment with their own country's policies. He refers to his honeypots as "Max" and "Moritz" (Reile, p. 216). If "Max" from the memoirs of both officers represented the same person, which is probable, Reile was dangerously close to discovering the Enigma secret, and was prevented from the success only by Allied landings in Africa, which forced Bertrand into hiding.

4 Summary

Functioning of P.C. Cadix was an exemplification of the complicated British-French-Polish relations after the defeat of France. None of the partners involved appears to have been fully sincere, and acting in good faith. The British simulated cooperation with Bertrand, trying to prevent the transfer of the Enigma secret into German hands. Bertrand was given by his Vichy supervisors unprecedented permission to contact SIS not to spy for the British, but on them. Polish HQ tried not so much keep a group of codebreakers in France, as simply away from London.

Work carried out in such conditions (not to mention the lack of technical equipment) could not bring spectacular results. The prospects for evacuation in the face of Allied landings in Africa were not much better. In a lecture given to a group of TR officers on June 6, 1942, Paillole described the priorities of his service as follows: '*Germany is danger number one* [and] *England is danger number two*' (SHAT 1942). The safe evacuation of the Polish codebreakers to London was in conflict with both the political priorities of the service that was supposed to assure it and the personal interests of Bertrand, who was supposed to oversee it directly. In the described conditions, the successful evacuation of at least the

[2] Arthur Hugh Clough, The Latest Decalogue

professional core of the group should be considered a success of the operation.

References

Bertrand Gustave, *Enigma, ou la Plus Grande Enigme de la Guerre 1939-1945*, Plon, Paris 1973

Borowiak Mariusz, Grajek Marek, *Enigma History and an Unexpected Treasure Trove*, Proceedings of the 5th International Conference on Historical Cryptology HistoCrypt 2022, Linköping University Electronic Press, Sweden 2022

GCR, Groupement des contrôles radioélectriques, Archives nationales (France), 65AJ/1-65AJ/1282

Grajek Marek, *Interrogation at Eisenberg castle. How two Polish officers saved the Ultra secret just before Overlord*, European Historical Ciphers Colloquium, Smolenice 2017

Jeffery Keith, *MI6: The History of the Secret Intelligence Service*, 1909-1949, London 2011

Kitson Simon, *The Hunt for Nazi Spies, Fighting Espionage in Vichy France*, Chicago & London, University of Chicago Press, 2008, p 69

Paillole Paul, *Chronologie de l'activité du 5ème Bureau de l'Armée (1940-1944).* Sans date, Archives nationales (France), Archives du Comité d'histoire de la Deuxième Guerre mondiale, Services spéciaux — Activités du commandant Paul Paillole, 72AJ/82 Dossier n° 4

Paillole Paul, *Organisation du 5ème Bureau et des Organes Derives*, op. cit.

Paillole Paul, *Résume de l'Action des Services de Contre-Espionnage militaire français de juillet 1940 à novembre 1944*, October 1946, op. cit.

Paillole Paul, *Services spéciaux : 1935-1945*, Éditions Robert Laffont Paris 1975, p. 397-398

Ramonatxo Hector, *Ils ont franchi les Pyrénées...*, La Plume d'Or, 1955

Reile Oskar, *Der Deutsche Geheimdienst im II. Weltkrieg. Westfront*, Weltbild Verlag, Augsburg 1989

Rejewski Marian, *Memories of my work at the Cipher Bureau of the General Staff Second Department 1930–1945*, Wyd. UAM, Poznań 2011

Romon François, *Les écoutes radio dans la Résistance française 1940-1945*, Noveau Monde, 2017

Service historique de l'armée de terre (SHAT), Fonds de Moscou, 784/381, conférence du commandant Paillole de 6 June 1942

Stead Philip John, *Second Bureau*, Evans Brothers, London 1959

The History of the Development and the Analysis of the Cipher Machine T-310/50 and the Procedure ARGON by the ZCO

Wolfgang Killmann
Neuenhagen, Germany
wkillmann@gmx.de

Abstract

This paper describes important aspects of the 10 years of development and the 18 years of security analysis of the cipher machine T-310/50 and the procedure ARGON by the Central Cipher Authority of the GDR. The threat model of the analysis is pictured. Examples of the security analysis of the cipher algorithm, machine, procedure and key management are provided. The focus is on the analysis of the operating functions of the T-310/50 (operating analysis), including the analysis of operating errors, as well as the analysis of the instruction manual for the cipher procedure ARGON. The possibilities of obtaining information by an attacker from traffic reconnaissance in general and decryption attacks in particular are assessed.

1 Introduction

The cipher T-310 was developed by the Central Cipher Authority ("Zentrales Chiffrierorgan", ZCO) of the German Democratic Republic (GDR) in the 1970s. It was used widely for the protection of teletype communication up to security level secret ("Geheime Verschlusssache") in the 1980s.[1] By the end of the GDR almost all cipher machines T-310/50 were destroyed in 1990. The last use case of the cipher T-310 was for secure governmental communication between GDR and German Federal Republic (Stephan, 2022). The cipher T-310 was kept secret until October 2003.

The development and the analysis of all components and procedures were closely linked together in order to ensure the security of the cipher T-310. The ZCO developed the algorithm together with all long-term keys and the guidance documentation. It also produced the short-term keys for the distribution by the cipher services and their specialists analyzed the cipher and the operational usage of the cipher T-310 over its whole life time. The industry produced the T-310/50 machine.

Today, everyone can see what the T-310 looks like and how it works. It is now on display at the museums, e.g. NVA Museum Harnekop and the Deutsches Museum München. A software simulation of the T-310/50 is available on http://scz.bplaced.net/freeware.html. The algorithm T-310 is demonstrated by the open-source software CrypTool 2 available on https://www.cryptool.org/en/ct2/downloads.

To understand the reasons and details why it works this paper highlights important aspects of the ten years development and the eighteen years of security analysis by the ZCO. Section 2 provides a short historical overview of the development of cipher T-310. Section 3 describes the general threat model used for the security analysis of the T-310. The algorithm (Section 4), the machine (Section 5), and the procedure (Section 6) build together with the overlapping key management (Section 7) bottom-up the cipher.

They are addressed by separated, but dependent hierarchical aspects of the security analysis. The structure of this paper follows this layered approach.

We describe the cipher model of T-310 for the systematic of the security analysis. The *cipher* comprises all regulations and means for the encryption and the corresponding security functions including the key management. The *keys* are variable parts of the cipher. The cipher T-310 uses long-term keys implemented by circuit boards in the machine and short-term keys on punch cards. The *algorithm* is the mathematical model of the encryption and the data authentication. The algorithm T-310 is implemented in hardware and pro-

[1] There were four security level: This is the second hightest.

vides only the encryption (Section 4). The *cryptographic module* implements the algorithms and possibly other security functions (e.g. generation of random numbers) in hardware or software as part of a dedicated *cipher machine*, in our case the T-310/50. The *cipher material* comprises cipher machine and the key material. The *cipher procedure* defines how to use the cipher material by an human operator or by other devices. The T-310 cipher procedure is named ARGON (Section 6). The cipher includes also the *key management*, i.e. the generation of keys, the production and the distribution of key materials (not shown in Figure 1) and the key handling by the crypto officer in the exclusion zone.

2 Time Line of T-310 Development

The ZCO was built in 1951. It developed manual ciphers in the 1950s and cipher machines with the VERNAM cipher in the 1960s. The GDR cipher services used Soviet cipher machines with internal keys, i.e. the key is shorter than all the text encrypted withthis key. In the 1970s the ZCO developed their first own ciphers with internal keys: the cipher SKS for the signal command system SKS V/1 and the teletype cipher T-310.

The first version of the teletype cipher T-310 (code name PUMA) was developed by ZCO from 1973 to 1975. It was designed for encryption of texts on five-hole and eight-hole punched tapes. In 1974 two ZCO cryptographers developed two new algorithms: a linear recursion for the initial vector with prime length of the period and a substitution algorithm for five bit and eight bit codes. In addition, a new technical base was available (e.g. new TTL chips; 1 bit updated to 4 bit). This allowed for cryptographic improvements (e.g. longer short-term key, bigger internal state). The development of the cipher T-310 started with the tactical-technical requirements for a machine T-310 in 1974. The machine shall work with teleprinter and data communication devices in stationary and mobile stations. In 1977 the A-phase of the development ("A" stands for "applied research") was finished and an A-prototype was available for trials. It comes out that the functionality and the complexity of the machine must be reduced (e.g. only five-bit code encryption, doubling of the complication unit on three plates instead of four) in 1978. The algorithm of the machine was fixed in 1979. A cryptographic analysis of T-310/50 was produced in 1980 (ZCO, 1980). It followed extensive trials of the procedure ARGON with the K5-prototypes of machine T-310/50 ("K" stands for "development and launch of products"). The modified 50 K5-prototypes were used as T-310/51 with procedure SAGA for transmission of tactical reconnaissance data by the GDR navy. The use of ARGON with mass-produced T-310/50 began in 1982. There were as many as $3,835$ cipher machines T-310 in active service by the GDR government, army, security services and political organizations.

The T-310/50 was the last cipher machine made up of small-scale integrated circuits in transistor–transistor logic (TTL). Only the optional code converter of T-310/50 used a microprocessor. The maintenance service used a special computer "Prüfrechner PR310/2" for functional checks of T-310/50 and fault finding in 1985. Telex was the standard form of text communication in governmental networks of the GDR in the 1970s and 1980s. The T-310/50 was used also for slow data transmission with "teletype modems". The next generation of ZCO cipher machines were dedicated for encryption for data storage and transmission (e.g. T-325/POLLUX) including PCM30-base systems for speech (T-311/SELEN) (Drobick, 2023). The new machines used microcomputers at least for the control of the cryptographic module.

The ZCO analyzed the security of the algorithm T-310, the machine T-310/50 and the procedure ARGON until the end of their use in 1990. The ZCO starts the development of the algorithm T-310 with two mathematicians in 1973. The core of algorithm T-310 was derived from the algorithm SKS. The group of cryptologists working on the algorithms T-310 grows to about ten mathematicians in the late 1970s. Additional fifteen cryptologists worked on security analysis of the machines T-310/50 and T-310/51, the procedures and the oversee of the technical development. The analysts were supported by two programmers, engineers, technicians and other staff members of the ZCO building and running special programs and devices, providing literature and other services.

3 General Security Model

The starting point of the security analysis is the security model. The security model describes the threats to the assets (threat model) to be mitigated

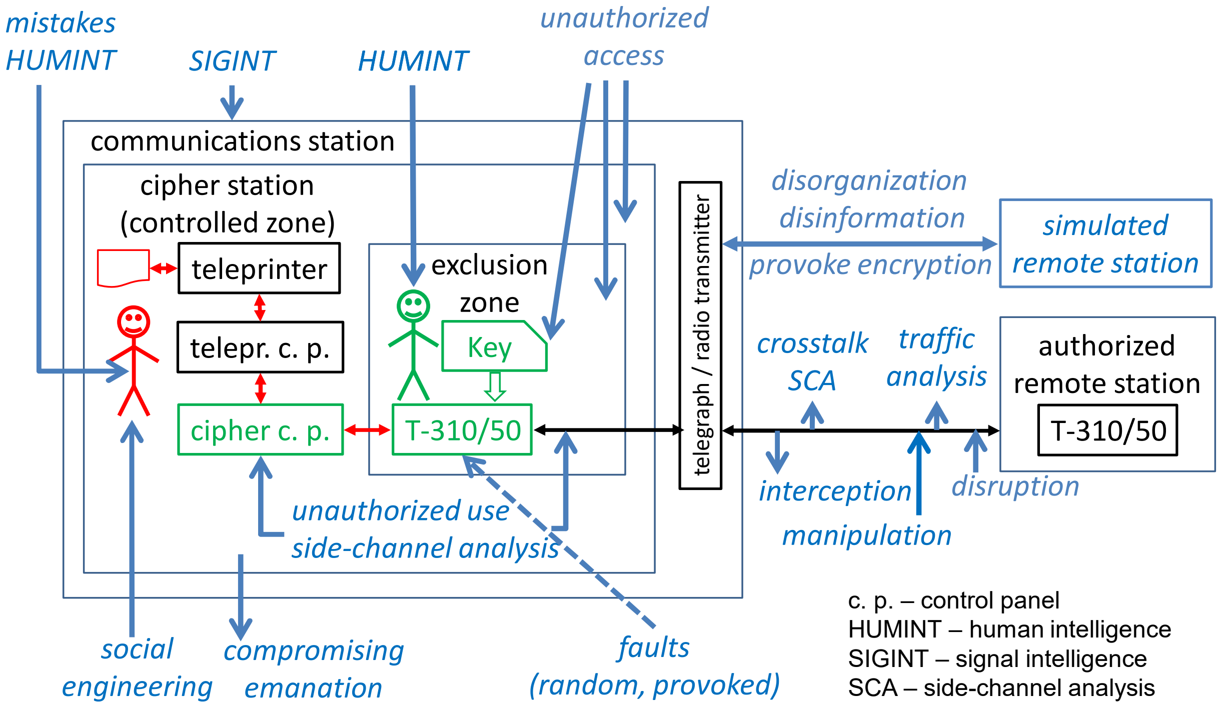


Figure 1: Threat model for the cipher machine T-310/50

by the security measures (cipher model) according to the enforced security policy. The security analysis needs a comprehensive and detailed understanding of the threats in order assess the effectiveness and the security of the cipher and if necessary of the additional countermeasures.

Figure 1 illustrates the threat model used for the cipher T-310. The primary asset to be protected are the state secrets represented in the plaintexts. The security policy requires the protection of the secrecy, integrity and availability of the state secrets. The cipher T-310 (drawn in green) was designed to ensure the secrecy of plaintexts (drawn in red) by encryption into ciphertext (drawn in black).

The operator transmits and receives non-classified and classified texts by teleprinter using cipher machine T-310/50 and the cipher procedure ARGON. In case of encryption the knowledge of the ciphertexts gained by interception of the communication line is always assumed. The adversary may also disrupt or interfere with communication and even imitate authorized communication.

The key has the same value for the adversary as all plaintexts encrypted with this key. Therefore, the cryptanalysis and the countermeasures distinguish between attacks on the plaintexts and on the keys (Section 4). Any information about the keys, the plaintexts or the internal processes of the cipher machine support the cryptanalysis. Therefore, the attacker analyzes not only the intercepted ciphertext but also the compromising emanation, the transmitted signal, the traffic, and so forth (Section 5.3). In case of state secrets any information about the transmission is of interest as well, e.g. the direction, the time and the priority of the transmission, the length of the plaintexts etc. Because of the general rules of telex and the combination of plaintext and ciphertext in the message the T-310 cannot prevent traffic analysis. But the T-310 must not provide additional marks supporting the traffic analysis (Section 5.1).

As a rule, the cipher is applied correctly. Any aberrance by technical fault in the machine or by mistake by applying the procedure may result in weak encryption and may enable attacks. The analysis of the security impact of every possible aberrance is not traceable. Any aberrance shall be avoided as potential vulnerability. Section 5.2 describes self-protection against technical faults. Section 6.2 discusses the robustness of the procedures and some security measures against operational errors.

In case of state secrets the security analysis shall identify possibilities and indication of covert adversary actions by operators or other persons despite of the personnel and organizational security measures. Section 6.2 provides an example.

The security analysis of the cipher may start with the algorithm but shall comprise all components up to the cipher network.

4 The Cipher Algorithm T-310

The strength of the cipher algorithm is the absolute necessary condition for the strength of the cipher (but as we see later not the only one). The cipher algorithm T-310 is defined in (Killmann and Stephan, 2021; Killmann, 2023). The T-310 is a stream cipher as a symmetric encryption system combining a sequence of teletype characters with the keystream by one character at a time, using a invertible function (like XOR). The structure of the cryptographic module is depicted in Figure 2 (Killmann and Stephan, 2021).

The short-time key of 240 bits (including 10 parity bits) is stored on a punch card. The input unit repeats cyclically the key components $S1$ and $S2$ building the s-sequences (Figure 3). The initial vector of 61 bits is randomly generated for encryption and derived from the message for decryption. The synchronization unit generates with the initial vector the linear shift register sequence f with period length $2^{61}-1$ (which is a prime number). The complication unit generates the keystream a controlling the substitution. For each teletype character 10 out of 13 bits of the a-sequence are used. The substitution combines each teletype character of five bits with ten bits of the keystream a. The mappings of the substitution algorithm build a double transitive permutation group of the teletype characters.

The complication unit is the core of the algorithm T-310. It implements a nonlinear shift register with the transition function φ of the states U as depicted in Figure 3 (Killmann and Stephan, 2021).

The sequences s and f act as parameters of φ. The long-term key defines the structure of the mapping φ. It defines the selection and the permutation of the 27 bits of the internal state U as inputs of the Boolean functions, the 9 bits XOR-ed to the feedback and the place of the output bit α of the register U. After 127 internal clocks one output bit is read for the a-sequence.

The cryptographic strength of the complication unit is crucial against attacks on the short-term keys. It comes out that the long-term key is critical for cryptographic strength of the cipher. Although kept secret the long-term keys are assumed fixed and potentially known to the attacker (Section 7). The specialists of the ZCO analyzed deeply the function φ depending on the long-term keys. The ZCO applied a procedure for the approval of long-term keys intended for application in the field (Killmann and Stephan, 2021; Killmann, 2023). These carefully selected long-term keys allowed to prove important security features of the cipher. If the parameters of φ are freely chosen, then the bijective φ generates a permutation group over the set of internal states U. The cryptologists proved that this group is transitive in the late 1970s. In the early 1980s they ensured that the group is the alternating group. These features prevent some simplified models of the complication unit. In the mid 1980s a new quality of results were reached in case the parameters of φ are deterministic sequences derived from the short-term key and the initial vector. The period of the a-sequence controlling the substitution is with high probability a multiple of the period of the f-sequence (i. e. $2.3 \cdot 10^{18}$). The specialists estimated the cardinality of equivalent keys as sufficient small depending on the long-term key.

The cryptologists of ZCO provided manual paper-and-pencil proofs and used computer programs for calculating specific long-term keys. They built special devices connected to a computer for time-consuming calculation. The special device T-032 were used for the calculation of cycles of the permutation φ since 1980. The special device T-037 generated internal sequences for statistical tests since 1982.

The substitution played a special role in protection of the plaintexts and the short-term keys too. If the keystream of a stream cipher with a simple invertible function is used more than once then the keystream may be withdrawn. Such attacks are well known for the VERNAM cipher and constitute a potential vulnerability of stream ciphers. If two ciphertexts are encrypted with the same keys and initial vector (disregarding of equivalences) the a-sequence can be determined by guessing the corresponding plaintexts. The a-sequence is the (necessary) base for attacks on the short-term key. It can also be used for encryption of another text imitating an authorized cipher station. If three such ciphertexts are intercepted and two corresponding plaintexts are known (or guessed) then the third plaintext may be determined independent on the strength of the keys and the complica-

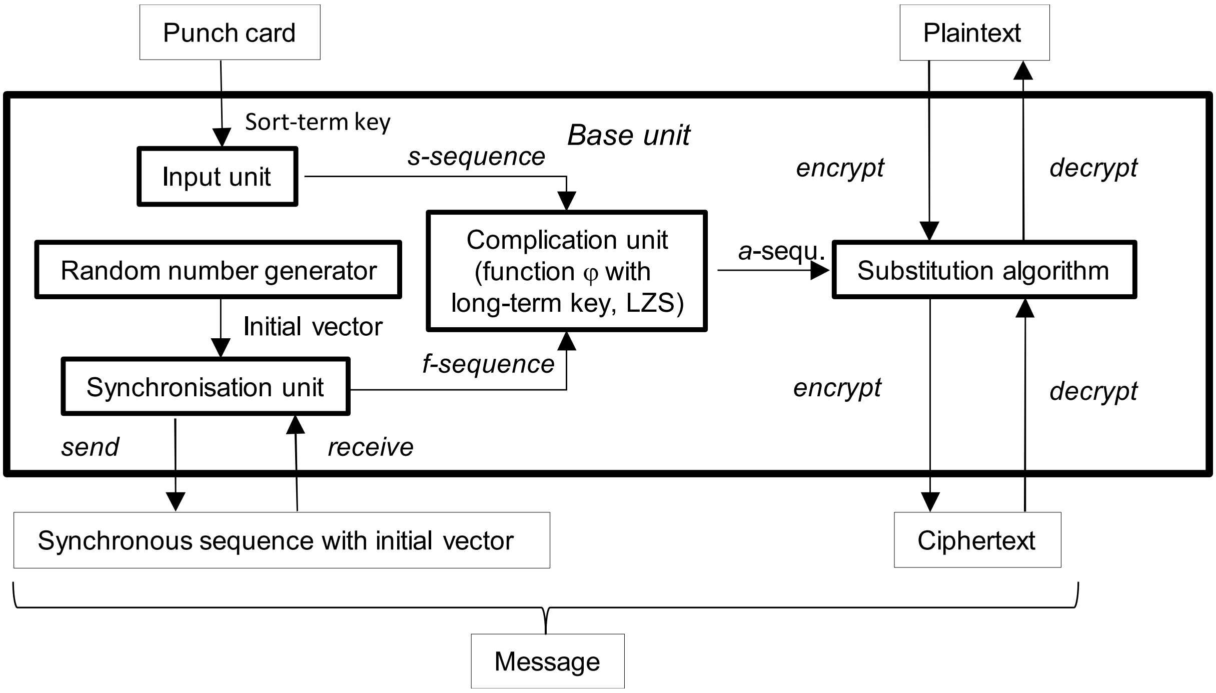


Figure 2: The cryptographic module

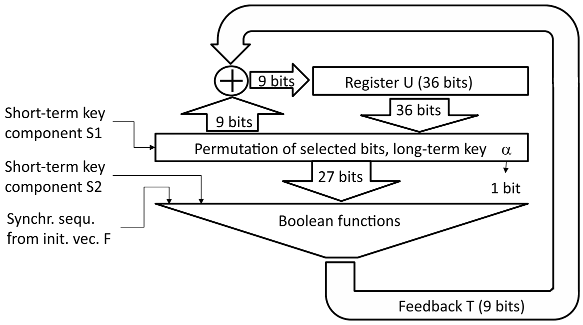

9 bits
Register U (36 bits)
Short-term key component S1
9 bits
36 bits
Permutation of selected bits, long-term key α
Short-term key component S2
27 bits
1 bit
Synchr. sequ. from init. vec. F
Boolean functions
Feedback T (9 bits)

Figure 3: Structure of the complication unit

tion unit (Killmann and Stephan, 2021; Stephan, 2022). The countermeasures against such attacks are (1) the generation of the initial vector by means of strong random number generators (Section 5.1), and (2) the robust cipher procedure preventing the encryption with an untrustworthy initial vector (Section 6.2).

The cryptologists of ZCO assessed the algorithm with approved long-term keys as secure and appropriate for encryption of state secrets.

According to Crypto Museum homepage[2], the US Army and the NATO widely used the SAVILLE cryptographic algorithm in high-level encryption devices including for teleprinter in the 1980s. SAVILLE is a stream cipher based on a nonlinear finite state machine, that has an internal cycle of several tens of iterations per output bit. The short-term keys consists of 128 bit key including 8 bit checksum (2^{120} keys), which is much less than T-310 short-term keys. The SAVILLE cryptographic algorithm is still secret. Thus, we cannot compare SAVILLE and T-310 in details. Some cryptologists are still interested in the strength of the algorithm T-310, even the machine T-310/50 has historical value only. The currently published attacks do not break the algorithm with approved long-term keys (Killmann, 2023).

5 The Cipher Machine T-310/50

The cipher machines T-310/50 were developed by the “Institut für Regelungstechnik” (IfR) and the ”VEB Steremat Berlin Hermann Schlimme“, and they were produced by the ”VEB Steremat Strausberg“. The ZCO as contracting authority defined the tactical-technical requirements, affirmed the specifications and the results of the development steps A and K, and the serial production of T-310/50. The developers provided a complete documentation (seven books from technical descriptions, circuit diagrams up to engineering drawings), A-prototypes, K-prototypes and serial-product machines. The T-310/50 consists of (1) the cipher control panel (green box in Figure 1) for the operator, (2) the basic unit implementing the encryption, (3) the power supply unit, and (4)

[2] https://www.cryptomuseum.com/crypto/usa/saville.htm

the cables connecting the panel and the units. It is connected between the teleprinter control panel ("Fernschaltgerät") and the communication line.

The cipher machine is much more complex than the abstract cipher algorithm. It shall implement the security functionality for encryption and decryption and may provide intended but security irrelevant functionality (like encoding the binary ciphertext to numbers). An (often hidden) functionality or feature of the cipher machine could also build a security vulnerability. The ZCO analyzed deeply the correctness and the security of the T-310/50. The following three examples illustrate the security analysis of the cipher machine T-310/50 performed by ZCO.

5.1 Random Number Generator

The generation of the initial vectors is an important aspect of the stream cipher T-310. From the pure algorithmic point of view the initial vectors must be different for all texts encrypted with the same short-term key (disregarding of equivalences). At first glance a deterministic generation of initial vectors would be possible e.g. by a linear shift register with a long period. But this approach requires random start vectors. The generation of initial vectors at random is an appropriate solution. The random number generator (RNG) is a non-algorithmic part of the cipher at the boundary between mathematics and engineering.

The RNG shall produce initial vectors of 61 bits for each message to be encrypted. The cipher machine of SKS V/1 implemented a physical random number generator. The A-protoype of T-310 used external input for the initial vector from teleprinter or punch cards in 1977. For the sake of reduction of the required punch cards and simplification of the machine the T-310/50 K-prototype used a non-physical true RNG[3] called System-RNG in 1978. The System-RNG used as noise sources (1) the clock for reading the punch card with the short-term key, (2) the time of the third step of the prophylactic check (Section 5.2), and (3) the telex characters read from the local and line interfaces. The System-RNG was analyzed by means of a comprehensive stochastic model and tests with special devices built by ZCO. Unfortunately the stochastic model could not be substantiated by sufficient amount of test data (e.g. reading of thousand punch cards). It was found that the evidence for the security of the System-RNG is sufficient for the procedure SAGA, but not for the procedure ARGON (ZCO, 1982). SAGA was used for only 50 clients and unilateral encryption from surveillance stations to the center. The System-RNG was secure for SAGA. In case of ARGON up to 150 client may communicate to each other sharing the same short-term key and generating initial vectors for each message. Therefore, the decision was made to equip the serial-production T-310/50 with a physically true RNG based on a transistor noise source in 1983. The final security analysis assessed the final RNG as secure in 1984.

The self-test of T-310/50 includes a health test of the physical RNG apparent during the prophylactic checks. The noise source is tested by counting the occurrences of a fixed pattern of six bits in its output. The health test is performed by the crypto officer when the punch card reader is switched on for input of the short-term key. When the pattern is detected 16 times then the H-OFF miniature lamp of the control panel is switched on or off. The decision can be made on the health of the noise source: (1) if the H-OFF does not blink then the noise source is broken and the machine is blocked (or the machine is T-310/51), (2) if the H-OFF blinks then the noise source is working, and (3) if the H-OFF blinks between 43 and 50 times in one minute then the noise source is working well (ZCO, 1983a).

Modern cipher machines implement physical RNG for the generation of keys and initial vectors. Standards exist only for deterministic RNG, which shall be used in combination with physical RNG. The stochastic model of the noise source is still a necessary but difficult part of the security analysis of physical RNG (BSI, 2013b).

5.2 Self-Protection against Technical Failure

Technical faults of a cipher machine may result in weak encryption of the secrets (Figure 1). The self-protection of T-310/50 shall detect every security critical single technical fault and prevent adverse aftermath by blocking the output. The security analysis of the self-protection required a detailed analysis of the implementation up to the level of the electric scheme.

The T-310/50 consists of

- twice the complication unit (their output are compared by a control unit),
- the control units for security critical internal

[3] (BSI, 2013b) for definition.

sequences,
- the control units of the module connecting the peripheral devices and the line,
- the blocking unit which blocks the output when faults were detected, and
- the prophylactic checks of the control units and blocking.

The two complication units are implemented on three identical plates with two types of small plates implementing the long-term key. One of these plates implements function for both duplicating each other complication units. The ZCO specialists detected that a single fault on this plate may cause the same undetected deviation from the algorithm. Therefore, the plates were reworked. Finally the analysis found the self-protection effective (ZCO, 1983c).

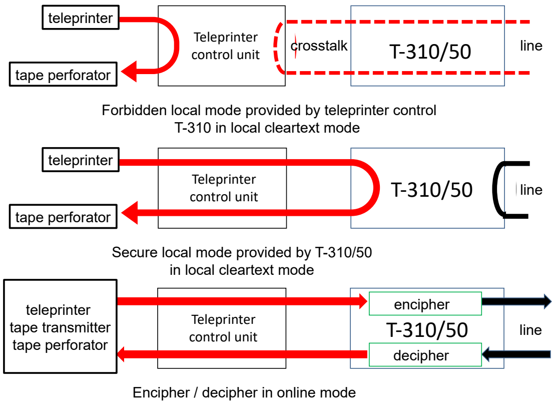
The prophylactic checks are enforced after input of the short-term key. The crypto officer shall run the prophylactic checks step-by-step (ZCO, 1983a). Each step simulates a fault that shall be detected by the control unit and cause indicated blocking of the output.

The analysis of ciphertexts in case of technical faults is a standard method called failure analysis. Today, the self-protection functionality must (or at least should) be a standard countermeasure against breaches of security caused by technical faults. The fault resistance of the hardware must be accompanied by software robustness, which is even more important and difficult to achieve for computers today.

5.3 Protection against Side-channel

The terminology document of ZCO (ZCO, 1971) states that emanation of electromagnetic, acoustic or other forms of energy by cipher machines, teleprinters, typewriters etc. may compromise secret information which are helpful for cryptanalytic attacks. Therefore, compromising emanation was an important topic of the security analysis not only of the cipher machine itself but also in combination with the teleprinter.

One of the vulnerabilities to mitigate was the crosstalk of secret plaintext from peripheral devices to the line. Figure 4 illustrates the problem. The teleprinter caused sharp edges of the plaintext signal on its output interface. The teleprinter control panel did not prevent crosstalk during local operation, e.g. preparing tapes with secret plaintext (upper picture). The operational manual of ARGON (ZCO, 1983a) prohibits the use of local mode using the teleprinter control panel. The T-310/50 implements a secure local mode of operation separating securely the local connection and the termination of the teletype line (middle picture). T-310/50 prevents also crosstalk of the plaintext on the local interface to the line interface during online encryption (bottom picture).



Forbidden local mode provided by teleprinter control T-310 in local cleartext mode

Secure local mode provided by T-310/50 in local cleartext mode

Encipher / decipher in online mode

Figure 4: Forbidden and secure local mode

The T-310/50 mitigates crosstalk by means of an “active electronic suppression” (“Aktive elektronische Entstörung (AES)”).

The AES was effective for electro-mechanical teleprinters (mass-produced machines since 1984, other machines were reconditioned). Electronic teleprinters like F1000 caused even sharper signals of the plaintext. Thus the T-310/50 cannot prevent crosstalk in this case. The use of T-310/50 with electronic teleprinter (e.g. F1300, F2000) protected against emanation and crosstalk was investigated and planed. That is why the installation manual (ZCO, 1986) allowed the use only of T-310/50 with modified electro-mechanical teleprinters T51 and T63 with attached tape punch T52, tape transmitter T53/x, $x = 3,4,5,6$, and the line switch T 57/4 of T 57/8. A exhibition of T-310 with electro-mechanical teleprinter is historical correct although the reason of the lack of electronic teleprinter is not obvious.

The emanation security and side channel protection were intensive investigated for cipher machines and communication technique by the ZCO. The analysis and the assessment followed a Soviet standard which was similar to American TEMPEST documents. The analysis of side channels and emanation are hot topics even today. The German Bundesamt für Sicherheit in der Informationstechnik published corresponding guidance (BSI, 2013a; BSI, 2008).

6 The Cipher Procedure ARGON

The cipher procedure ARGON defines how the operator and the crypto officer shall securely use both the cipher machine T-310/50 and the teleprinter. General security regulations and the technical provisions prescribed the organizational, personnel, physical and technical security measures for the premises, the installation and the operation for the use of ciphers in stationary and mobile stations. Based on this the ZCO specified the manuals for the installation of the T-310/50 and of the procedure ARGON as mandatory guidance documentation. The ZCO performed the trial of the prototypes and the procedures. It guided the training of the crypto officers and the operators. The ZCO and the cipher services inspected the application of ARGON in real life.

ARGON regulates the use of the T-310/50, the key material, the texts, the teletype components and the security measures. ARGON and SAGA are two different procedures for the very similar machines T-310/50 and T-310/51.

6.1 The Installation Manual of T-310/50

The installation manual (ZCO, 1986; ZCO, 1983b) defines the technical security measures for the installation of the T-310/50 and teleprinter components. The teleprinter components are connected to the cipher control panel. An additional teleprinter may be connected trough an additional control unit and the first control unit to the base unit. The control units may be installed in distance of up to 100 m from the base unit.

The basic unit and the power supply were located in an exclusion zone (“Sperrzone”) surrounding the base and power supply units of at least 0.5 m. The exclusion zone may access only explicitly authorized person e.g. the crypto officer for key management or the service personnel for maintenance of the machine. The controlled zone (“kontrollierte Zone”) surrounded the exclusion zone, all components of T-310/50 and the teleprinter components of at least 10 m. The plaintexts are input into and output from the T-310/50 within the controlled zone. The controlled zone prevents sojourn of vehicles, unauthorized persons or interception devices near to the devices and cables transmitting or operating secret information.

The installation manuals consider the results of the analysis of the emanation security and the physical protection of the T-310/50. The exclusion zone and the controlled zone are standard security measures protecting state secrets.

6.2 The Operational Manual of ARGON

The operational manual of ARGON shall ensure the general rules for the protection of state secrets by means of telex communication with T-310/50. The specialists of the ZCO performed a detailed analysis (ZCO, 1985) of every elementary reaction on any interaction through all interfaces of T-310/50 up to complex processes of cipher operation and network analysis. The machine allows for (1) the *transparent mode* (like in absence of the cipher machine), (2) the local and online *cipher mode* (encryption resp. decryption depending on the direction of the transmission), and (3) the local (offline) and online *monitoring mode* (entering decryption after receiving a synchronization sequence) (ZCO, 1983a). The T-310/50 implements some automatic processes e.g. the encryption of four fixed characters “Maschinenbefehlsfolge 2” BU, CR, LF, LF in front of the provided plaintext.

It followed that ARGON shall cope with two major issues: (1) the switchover between transparent, monitoring and cipher mode of operation, and (2) the change between sending/encryption and receiving/decryption of text without new synchronization in the cipher mode. These vulnerabilities must be mitigated by organizational countermeasures.

The operator may send by mistake secret plaintext in transparent and monitoring mode of operation. The analysts of ZCO found a precaution against unintended sending of plaintext if the T-310/50 is in monitoring online mode. The T-310/50 starts encryption by sending a “Maschinenbefehlsfolge 1” *bbbb* and the synchronization sequence containing the initial vector. When receiving this prefix at the peripheral interface or line interface in monitoring online mode the T-310/50 expects the synchronization sequence at the line interface and blocks any input at the peripheral interface. Therefore, the analysts suggested to add the prefix *bbbb* at the beginning of secret plaintext, e.g. on the punched tape prepared for sending by online encryption. In case of unintended transmission of the punched tape in monitoring mode the transmission will be stopped. The suggestion was followed in the operational manual (ZCO, 1983a, sec. 5). The input of secret plaintext with prepared punch tapes was mandatory if the ciphertext was send by radio transmission.

The T-310/50 being in cipher mode changes between encryption and decryption depending on the interface receiving the characters, i.e. encrypt characters input on the peripheral interface and decrypt characters input on line interface. This feature enables an encrypted dialog between the cipher stations. But this feature might be misused to provoke ciphertexts encrypted with the same short-term key and initial vector. Suppose the following scenario: An attacker *Eve* intercepts a message containing probably an interesting encrypted plaintext with short-term key K and a synchronization sequence with the initial vector V. *Eve* assumes that a cipher station operates a T-310/50 with the same short-term key K in monitoring online mode, but unwatched by the operator *Alice*. *Eve* establishes a connection with this T-310/50, imitates a legitimate cipher station, synchronizes the T-310/50 with V and waits for receiving a ciphertext from *Alice*. *Alice* sees the T-310/50 in cipher mode without any expected text. If *Alice* asks for clarification in cipher mode then *Alice* provides *Eve* a ciphertext encrypted with the same K and V. If *Eve* repeats successfully this procedure twice and can guess the plaintext sent by *Alice* and *Eve* can guess the corresponding plaintext *Alice*, then *Eve* may decipher the intercepted ciphertext. Such attacks are known for a long time and called nowadays social engineering.

The described scenario maybe used also as deliberate attack of an fraudulent operator compromising a specific plaintext. Instead of direct decryption of this plaintext the operator does not know the compromised plaintext and there is no evidence of the treason except the encryption of arbitrary text. Thus such actions must be explicitly forbidden in order to make the operator accountable for the adverse action.

In order to mitigate such attacks the operational manual of ARGON requires (1) the station dialing to the remote station must initiate the cipher mode, and stop any communication if the dialed in remote station starts synchronization, (2) watch the behavior of the machine during establishment of the cipher connection and the indication “C” of the cipher mode on the control unit during the encrypted communication, (3) enforce the required exchange the names of the station by the answer-back unit, and other required information, (4) to enter the transparent mode when leaving the T-310/50 unwatched (ZCO, 1983a, sec. 13).

The guidance for the training addressed the security requirements and regulations identified by the security analysis (ZCO, 1984). The analysts got feedback from the supervision of the application of the T-310/50 and ARGON by the cipher services.

The procedure ARGON may be demonstrated partly (local mode) if at least one T-310/50 is operational. Only two operational T-310/50 allows for detailed demonstration of ARGON and the potential problems as discussed above. Security problems of the operation may be solved by robust procedures reducing the probability and the affect of mistakes. Computers allow better ease of use, but the complex applications make analysis much more difficult.

7 Key Management

The cryptographic role of the long-term keys comprises the secrecy of the algorithm, the cryptographic reserve for strength, and the separation of networks (Stephan, 2022). The ZCO approved six long-term keys, but only three of them were implemented in hardware for ARGON and SAGA. The production of the plates with the long-term-key was controlled by the ZCO. The change of the long-term key for mass-produced T-310/50 would be difficult because many machines needed to work together. Thus the long-term keys were never changed in the field.

The ZCO produced all the short-term key material for distribution by the cipher services. The short-term keys were changed every week. The manufacturing of the short-term keys was developed, build, run, and continuously checked by the ZCO. The packets of punch cards were distributed over secure channels of the cipher services and stored in safes of the cipher stations. The packaging of the punch cards provided only a known limited physical security. The number of clients sharing the same short-time key was limited to 150 in order to reduce the security impact in case of compromise. The short-term key were put into the base unit by crypto officers. The operator of the teleprinter does not need to know the short-term key. But the operator shall destroy the short-term key stored electronically in the basic unit by push-ing a button “GG AUS” on his cipher control panel in case of emergency.

One should have these circumstances in mind when examining an exhibited punch card package or watching the key import in a museum. The key-distribution method was specific and appropriate for the cipher services. The cryptologists of ZCO knew the public-key methods. But public-key cryptography was not needed for the key management of the cipher services at the time of ZCO, because all cipher stations were under sole control of cipher services, including the distribution of the cipher machines and the key material.

8 Conclusion

The machine T-310/50 and the procedure ARGON are historical objects. Their development and the analysis provided a lot of insight for the cryptologists of ZCO. The cryptologists used their experience for the next generation of ciphers in the 1980s. The strength of the algorithm T-310 is still of interest, even the machine T-310/50 has historical value only. The security analysis of the machine T-310/50 and the procedure ARGON then provided by the ZCO is comparable to the modern Common Criteria evaluation on EAL 4 augmented with vulnerability analysis against high attack potential (AVA_VAN.5) (https://www.commoncriteriaportal.org/).

Acknowledgments

The author thanks Winfried Stephan and Franz-Peter Heider for the fruitful discussions and Nils Kopal and Bernhard Esslinger for proofreading. The author would also like to thank Jörg Drobick publishing additional information about the cipher service of GDR on his website http://scz.bplaced.net/.

References

BSI. 2008. *TR-03209: Elektromagnetische Schirmung von Gebäuden, Theoretische Grundlagen.* Technical report, BSI.

BSI. 2013a. *AIS 46.* Technical report, BSI.

BSI. 2013b. *Evaluation of random number generators.* Technical report, BSI.

Jörg Drobick. 2023. *Homepage Der SAS- und Chiffrierdienst.*

Killmann and Stephan. 2021. *Das DDR-Chiffriergerät T-310.* Springer Spektrum, Berlin. 978-3-662-61896-7.

W. Killmann. 2023. *On security aspects of the ciphers T-310 and SKS with approved long-term keys. Cryptologia*, pages 1–33.

W. Stephan. 2022. *Use of T-310 Encryption During German Reunification 1990.* In *Proceedings of the 5th International Conference on Historical Cryptology HistoCrypt 2022*, Linköping Electronic Conference Proceedings 188.

ZCO. 1971. *Fachbegriffe des Chiffrierwesens.* VVS - ZCO/407/71.

ZCO. 1980. *Kryptologische Analyse des Chiffriergeräts T-310/50.* Technical Report GVS ZCO Nr. 402/80, ZCO. BStU Archiv der Zentralstelle MfS - Abt. XI, Nr. AR3 594.

ZCO. 1982. *Das Auftreten gleicher Spruchschlüssel bei Geräten T 310/50.* Technical Report VVS-o020 MfS XI/393/82, ZCO. BStU Archiv der Zentralstelle MfS - Abt. XI, Nr. 596.

ZCO. 1983a. *Gebrauchsanweisung ARGON T-310/50.* Technical Report GVS B 434-081/83, ZCO. Harnekop NVA Museum.

ZCO. 1983b. *Gerätesystem T310/50 Installationsvorschrift (1. Ergänzung – Zentrale Chiffrierstellen).* Technical Report VVS B 434-065/83, ZCO. Harnekop NVA Museum.

ZCO. 1983c. *Technische Analyse des PBS des Chiffrators des Geräts T 310/50.* Technical Report GVS-o020 MfS XI/356/83, ZCO. BStU Archiv der Zentralstelle MfS - Abt. XI, Nr. 596.

ZCO. 1984. *Schulungsanleitung Verfahren ARGON (T 310/50).* Technical Report VVS B 434-416/84, ZCO. Harnekop NVA Museum.

ZCO. 1985. *Analyse der Bedienhandlungen am Gerätesystem T 310/50 und deren Konsequenzen für die Gewährleistung der Sicherheit der zu übertragenden Informationen.* Technical Report GVS MfS-Nr. XI/113/85, ZCO. BStU Archiv der Zentralstelle MfS Abt. XI, Nr. 665.

ZCO. 1986. *Gerätesystem T-310/50 Installationsvorschrift.* Technical Report VVS B 434-143/86, ZCO. Harnekop NVA Museum.

Deciphering Secrets Throughout History: An Interdisciplinary Linguistics and Cryptology Course

Eunice Kim
Classics Department
Furman University
Greenville, SC 29613
eunice.kim@furman.edu

Christian Millichap
Mathematics Department
Furman University
Greenville, SC 29613
christian.millichap@furman.edu

Abstract

This paper describes an interdisciplinary approach to teaching a linguistics and cryptology course. The authors, a Classics professor and a Mathematics professor, co-taught a three-week course, entitled "Deciphering Secrets Throughout History," to undergraduate students of varying backgrounds in mathematics and the humanities. Students were taught to apply tools from linguistics, statistics, and cryptanalysis to examine ancient texts, languages, and ciphers. The course culminated in an extended analysis of the fifteenth-century Voynich Manuscript, where students proposed their own original analyses of the text.

1 Introduction

In the past couple decades, a plethora of cryptology-related courses have emerged in the undergraduate curriculum. Many of these courses are traditionally taught in mathematics and computer science departments, with topics ranging from historical ciphers to post-quantum cryptography, and target audiences ranging from introductory students to graduating majors. Course topics, in-class activities, projects, and pedagogical approaches for teaching cryptology courses have been highlighted (though not exhaustively) in *PRIMUS* (Kaur, 2008), (Aydin, 2009), (Karls, 2009), (Schembari, 2020), *Cryptologia* (Winkel, 2008), (Kurt, 2010), (Glass, 2013), and previous HistoCrypt Conference Proceedings (Musílek and Hubálovský, 2018), (Krapp, 2019).

Some of these courses have gone beyond the typical mathematics and computer science audiences. For instance, Koss (2014) taught a first-year college writing seminar using cryptology as a vehicle for analyzing information literacy, critical thinking, and writing. In some cases, interdisciplinary courses co-taught by professors from two different disciplines have emerged. For example, Karst and Slegers (2019), an applied mathematics professor and a philosophy professor, co-taught a course partially focused on ethical issues in modern cryptography. Despite the wide range of interdisciplinary connections found within undergraduate cryptology courses, the literature lacks discussion on courses that have both a significant linguistics and cryptology focus. To help provide a novel contribution to the literature that fulfills this need, the authors, a Classics professor and a Mathematics professor, describe the structure of a three-week interdisciplinary linguistics and cryptology course that they co-taught and share some of their successful in-class activities.

2 Context

The setting for this unique interdisciplinary course was Furman University which is a small liberal arts university largely focused on undergraduate education. This school encourages interdisciplinary learning by annually offering a three-week May Experience (MayX) term after the spring semester ends. MayX classes must be non-traditional courses that are not taught during the semester, and they can be co-taught between two faculty members from different disciplines. While the authors have very different academic backgrounds (one is a theoretical mathematics Ph.D. whose research is in geometric topology but he has also taught some introductory college-level cryptology courses; the other is a Classics professor specializing in archaic Greek literature and historical linguistics, the study of language change, particularly within the Indo-European language family), co-teaching a MayX course that focused on applying tools from both cryptanalysis and linguistics in the context of ancient texts, languages, and ciphers seemed like a fantastic opportunity to combine our skill sets and academic interests. This resulted in the creation of "Deciphering Secrets Throughout History," (referred to as Deciphering Secrets moving forward), which we co-taught in May 2019.

For such a course, we hoped to attract students from a variety of majors to encourage interdisciplinary learning. The twelve students that enrolled in the course had a wide range of background knowledge in mathematics, cryptology, and linguistics. As a result, we had to introduce basic background material in cryptology and linguistics, do our best to make immediate connections between these two disciplines, and leverage student skill sets to encourage productive group work.

3 Course Components

In this section, we will discuss the major topics in cryptology and linguistics covered in this course and how this material was synthesized into a cohesive structure. Topics were chosen based on the audience, time constraints, and our course goals, which included: acquaint students with the history, development, and methods of cryptology, and linguistics; apply tools from cryptanalysis and linguistics to analyze historical codes and texts in order to draw informed conclusions about the structure of these texts; further develop independent research skills and group work skills; reflect on and communicate the importance of interdisciplinary collaboration in academics and non-academic careers. Below, we describe the key introductory topics emphasized in each discipline separately before turning to how students were encouraged to synthesize these complementary fields in an activity involving the Voynich Manuscript.

3.1 Cryptology and Mathematics Components

For the cryptology components of this course, we focused on the Caesar shift cipher, monoalphabetic substitution ciphers (MSCs), nomenclators, and the Vigenère cipher. These topics were partially chosen because many of them introduced important cryptanalysis tools, such as frequency analysis, vowel recognition algorithms, and the Index of Coincidence, that could be applied more broadly to linguistic analysis. For most of these topics, our main resources were Bauer's undergraduate cryptology textbook *Secret History: the Story of Cryptology* (Bauer, 2013) and Singh's *The Code Book* (Singh, 2000).

The cryptology portion of our course started with the (Caesar) shift cipher, which is a basic substitution cipher where each letter in a text is replaced with another letter by shifting down the alphabet by a fixed amount. Because of the small key space (only 25 nontrivial keys when using the English alphabet), brute force attacks can easily be implemented to break shift ciphers. However, we challenged students to develop their own tactics for breaking shift ciphers by hand based on the linguistic structure of the underlying language and the rigid structure of shifting. To this end, we provided cryptanalysis activities where the plaintexts were written in a variety of languages (and without knowledge beforehand of which language was used): English, French, German, and Latin. This led to a discussion on differences in linguistic structures of these languages: most common letters, most common words and letters that begin a sentence, most common short words, etc., and how these differ among the languages.

We then transitioned to MSCs, which encompass any encryption method where each letter or symbol used in a text is replaced with one and only one letter or symbol. Unlike the shift cipher, the number of possible keys can be quite large for an MSC. For instance, if the English alphabet is used for writing, then there are 26! possible keys, making a brute force attack not possible. For cryptanlaysis, this naturally led to introducing frequency analysis and expanding upon the linguistic attacks used on shift ciphers. For this topic, we again wanted to give students examples of ciphertexts to break where they were required to analyze the linguistic structures of languages beyond English. One MSC example we used that met these goals was Challenge 1 of the 2016 Kryptos competition. This challenge featured an intercepted encrypted telegram between rumrunners in the 1920s where context can lead to conjecturing that the underlying language is Spanish, and context clues can assist with looking for key words in the text; see `http://www.cwu.edu/math/previous-challenges`, for the corresponding exercise and solution.

Beyond these traditional topics in MSCs, we also discussed nomenclators and vowel recognition algorithms. A nomenclator implements an MSC for parts of the encryption process, combined with a list of code words and symbols used to replace more common words, bigrams, trigrams, and names, or represent nulls. This topic came up since we had students read parts of the *The Code Book* (Singh, 2000), and nomenclators are used in the Babbington Plot, which is discussed in Chapter 1. We also worked on implementing Sukhotin's vowel recognition algorithm, as described in Section 1.12 of Bauer (2013). Vowel recognition not only helps with breaking an MSC, but also can assist with conjecturing vowels in an ancient text where the underlying language or writing system might be unknown at the start. In particular, vowel recognition algorithms have been used to analyze the Voynich Manuscript, as noted in chapter 2 of Bauer (2017).

We next considered the Vigenère cipher, a polyalphabetic substitution cipher that requires a keyword to be shared between parties for encryption. This keyword designates a sequence of shifts to be used for substituting letters, with the keyword repeating as needed to complete the encryption process. For instance, if your plaintext said "Histocrypt" and your key was "dog," then a shift of $A \to D$ would be used to encrypt plaintext letter H to ciphertext letter K, a shift of $A \to O$ would be used to encrypt plaintext letter I to ciphertext letter W, and so on.

A major step in breaking the Vigenère cipher is finding the key length. We examined two methods for finding the key length: The Babbage–Kasiski test and the Index of Coincidence. Since the Babbage-Kasiski test did not come up in our Voynich Manuscript activity (see Section 3.3), we refer the reader unfamiliar with the technique to chapter 2 of (Singh, 2000). The index of coincidence (IC) measures the probability that two different letters chosen at random from a text will be the same. More formally, let N be the length of a given text and let F_α be the number of times a letter α occurs in that text. Then the index of coincidence for that text can be calculated as

$$IC = \sum_{\alpha\in\Omega} \frac{F_\alpha(F_\alpha - 1)}{N(N-1)},$$

where Ω represents the set of all letters in your alphabet, or more generally all graphemes in your writing system; see Section 3.2 for further discussion on linguistic terms. Since using a long keyword for the Vigenère cipher flattens the frequency distribution of the ciphertext letters, the IC can be used to provide an estimate on the key length.

In class, we covered the necessary background on combinatorics and probability to formally describe the IC and justify its connection to key length. We then worked on some basic examples of calculating the IC, where students created Excel worksheets to assist with calculations. Afterwards, students implemented both the IC and the Babbage–Kasiski test to assist with Vigenère cipher cryptanalysis exercises.

At the same time, by calculating the IC over numerous texts in a common language, one can calculate an expected value IC for that language. For instance, the expected value IC for English is approximately 0.0667, while the expected value IC for Spanish is approximately 0.0775; see Chapter 2 of Bauer (2013) for some more examples. Thus, this was a tool that naturally fit into our course since it could be applied to both breaking the Vigenère cipher (and more broadly, any polyalphabetic substitution cipher) and analyzing ancient texts to make conjectures about what language underlies a text or if the text was possibly encrypted using certain methods.

Our final mathematical topic was entropy, which supplies a statistical measure of information contained in a text or language. This concept was first introduced by Shannon (1948) and we refer the reader to Section 11.2 of Bauer (2013) for an introduction to this topic. One can theoretically compute the (expected value) entropy of a language as

$$H = -\sum_{i=1}^{n} M_i \log_2(M_i),$$

where M_i indicates the probability of message i, and the summation is taken over all possible messages in that language. In practice, such a calculation is unreasonable to compute. However, it is reasonable to calculate the n^{th}-order entropy, H_n, for n sufficiently small, which provides an estimate for H. For instance the first-order entropy of English can be calculated as $H_1 = -\sum_{i=1}^{26} p_i \log_2(p_i)$, where p_i is the probability of the i^{th} letter of English occurring on average in a text written in English. Similarly, H_2 is an expected value where probabilities of bigrams in the relevant language are used, and so on. In a similar manner, one can calculate the n^{th}-order entropies of a given text and make a comparative analysis with the expected value entropies of languages and other texts. See Bennett (1976) for a chart of first-, second-, and third-order entropy values for various languages and writers.

While we did not apply ideas from entropy to any cryptanalysis assignments, this topic naturally built off the probability background established from defining the IC and the use of frequency analysis in breaking MSCs. The only mathematical background needed to be introduced (or reviewed) were properties of logarithms. Like the IC, n^{th} order entropies can be used to conjecture if a text has been encrypted and which languages possibly underlie a text. Furthermore, entropy has been applied as a powerful quantitative tool in linguistic analysis, including the Voynich Manuscript, making this an ideal topic for our course.

3.2 Linguistics and History Components

To complement the cryptological principles, tools, and historical ciphers emphasized in the previous section, the linguistics component of the class introduced core concepts from the basic subfields of phonology (study of sounds), morphology (study of word form), and grammatology (study of writing systems and scripts). We also highlighted historical texts and examples of scripts and their decipherment through linguistic analysis, particularly the cases of Egyptian hieroglyphics and Linear B, as well as scripts still yet to be deciphered, including Linear A. Before discussing how these famous examples were deciphered or have been attempted to be deciphered, students were first introduced to the basic building blocks of language and writing.

We began with the question of what makes a language a language, and how spoken languages in particular are based on a system of sounds, as opposed to sign languages which are based on a system of gestures. Spoken languages rely on the human vocal tract's ability to produce sounds, which can be divided into two essential categories, vowels and stops (consonants). When combined, consonants and vowels form sound sequences in human speech that we call syllables. Syllables in turn serve as the phonological building blocks of words. For the purposes of time, we did not go into detail about additional types of sound distinctions that some languages employ, such as stress and pitch, among others. It was important to establish a basic understanding of vowels, consonants, and syllables before discussing how writing works. Starting in this way also reinforced the class's grasp of vowel recognition algorithms aiding in decryption.

Our next unit turned to the issue of writing systems and how they encode the sounds and ideas of a language through graphemes (or characters), which are the most basic contrastive unit of a writing system. We then provided a historical survey of the world's writing systems and the basic typology used by linguists to categorize them; see Daniels (1990) and Rogers (2005) for a more detailed overview of writing systems. All writing systems can be categorized as one of the following:

1. morphosyllabary: each grapheme stands for the sound of a morpheme, the most basic meaningful unit of a language (e.g. Chinese characters, or hànzì, for Chinese)

2. syllabary: each grapheme stands for a syllable (e.g. Katakana for Japanese)

3. abjad: each grapheme stands for a consonant only, while no vowels are represented (e.g. Hebrew and Arabic, although these are no longer pure abjads)

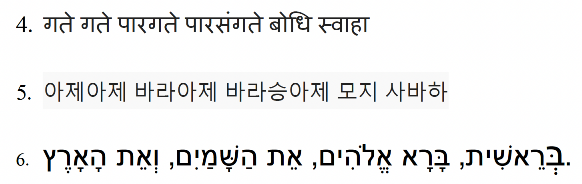
4. abugida: each grapheme stands for a consonant, while additional flourishes may be added to the character to indicate particular vowels (e.g. Devanāgarī for Sanskrit)

5. alphabet: each grapheme stands for either a vowel or a consonant (e.g. Latin alphabet for English, French, and many more)

6. featural script: each grapheme represents a phonological feature of a sound segment (e.g. Hangul for Korean)

As these writing systems represent different segments of sounds, each system consists of a different number of graphemes. Syllabaries, for instance, tend to feature a high number of graphemes, ranging from 50 to 200 characters, while alphabets tend to feature a lower number, ranging from 20 to 40 characters. Knowledge of these statistics helped the students to predict what type of writing system they were encountering, even if they had never been exposed to the script before. We also considered how each writing system posed different challenges and advantages to codebreaking methods based off frequency analysis. For instance, students needed to make the connection that the larger the set of graphemes in a writing system, the longer a text would need to be in order to approximate the average frequencies of the underlying language.

With both the types of world writing systems and the types of graphemes employed for each system established, we turned to other key elements of writing, including orientation of writing, syllable division, and word division. We discussed what visual cues we might use to determine each of these elements. We tasked students with analyzing known scripts (although the students did not necessarily recognize all of them) and identifying their graphemes, writing orientation, syllable/word divisions, and ultimately writing system type. The students had to explain concretely in linguistic terms how they were able to discern each of these elements. Examples of scripts we used for this exercise are provided in Figure 1 below.



4. गते गते पारगते पारसंगते बोधि स्वाहा

5. 아제아제 바라아제 바라승아제 모지 사바하

6. בְּרֵאשִׁית, בָּרָא אֱלֹהִים, אֵת הַשָּׁמַיִם, וְאֵת הָאָרֶץ.

Figure 1: Writing Systems Exercise

In item 4 of the figure above, for example, students were expected to recognize distinct recurring graphemes, to which some alterations were made above a key horizontal line, which should allow students to infer that they were dealing with an abugida script type.

We also considered why different languages adopted a particular writing system. This served as an introduction to the important concept of language classification by language families, which refer to groups of languages deriving from a common ancestral language. Some of the most well-known and well-established language families include Indo-European, Sino-Tibetan, and Afro-Asiatic, among many more. Semitic languages such as Hebrew and Arabic, a subset of the Afro-Asiatic family, have tended to employ abjad scripts, which lend themselves well to encoding the Semitic word structure, which is heavily based on a triconsonantal root and vowel patterns. Indo-European languages such as Greek and Latin, on the other hand, have less predictable vocalizations and have therefore generally eschewed abjad scripts in favor of other segmental scripts like the alphabet, since they can more clearly represent and distinguish vowels and consonants in writing. With an understanding of which writings systems different languages tend to employ, students could now make good hypotheses about potential underlying languages of an unknown script based purely off the type of writing system being used. For instance, if students identify an unknown script as an abjad type, they might reasonably assume the underlying language to be Semitic.

At this stage, students had a good grasp of essential linguistic concepts to consider codebreaking in a new light, and thus we prepared them to begin synthesizing the cryptological and linguistic components of the class by showcasing different historical examples of script decipherment. Our first case study was the decipherment of Egyptian hieroglyphics, which was an example where researchers did not recognize the script but knew the underlying language (Egyptian). The successful decipherment of hieroglyphics was ultimately made possible by the existence of the Rosetta Stone, which also included Demotic and Ancient Greek inscriptions that presumably translated the hieroglyphics, and hence provided a crib for the unknown script; see Robinson (2012) for a detailed history of this decipherment. The French historian and linguist Jean-François Champollion ultimately cracked the writing system by understanding that what appeared to be pictographs (drawings representing ideas) actually represented sounds. He was able to match the glyphs to particular sounds by linking cartouches to the phonetic representation of names such as Alexander (Alexandros) and Ptolemy (Ptolemaios), which appeared in the Demotic and Ancient Greek inscriptions on the Rosetta Stone. Names being names cannot be translated but must rather be phonetically mimicked in other languages. This historical example showcases the problems that arise in considering what a particular grapheme might represent: a vowel, a consonant, a syllable, or even an entire word or idea, in which case it might not be possible to discern the sounds of a language that correspond to a word represented in a pictograph.

Students were then tasked with identifying sound values for Egyptian hieroglyphics, which is primarily an abjad that does not represent any vowels in its glyphs, through a worksheet where we included cartouches recording the names Ptolemy, Berenike, Cleopatra, and Alexander.

This exercise provided a language complement to MSC problems emphasized during the cryptology-focused portion of the course. In both cases, students had to determine a direct one-to-one correspondence between two separate value sets, whether it was matching a sound to a grapheme or matching plaintext letters to ciphertext letters. The exercise also built upon previous linguistic concepts, since students once again had to isolate what constituted a single grapheme and figure out how these graphemes combined together to produce words. We provide below in Figure 2 the cartouches for Cleopatra (left) and Alexander (right) used for the worksheet.

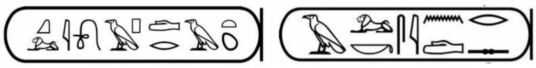


Figure 2: Egyptian Hieroglyphics Exercise

Our second case study in decipherment was that of Linear B, an ancient syllabic script that was preserved on baked clay tablets and used to record Mycenaean Greek in the Bronze Age. This example is significant in that researchers did not know the script or the underlying language recorded by it, nor were there any known translations that could serve as a potential crib. The decipherment of Linear B thus offers one of the most exciting examples of linguistic breakthroughs, which was made possible through the combined efforts of three individuals (Alice Kober, John Chadwick, and Michael Ventris), who each put forth unique contributions based on their professional experiences as a classicist, codebreaker and linguist, and architect; see Chadwick (1958) for a thorough overview of the Linear B decipherment. In particular, statistical methods (essentially frequency analysis) revealed underlying patterns of regularities that made it possible to associate phonetic and semantic values with the symbols. Through a form of brute-force attack, syllabic sound values were variously cross-checked (purely by guessing at first) across all the symbols until the combination of sounds for a particular word produced recognizable place names in Greece. Once the sound values for just a few symbols of the Linear B syllabary were established, the remaining number (with a few exceptions) were soon after resolved. Even before the sound values could be determined, linguistic analysis of the script had already proved that the encoded language was inflected. Linear B words with the same stable set of initial symbols (i.e. word stems), but with regular changes to their endings, signalled changes in word form that appeared to indicate different grammatical functions (i.e. inflection). This breakthrough allowed linguists to reasonably assume that the underlying language was Indo-European, and through knowledge of the historical development of sounds from Proto-Indo-European, the underlying language was ultimately determined to be Mycenaean Greek. The incorporation of both statistical and linguistic considerations was essential for developing a thorough understanding of the language and script of Linear B. This historical example modelled the type of interdisciplinary work we encouraged our students to emulate in our class.

To simulate the processes involved in the decipherment of Linear B, we created an exercise that involved deciphering a fictional script called Linear C. This script encoded another fictional language called Yomama. Students were provided a list of symbols that could be determined to be words through writing, but their pronunciation and meaning were still unknown. The words are provided in Figure 3 below.

a) ♥ ◊ ⊗ ♣

b) ♥ ◊ ♠ ◊

c) ♥ ◊ ♣ ∇

d) ∅ ♣ ♥ ◊

e) ∅ ♣ ⊕ ∇

f) ∅ ♣ ∇ ♣

g) ♦ ♦ ◊ ∇

h) ♦ ♦ ∅ ♣

i) ♦ ♦ ♦ ◊

Figure 3: Linear C Data

Students were then provided with an additional piece of data: a recently unearthed but fragmentary text revealing that the Yomama word for gold was pronounced *gobade*. We also shared the word in Linear C, provided in Figure 4 below. The students were tasked with finding the phonetic values for all the symbols of Linear C and were told to

⊗ ♥ ◊

Figure 4: Linear C Problem

assume the following: There are three consonants and three vowels, giving Yomama the simplest phonological system known; reading and writing orientation is left-to-right; every symbol stands for a CV (consonant-vowel) syllable; if two distinct symbols share a consonant, they must differ in vowels; if two distinct symbols share a vowel, they must differ in consonants; all words consist of a stem and suffix; stems are of the form CVCVC; all suffixes are of the form VCV, and it may be assumed that suffixes sharing their final syllable are of the same suffix.

The goal of this exercise was to get students to recognize the unique challenges posed by a syllabic script, since the individual graphemes do not easily reveal the vowel sounds that may be used for each syllable. But by a systematic approach modelled by the example of the decipherment of Linear B, the full CV sound values for each symbol in the Linear B syllabary could still be identified. Students were then able to better understand both the linguistic and cryptological implications of different writing systems that encode sounds in different ways, specifically how consonants and vowels may be variously encoded in writing (or not), and how this may ease or frustrate attempts to decipher an unknown script or to decrypt a challenging code.

In conclusion to the linguistic portion of the class, we shared some additional case studies of scripts that still have yet to be deciphered, such as Linear A. In this case, the script is the same as Linear B; however, the underlying language is unknown. All attempts to decipher it have been unsuccessful so far; see Salgarella (2020) for an overview of the unique challenges posed by Linear A. Ultimately, for such cases, we don't have enough surviving textual examples or data to allow for a securable decipherment. Exposure to this example encouraged students to consider the minimum amount of text necessary to allow for successful decipherment or decryption.

The goal of studying all of these historical case studies was to demonstrate that methods employed to decipher an unreadable script can also be employed to break a secret code or cipher and vice versa. Students could now attempt to approach a practical and famously longstanding problem of decipherment/decryption themselves.

3.3 Course Synthesis: The Voynich Manuscript Activity

Following the instruction of the individual components of cryptology and mathematics on the one hand, and linguistics and history on the other, the authors devised an original group activity that synthesized all of these topics as the culminating course experience. This activity required students to conduct an original analysis of the Voynich Manuscript (hereafter referred to as VM), a fifteenth-century illustrated codex hand-written in an unknown writing system with an equally unknown language underlying the script. A printed color copy of this manuscript consisting of all 116 folios (leafs, or pages, of the manuscript) can be found in Clemens (2016) and we refer the reader to Figure 5 for a visual of a folio. Though the VM has defied all attempts at decryption, it provided the perfect practical testing ground for the students to experience firsthand an attempt at code-breaking and script decipherment through interdisciplinary collaboration and synthesis of their newly acquired skill sets.

For this assignment, students worked in groups of three to perform a statistical and linguistic analysis of a folio of the VM, applying the different methods and concepts they had learned throughout the course. At the end of the analysis stage, each group gave a ten-minute presentation to share their conclusions. Four folios were preselected for the students to examine: folio 42 recto, folio 81 recto, folio 93 recto, and folio 99 verso. These four were selected due to the clearly discernible writing components and additional remarkable elements of each folio, particularly illustrations, which could provide useful context to the text and further aid the students in their original analyses. Students were able to examine these four folios in-depth through the high-definition images provided in `https://www.jasondavies.com/voynich/#f1r/0.5/0.5/2.50`. We also shared additional online resources to aid in their analysis (see Section 3.4 below).

The students were first tasked with isolating the set of graphemes that appeared in their folio. They had to consider how many unique graphemes they could identify, and what type of writing system

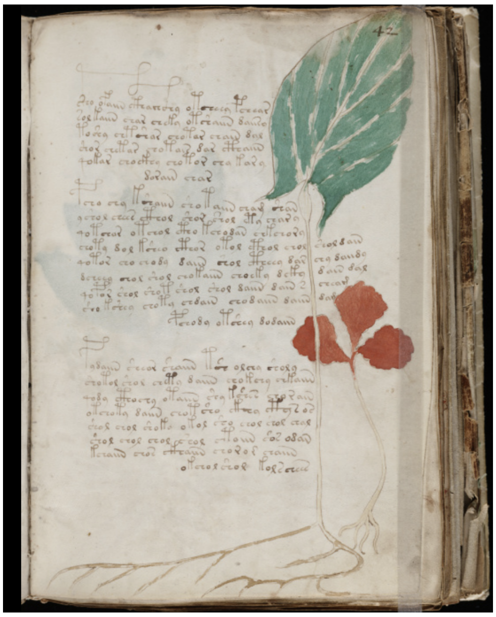


Figure 5: Folio 42 recto from the Voynich Manuscript

the number of existing graphemes might indicate. They also took care to note any patterns in the use and appearance of a single grapheme in the text, particularly if it appeared in a restricted environment, such as in word-initial or word-final positions.

Historically, this grapheme identification process for the VM has been a challenging but important task, since so many of the statistical tools discussed in Section 3.1 are dependent on having a clearly defined set of graphemes. There are theories and evidence that the manuscript consists of different sections authored by different individuals; see Chapter 2 of Bauer (2017) for more details. This challenge led to a productive class discussion on how many of the cryptanylsis tools we had discussed (e.g.frequency analysis, vowel recognition algorithms, index of coincidence, and entropy) could be applied to help identify if a text was encrypted, the type of encryption, or the underlying language. But these tests are all dependent on first identifying your set of graphemes. Thus, this led to a clear order of operations for an analysis of the VM: students first needed to identify the graphemes for the folio under analysis, and then look to apply the relevant statistical tools from cryptanalysis to draw conclusions about the text.

After this initial linguistic analysis, the students provided arguments and counterarguments for the use of a particular writing system being employed in the VM, considering first if it was a phonetic writing system and if so, which one: abjad, abugida, alphabet, or featural.

The next stage of analysis involved calculating the grapheme frequencies of their assigned folio. The students first created their own transliteration systems to make the VM text machine-readable and recorded their graphemes in an Excel sheet. Using their newly created datasets, they made the appropriate calculations, including the first order entropy and the index of coincidence for their folio. Results were interpreted through comparison to the entropies and indices of coincidence calculated for other known languages that they thought might be the underlying language of the VM. The students were also asked to consider what statistical tests other than average word length and word length distributions they might apply to their folios, and what complications might arise.

Following this close study of a single folio, the students expanded their investigation to the rest of the VM and attempted to find another folio that might come from a different hand, use a different writing system, or record a different language. Students had to justify their reasoning if they did or did not find evidence for different languages and writing systems within the VM. Finally, the students drew conclusions on what they thought the text of the VM represented at this point: a hoax, an MSC text, a PSC text, a universal language, a natural language (and if so, which language family), multiple languages, or some combination of these options.

This activity was clearly the highlight of the course, with students noting that this was a concrete area where they could enjoy and experience firsthand the rewarding coordination and synthesis of cryptology and linguistics.

3.4 Voynich Manuscript Worksheet

Below, we include the setup and questions from our VM activity (without commentary) to assist educators interested in incorporating or modifying this activity.

Directions: For this assignment, each student will work in a group of three to perform a statistical and linguistic analysis on a folio of the Voynich Manuscript. At the end of this analysis, each group will have 10 minutes to share their work with the class. Use the link `https://www.jasondavies.com/voynich/#f1r/0.5/0.5/2.50` to find the necessary folios of the manuscript. Group 1 is assigned Quire 6 f42r, group 2 is assigned Quire 13 f81r, group 3 is assigned Quire 13 f93r, and group 4 is assigned Quire 19 f99v.

You may find the following links helpful for your analysis:

`https://voyant-tools.org/` - text analysis, including word frequencies

`http://textalyser.net/` - letter frequencies, word frequencies, word length (don't rely on syllable count)

`https://md5decrypt.net/en/Letters-frequency-analysis/` - frequency count by letters, digraphs, trigraphs, etc., with comparisons to other languages

1. For homework, each member of your group should have determined the set of graphemes for your folio.

   a. Decide on a single set of graphemes for your group that you will use for the rest of your analysis. How many unique graphemes did your group identify for your folio of the VM?

   b. Which kind of writing system does the number of existing graphemes indicate?

   c. Provide arguments and counterarguments for the use of the following writing systems in the VM:

   i. Abjad

   ii. Abugida

   iii. Alphabet

   iv. Featural

2. Calculate the grapheme frequencies of your assigned folio. You can use Excel or other programs listed above. Also, it might help to develop a transliterative system and make the VM text machine-readable.

3. Calculate the first order entropy (H1) and the index of coincidence for this text and interpret your results. Compare these calculations to those of known languages that are possibly the underlying language of the VM.

4. What other statistical tests could you run to gather data on your page? For instance, your first reading on the VM showed comparisons on average word length and word length distributions to draw some conclusions about the script. Come up with one other statistical test that could be useful and apply that to your folio. Interpret your results and draw comparisons to other languages.

5. Do you notice any patterns in the use and appearance of a particular grapheme in your text? Do any graphemes appear in a particularly restricted environment, such as word-initial or word-final positions only? What can you conclude from the observation of such occurrences?

6. Look through the manuscript and find one other folio that your group thinks was written by the same hand, using the same writing system, and the same language. Justify your reasoning.

7. Look through the manuscript and find one other page that you think either came from a different hand, a different writing system, or uses a different language. Justify your reasoning. If no such folio exists, explain why.

8. If you wanted to do a more thorough statistical analysis of this text, what would you do? What complications might arise?

9. Based on your analysis of the VM with your team, what do you think the text represents at this point: a hoax, an MSC text, a PSC text, a universal language, a natural language (and if so, which language family), multiple languages, or some combination of these (specify)?

4 Conclusion

This paper has demonstrated how a cryptology course taught through multiple disciplinary perspectives can contribute to the current range of pedagogical approaches employed at an undergraduate institution. An interdisciplinary approach holds great appeal for students of broad disciplinary backgrounds and interests, and offers a promising way to enrich current undergraduate course offerings that focus exclusively on cryptology or linguistics in separate courses. The thoughtful implementation and combination of different disciplinary skill sets to the same problem can enrich student engagement, facilitate collaborative learning, and raise greater metacognitive awareness of undergraduate learning and its applicability to practical problems.

Finally, while this course did have a specialized format, there are elements that could easily be transferred to a variety of introductory college classes. Since no prerequisites were required for our students and there was not a significant amount of content covered from any one discipline, our major course activity on the Voynich Manuscript (discussed in Section 3.3) could easily fit into an introductory cryptology course taught in a mathematics or computer science department, a linguistics course, or as a student project in any such courses. We hope this paper inspires other educators to incorporate interdisciplinary approaches into their cryptology and linguistics courses.

Acknowledgments

This work was funded by Furman University's Interdisciplinary May Experience Course initiative. The authors would like to give special thanks to Christopher Hutton and the MayX Committee for their guidance and support.

References

Nuh Aydin. 2009. Enhancing undergraduate mathematics curriculum via coding theory and cryptography. *PRIMUS*, 19(3):296–309.

Craig P. Bauer. 2013. *Secret history*. Discrete Mathematics and its Applications (Boca Raton). CRC Press, Boca Raton, FL. The story of cryptology.

Craig P. Bauer. 2017. *Unsolved! The history and mystery of the world's greatest ciphers from ancient Egypt to online secret societies*. Princeton University Press, Princeton, NJ.

William Ralph Bennett, Jr. 1976. *Scienfitif and Engineering Problem-solving with the Computer*. Prentice Hall, Englewood Cliffs, NJ.

John Chadwick. 1958. *The Decipherment of Linear B*. Cambridge University Press, Cambridge, UK.

Raymond Clemens, editor. 2016. *The Voynich Manuscript*. Yale University Press, Yale, CT.

Peter T. Daniels. 1990. Fundamentals of grammatology. *Journal of the American Oriental Society*, 110(4):727–731.

Darren Glass. 2013. A first-year seminar on cryptography. *Cryptologia*, 37(4):305–310.

Michael A. Karls. 2009. Codes, ciphers, and cryptography—an honors colloquium. *PRIMUS*, 20(1):21–38.

Nathaniel Karst and Rosa Slegers. 2019. Cryptography in context: Co-teaching ethics and mathematics. *PRIMUS*, 29(9):1039–1059.

Manmohan Kaur. 2008. Cryptography as a pedagogical tool. *PRIMUS*, 18(2):198–206.

Lorelei Koss. 2014. Writing and information literacy in a cryptology first-year seminar. *Cryptologia*, 38(3):223–231.

Peter Krapp. 2019. Beyond schlock on screen: Teaching the history of cryptology through media representations of secret communications. *Proceedings of the* 2*nd Conference on Historical Cryptology, HistoCrypt 2019*, pages 79–85.

Yesem Kurt. 2010. Deciphering an undergraduate cryptology course. *Cryptologia*, 34(2):155–162.

Michal Musílek and Stepán Hubálovský. 2018. Teaching and promoting cryptology at faculty of science university of hradec králové. *Proceedings of the* 1*st Conference on Historical Cryptology, HistoCrypt 2018*, pages 137–143.

Andrew Robinson. 2012. *Cracking the Egyptian Code: The Revolutionary Life of Jean-Francois Champollion*. Oxford University Press, Oxford, UK.

Craig P. Rogers. 2005. *Writing Systems: A Linguistic Approach*. Blackwell Publishing, Malden, MA.

Ester Salgarella. 2020. *Aegean Linear Script(s): Rethinking the Relationship Between Linear A and Linear B*. Cambridge University Press, Cambridge, UK.

N. Paul Schembari. 2020. A half-rotor cipher for the classroom. *PRIMUS*, 30(5):552–570.

C. E. Shannon. 1948. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656.

Simon Singh. 2000. *The Code Book: The Science of Secrecy from Ancient Egypt To Quantum Cryptography*. Anchor, New York, NY.

Brian Winkel. 2008. Lessons learned from a mathematical cryptology course. *Cryptologia*, 32(1):45–55.

Situating ciphers among alchemical techniques of secrecy

Sarah Lang
University of Graz
Elisabethstraße 59/III
8042 Graz
Austria
sarah.lang@uni-graz.at

Abstract

This paper offers a contextual framework for the historical analysis of alchemical ciphers. It argues that they differ from other ciphers due to their unique context: the alchemical tradition embodies a performative culture of secrecy, which employs a variety of techniques to achieve this performance. This paper contends that the distinction between 'secret as content' versus 'secrecy as practice' presents a useful framework for understanding alchemical rhetorics of secrecy and their relationship to alchemical cryptography. Additionally, it demonstrates how these principles can be applied in interpreting several examples.

1 Introduction

Alchemical ciphers are something of a *terra incognita*. While some are known in alchemy research circles, most have not received the detailed attention they deserve. However, they distinguish themselves from other contexts in which ciphers were utilized, as the communicative traditions of alchemy are steeped in intricate techniques of secrecy, of which ciphers constitute only a single facet. The alchemical tradition is rife with practices of secrecy. Yet because of its distinctiveness, it is hard to integrate conceptually with other more well-known practices of secrecy such as diplomatic ciphers.

On the surface, it may seem that cryptography, which is quantitative and highly systematic, has little in common with other alchemical techniques of secrecy, which are qualitative in nature and often used creatively. However, when looked at in their historical contexts, it becomes obvious that these superficially disparate traditions actually share the same historical backdrop and are, most likely, used for the same reasons by alchemists and chymists. Seen through this lens, non-quantitative methods of encipherment, which are not strictly cryptography, are closely related. In essence, both pertain to what Katherine Ellison has termed 'cipher literacy' (Ellison, 2017): alchemists and chymists constituted a group of historical individuals in which methods of secrecy (including ciphering) flourished, and those alchemical experts defined themselves significantly by their hermeneutic abilities and mastery of a high degree of 'cipher literacy'. Many alchemists were scholars "trained in the tropological interpretation of texts" who demonstrated spectacular command of their encipherment techniques (Newman, 1996, 188).

To gain a proper understanding of alchemical cryptography, it is crucial to possess a basic understanding of what we can define as cryptographical stylistic devices. Among the most well-known of these devices are the so-called alchemical *Decknamen* (see section 10). Consequently, the cryptographical community should not only be interested in alchemical ciphers but the entire arsenal of encipherment techniques used by alchemists and chymists. Within the alchemical tradition, these encipherment devices, whether qualitative or quantitative, share a common purpose and cannot be understood in isolation.

Historians of cryptology may question how alchemy rife with qualitative methods of linguistic and iconic obfuscation is relevant to their interests. For once, as shown by the study by Bean et al. (2022), one of the earliest known Bellaso/Porta/Vigenère ciphers outside of a cipher manual was part of an alchemical scribal culture and textual recipe tradition, firmly embedded in the intricate cultural context of alchemical techniques of secrecy (Piorko et al., 2023). This single example alone has yielded a gold mine of historical insights. Piorko et al. (2023) show just how rich a historical close reading of an alchemical cipher can be, particularly given how few alchemical ciphers have been studied in detail so far. The polyalphabetic Bellaso cipher (Buonafalce, 2006) discovered in Sloane MS 1902 (Lang and Piorko, 2021) is a rare and early example of this type of cipher outside of a cipher manual (Kahn, 1996, 151).

Despite its prevalence in the tradition, alchemical secrecy, particularly regarding the use of cryptography, is understudied and not yet well understood. As of yet, there is no framework for the systematic study of alchemical techniques of secrecy. This paper presents an initial attempt to establish a theoretical foundation and framework for their interpretation and classification, as well as provide context for the historical analysis of alchemical ciphers. It argues that the distinction between 'secret as content' and 'secrecy as practice' provides an effective framework for interpreting alchemical rhetorics of secrecy and how they relate to alchemical cryptography.

2 Alchemy in the history of cryptography

Alchemy still lacks contextualization and adequate representation in cryptographical contexts: David Kahns *The Codebreakers*, a classic reference for cryptography studies, only mentions in passing that "[m]ysterious symbols were used in [. . .] astrology and alchemy [. . .] just as they were in cryptology. Like words in cipher, spells and incantations [. . .] looked like nonsense but in reality were potent with hidden meanings" (Kahn, 1996, 91). B. Láng conjectures, like many others have probably found, that alchemical or chymical encipherment does not seem to function in the same way as other types of secrecy in science do (Láng, 2018, 163, 165–166). He also remarks that "only a few ciphers applied in alchemical texts from before 1600 are known" (Láng, 2018, 165). Agnieszka Rec laments that alchemical ciphers remain a seriously understudied topic, especially given the abundance, even omnipresence of such devices in alchemical literature (Rec, 2014).

Some dedicated studies exist on the 17th century secretive practices and ciphers of chymists Robert Boyle (1627–1691) who is known as a public advocate for open communication in chymistry (Principe, 1992; Hunter, 2016) and the ones in George Starkey's (1628–1665) laboratory notebooks (Starkey, 2004) who therein "employs the full panoply of traditional alchemical cover-names – *Decknamen* – to describe the veiled processes that he employs" (Newman and Principe, 2003, 25). The goal of this article is to give context on alchemical techniques of secrecy as an explanation for why alchemical ciphering may be different from other kinds of historical cryptography. It proposes the distinction between 'secret as content' and 'secrecy as practice' as a fruitful explanatory framework for alchemical practices of secrecy.

3 Current trends in the historiography of alchemy

The switch from alchemical language to chemical nomenclature is generally considered a pivotal turning point in the history of chemistry. Some even argue that it is only with the *Méthode de nomenclature chimique* (1787) that modern chemistry was born during the 'Chemical Revolution' (Lefèvre, 2018). The long-held opinion that there was a 'Scientific Revolution' in the seventeenth century during which the scientific method emerged and amongst other things, secretive and obscure language in science was replaced by scientific openness is now widely contested within the history of science (Principe, 2011; Vermeir and Margócsy, 2012). The scholarly movement called 'The New Historiography of Alchemy' pioneered by Lawrence Principe and William Newman (Martinón-Torres, 2011) has demonstrated that the caesura between alchemy and chemistry is an artificial one (Newman and Principe, 1998; Newman, 2006). Importantly for the discussion of alchemical language, they have shown that most of the *Decknamen* which were formerly read as nonsensical products of the unconscious in Jungian and occultist interpretations of alchemy could, in fact, be translated to actual chemistry and the recipes they are contained in tested experimentally in a modern laboratory (Newman, 1996; Principe and Newman, 2001). This changed the historiography and public perception of alchemy so drastically that some have called it an 'Alchemical Revolution' in analogy to the metaphor of revolutions in earlier historiography of science (Reardon, 2011). While some have initially contested certain opinions expressed by Principe and Newman, the methods and theories of the 'New Historiography of Alchemy' have laid the foundation for what has become the *de facto* standard for alchemy research today and substantially contributed to its revival.

As part of this new historiographical turn, scholars have opted to be more deliberate in their use of terminology surrounding the alchemical tradition: For example, practitioners in the early modern period tended to call themselves 'chymists' rather than 'alchemists' which had become a slur. It is for that reason that the term 'chymist' and 'chymistry' shall be used for the early modern period, which is situated between Ancient to Medieval alchemy and modern chemistry, which only begins in the 18th century. It is, however, still appropriate to speak of an alchemical tradition when speaking of alchemy as a whole (as opposed to modern chemistry) which is why this article speaks of cultures of alchemical secrecy. These persisted well into the period when chymists were publicly calling for the abandonment of alchemical secrecy in favour of open language as part of a 'rhetoric of openness' (Golinski, 1990) demonstrated, famously, by individuals such as Robert Boyle, author of the 1661 *The Sceptical Chymist* (Principe, 1992). "The portrayal of chemical language as having recently freed itself from the obscurities of the past became a central feature of chemists' rhetorical presentation of their discipline" (Golinski, 1990, 375). To help make sense of these contradictions, an interpretation framework from the field of secrecy studies will be used in this article.

4 Attempts at analyzing (alchemical) secrecy

Theoretical foundations have been laid by G. Simmel who established the secret's sociological role as a tool for structuring hierarchy in a society (Simmel, 1908), and S. Bok, showing the secret's difference from privacy as well as its philosophical and ethical implications (Bok, 1983). In the recent secrecy studies research of the history of crafts and science, Long has followed Bok and defined the secret as 'intentional concealment' different from the private or the unknown, focusing on the secret 'as content' (Long, 2001), whereas Vermeir follows Simmel in investigating secrecy as a practice and social phenomenon as well as its implications for group dynamics: Secrecy and openness, according to him, form a range rather than polar opposites, challenging scholarly work on the 'Scientific Revolution' which has implied a teleological move from secretive unscientific traditions to the openness of science (Vermeir, 2012). Vermeir and Margócsy also pointed out that not only the 'contents' of secrets are interesting, but maybe even more so is the act of secrecy, its related social practices, and psychodynamics (Vermeir and Margócsy, 2012, 153). Benedek Láng has recently criticized the fact that secrecy studies and cryptology studies have, thus far, seldom been connected (Láng, 2018). Cryptology studies have mostly focused on solving ciphers and revealing their algorithms. Rarely have they asked about the socio-historical contexts and reasons why ciphers were used and from whom information encrypted using a specific cipher was actually hidden because, ultimately, "secrecy can only be defined in relation to a community with which one wishes to share the secret information" (Láng, 2015, 126). Similarly, secrecy studies have neglected the concrete results of secretive practices, that is ciphered texts. His claim is particularly relevant to alchemy since earlier existing theories on alchemical language, such as Umberto Eco's 'hermetic semiosis' and 'alchemical discourse', consist of claims which are historically intangible in that they cannot be verified or validated using concrete examples of historical texts (Eco, 2016). Furthermore, their explanatory value for historical remnants of alchemical secretive practices, such as cryptography or other forms of veiled communication, is minimal.

5 The topos of the 'alchemical secret'

Alchemy has a tradition of guarding secrets (Bachmann and Hofmeier, 1999, 9). The assertion that secrecy is a central aspect in the perception and discussion of alchemy is widely agreed upon in secondary literature (Ebeling, 2001; Principe, 1992, 63). Ebeling stresses that the concept of the secret itself has to be clearly distinguished from the reasons given for concealment practices (Ebeling, 2001, 63–64) which are themselves part of a 'rhetoric of secrecy'. Principe, on the other hand, puts special emphasis on the question *from whom* a secret was supposed to be hidden, and conversely, for which audience it was intended to be comprehensible (Principe, 1992; Principe, 2000, 141). The type of secret most commonly associated with alchemy is the 'hermetic secret' or 'empty secret' as popularly criticized by Umberto Eco (Eco, 2016). However, this theory cannot stand any longer after the scholarly movement referred to as the 'New Historiography of Alchemy' has been able to show that, in fact, many of those supposedly ‘empty secrets’ were not empty at all – historians of chemistry were able to read the secretive alchemical language chemically and recreate the processes described in the recipes in their modern laboratories (Martinón-Torres, 2011). The encipherment of alchemical language was thus decrypted by means of ‘practical exegesis’ (Rampling, 2020, 63–64, 97–99, 354).

B. Láng suspects with regard to ciphers that, in some cases, a historical actor “might simply have regarded encrypting as a playful activity. He seems to invite readers for a game” (Láng, 2018, 159). Eamon called this the game of *venatio* (Eamon, 1994). Alchemical texts, too, tend to use encipherment in playful ways (Bilak, 2020), yet this is probably more pronounced in allegorical or emblematic contexts than with actual cryptography. As explanations for alchemical secrecy, scholars further cite the type of knowledge communicated or the fact that alchemical transmutation revolved around money and power (Eis, 1965) or stress that alchemical rhetoric of secrecy doesn’t differ from the oats of secrecy present in other *artes* (Telle, 1978, 211). Vermeir notes that “alluding to secrecy might be the best way to disseminate your ideas” (Vermeir, 2012, 188) and “secrets publicized in print were often viewed as less valuable or proprietary than those confined to manuscripts” (Leong and Rankin, 2016, 15) or those confined to oral transmission altogether. Self-promotion was likely a strong motivation for engaging in theatrical performances of secrecy (Leong and Rankin, 2016, 13). Vermeir stresses that such rhetorics of secrecy were a powerful aspect of patronage and salesmanship:

> To understand such phenomena, it is important not to be misled by the actors’ categories and not to take the rhetoric of secrecy at face value. There is nothing paradoxical, per se, in the dissemination of secrecy or the values of secrecy, and many of the secrets transmitted in the books of secrets were ‘open secrets’ that were already widely known and applied. [. . .] Cunning use of the rhetoric of secrecy was a powerful means of building a reputation, by advertising that one has a secret as widely as possible and at the same time carefully controlling access to the content of the secret (Vermeir, 2012, 180).

By framing their knowledge as precarious knowledge (Mulsow, 2012) and through self-fashioning as professors of this exclusive knowledge, entrepreneurial alchemists could make their knowledge and products seem more valuable in the ‘economy of secrets’ (Jütte, 2011). However, any alchemical techniques of secrecy have a dual function: not only do they promote their user as someone who may be in possession of valuable secrets, they also represent “performances of expertise in the marketplace” of entrepreneurial alchemy (Nummedal, 2007, 170–172).

6 Alchemical rhetorics of secrecy

Because one “cultural function of secrecy is to establish boundaries” (Eamon, 2006, 234), the content of the secret is sometimes secondary. Rather it is the fact *that* there supposedly is an information gap between different actors that matters. The intentional concealment needs to be made known to all parties involved and aims at generating a hierarchical imbalance of power. The rhetoric of secrecy is the strategic game that creates this asymmetric relationship between the one who has and the one who seeks knowledge (Lochrie, 1999, 93). Not all contexts where secrecy is performed involved actual secrets (Vermeir and Margócsy, 2012, 164). Early modern secrecy is theatrical and performative, oscillating between hiding and revealing; the secrets often only become meaningful when seen as performative acts: “[S]ecrecy and openness are norms or values that regulate behaviour” as well as “characteristics of practices” (Vermeir, 2012, 166). Vermeir criticizes a strong focus of past historiography on the contents of secrets while simultaneously disregarding their performative value in practices:

> In many instances, what is kept secret is not even relevant for studying the dynamics of secrecy, i.e. the practices of simulation and dissimulation, the rhetoric of secretiveness, or the strategies of hiding and revealing that are employed. [. . .] As objects of desire, secrets accrue a special value, even if their content would in itself be valueless. They hide the real value of the content by keeping it hidden (Vermeir and Margócsy, 2012, 160, 162).

Like Vermeir stresses, a “rhetoric of secrecy communicates not facts but certain expectations, attitudes, and feelings – it creates a fascination, a certain thrill – and invites certain behaviour." It should therefore be kept separate from the assumed 'contents' of secrets (Vermeir, 2012). A rhetoric of secrecy doesn't have to imply the presence of actual 'secrets as content'. In the same vein, only because authors publicly call to abandon traditional alchemical means of communication does not mean that they themselves stop using alchemical stylistic devices or cryptography in their texts – despite promoting the opposite of calls to secrecy and rhetoric of keeping the alchemical secrets, calls to openness are part of the same theatrical tradition (Golinski, 1990). Even public advocates for the abandonment of the obscure alchemical language such as Robert Boyle were still using ciphers and code in their notes or correspondence (Principe, 1992, 63–67).

7 'Books of secrets' and secret as content

A genre especially relevant to the question of alchemical secrecy are so-called 'books of secrets' (Eamon, 1994). When the term 'secret' is interpreted outside of its historical context, it can be misleading because we tend to associate meanings with it that may not have been as dominant historically as they are today: In the contexts of these so-called 'books of secrets',

> the word 'secret' could also refer more specifically to a set of procedures known only to a select group of initiated individuals – in other words, craft or trade secrets. [...] This kind of secret was more about technical know-how, or 'how to', than hidden knowledge (although the two concepts were by no means mutually exclusive). [...] A secret could [...] be a physical object (a remedy) as well as the knowledge required to make it. (Leong and Rankin, 2016, 8–9, 12).

'Books of secrets' were a historical form of 'how-to' literature which usually contained all sorts of recipes describing processes that ultimately consist of "a set of operations known to any metalworker or distiller" (Smith, 2016, 48), yet they are marketed towards a popular audience as instructional manuals of didactic value (Eamon, 1994; Eamon, 2016). These books are a material container for crafts knowledge which had been viewed as proprietary knowledge in Medieval times but became more profane with the advent of print, making techniques previously reserved to a select group available to anyone who was literate and thereby reducing the meaning of what used to be a 'secret' to a mere technique (Davids, 2005, 342–343). In the sixteenth century, the book market started to become flooded with *alchemica* and 'books of secrets', culminating in wide popularity during the seventeenth century, indicating that such books weren't only interesting to a narrow group of experts and their potential customers. Books claiming to share the most secret of secrets often became instant bestsellers (Eamon, 2013, 60).

8 Secret publications

Beyond practices and performances of secrecy within alchemical texts, alchemical techniques of secrecy can also pertain to the mode of publication itself: Books with missing publication information are not a rarity in the alchemical context. Sometimes the secrecy even pertains to the physical books themselves: While we today assume that all books of one edition must be the same, in the context of hand-press print, this is often not the case. Parts of books are missing or added in some copies which were not included in the main issue – this is especially true for all materials in the front matter! –, publication information is left out. There is even the curious case of Arthur Dee publishing a 'Rosicrucian issue' of his 1631 *Fasciculus Chemicus* (1631) which would have gone unnoticed if detailed bibliographical analysis had not been performed on it (Piorko, 2019). The unicality of copies is thus another crucial element to be taken into account when dealing with alchemical print culture. In other cases, alchemists themselves admit to publishing their books as if they had not been published (to avoid sharing alchemical secrets with too big an audience, or so they claim). As an example, let's consider alchemist Michael Maier's (1568–1622) first printed book *Coelidonia* (Maier, 1609): He states that the book was published as though it had not been published.[1] The publication date (1609) and publication place (Prague) are encrypted on the back of the title page. In

[1] Latin text: "Editus est enim hic liber, quasi non esset editus, cum nusquàm publicatus aut vulgo prostitutus sit, sed in doctrinae filiorum gratiam, rarissimis exemplaribus inter privatos parietes conservetur." Maier (1609), [*r].

his 1614 book *Arcana Arcanissima* ('Most Secret of Secrets'), Maier offers an introductory poem containing anagrams of his name (Maier, 1614; Tilton, 2003, 82). He also often contributed to the front matters of his friends' publications under the name anagram 'Hermes Malavici', highlighting the alchemists' tendency to publish anonymously. In the case of the anagram, Maier writes under a pseudonym, thus hiding his true identity from all except an initiated few. But it is also quite common in alchemy to publish eponymously, i.e. attributing one's work to an earlier authority to make it seem older than it is and more venerable. This is, for example, the case in the large corpus attributed to the 9th-century Arab alchemist Jabir ibn Hayyan (Principe, 2013, 33–45). A famous Western contribution to this corpus is so-called Pseudo-Geber who was actually a late 13th-century Italian monk (Newman, 1991).

9 Alchemical language and terminology

Now that we have seen rhetorics and practices of secrecy in advertising alchemical books, we will investigate the most prominent form of alchemical secrecy: its characteristic language and cryptographic stylistic devices employed by the alchemists. When creating terminology for the sciences today, scientists aspire to create unambiguous terms. However, most alchemical *Decknamen* are highly dependent on their context, even more than normal words. This is why alchemists and chymists can use them creatively to suit their own needs and occasionally also as an effective method of hiding the true meaning of their recipes from the uninitiated. This may initially seem arbitrary, however, the word substitutions are usually based on common properties of what they actually mean and the word they use in its place. Lawrence Principe shows an example of such concealment in Robert Boyle's laboratory notes.[2] Similarly, Michel Butor (Butor, 1990) and William Newman have stated that "[...] *Decknamen* are not arbitrary, they change their meaning with context" (Newman, 2018, 33). In the historiography of alchemy, those word substitutions or cover names omnipresent in alchemical texts have come to be referred to as 'Decknamen'. The term is meant as a neutral term, thus the use of a German loan word. The problem with understanding alchemical language as terminology (*termini technici*) is that this does not adequately reflect its nature, as can easily be seen in the example of *Decknamen*. Our modern understanding of terminology is that of fixed unambiguous meanings, it evokes thoughts of chemical nomenclature, yet this is not at all what we encounter in alchemy.[3] It shall be argued here, that it is more fruitful to conceptualize alchemical language as a specialist sub-language and its specificities as (cryptographical) stylistic devices.

Many older publications and lexica on 'alchemical symbols' actually refer mainly to iconic symbols, e.g. Lüdy (1928). While these undoubtedly range amongst the particularities of alchemical texts, they are only one characteristic aspect beside many others such as *Decknamen*, technical terms which do not match the criteria of being *Decknamen*, stylistic devices, rhetorics of secrecy, allegorical images, riddle conundrums and even cryptographical encipherment. While iconic symbols catch the eye immediately, it is still alchemical *Decknamen* which are the most typical alchemical devices. And although *Decknamen* may have their epistemological advantages for communicating chemistry in an age where many chemical phenomena were hard to quantify or analyze, many chemists themselves felt frustration with their ambiguous terminology. A whole range of alchemical lexica testify to this impetus which increasingly took off during the seventeenth century (Ruland, 1612; Sommerhoff, 1701).

[2] "Name substitution is ubiquitous in alchemical treatises where common words like *mercury* or *sulphur* cause endless confusion by their broad application to a myriad of different substances. Boyle uses this standard technique, for example, in a laboratory account dated 29 April 1657. The text describes a process wherein copper is dissolved, distilled, and extracted into a tincture which, when digested with tin, is able to tinge that metal with a yellow colour. [...] In three of the four cases where the word *copper* appears, it has been crossed through and the alchemical symbol for gold written above it (in the fourth case the metal copper is actually meant). Wherever *tin* occurs, that word has been replaced with either *Silver* or *Lune*. Boyle's corrections reveal the text as a receipt for the transmutation of silver into white gold. This substitution scheme is undoubtedly founded upon the similarity between copper and gold – the only coloured metals – and between tin and silver – the most brilliantly white metals. Word substitution is Boyle's most common method of concealment" (Principe, 1992, 64).

[3] Incidentally, terminology as a science in the modern sense began only with Eugen Wüster (1898–1977), see Wüster (1991). It thus makes no sense to apply present-day standards for terminology to alchemical language which originated even long before the early modern lexicographical and terminological endeavour of the emerging natural sciences.

10 (Re-)Solving *Decknamen*

The term *Decknamen* originally stems from an early 20th-century German Arabist tradition (von Lippmann, 1919; Ruska and Wiedemann, 1924) and was reclaimed as a neutral *terminus technicus* in the context of the 'New Historiography of Alchemy' by Lawrence Principe and William Newman (Principe, 1992; Newman, 1996).[4] Contrasted with the somewhat related tradition of cryptography, *Decknamen* can be defined as symbol words employed as a substitute to avoid publicly naming which substances a recipe contains or to signify a specific chemical phenomenon for which there was no other adequate description. While some created neologisms or used somewhat consolidated terms to gesture to a certain substance or phenomenon with a relatively stable connotation (such as the 'green lion' for vitriol), others used allegories or figures from mythology creatively. This specific practice is referred to as mythoalchemy.[5] *Decknamen* and other related phenomena are better thought of as cryptographical stylistic devices rather than terminology. This term acknowledges that they are not quantitative ciphers but are often used to conceal, too, albeit using a qualitative substitution logic.

Through the use of so-called performative methods in the history of science, the research tradition called the 'New Historiography of Alchemy' pioneered by Lawrence Principe and William Newman (Principe and Newman, 2001) was able to show using historical-critical replicative experiments that alchemy and chymistry's peculiar language is a deliberate style, achieved by the use of encipherment techniques which, beyond merely hiding knowledge, also serves to appropriately communicate the multi-sensory experience which is chemical experiment. Like a cryptographical system, alchemical tropes such as *Decknamen* are employed in a way that may seem confusing, even inscrutable to outsiders, yet has a logic to it that can be utilized to get at their hidden meaning. Cryptography is a less embellished and more practical way of hiding information, in contrast to the playfulness of *Decknamen* which is a performative gesture (secrecy as practice) as much as it is a means of encrypting information (secret as content). The importance of hermeneutics as a key skill of the able alchemist was already stressed by Zosimos who insisted that "only a correct interpretation of the earliest writings and of their hidden meaning could disclose the right way to perform alchemical procedures." (Martelli, 2016, 227). Newman argues that showing off one's ability to decode and write complicated alchemical language, thus the adept's hermeneutic skill, served to establish credibility in an alchemist's practical skills, establishing authority and showing that one had a righteous part in the tradition of older alchemists. He writes:

> Alchemical writers delighted in announcing that they were going to explain a riddle – only to give the answer in the form of a conundrum. [. . .] The alchemists themselves maintained that a diligent reader could decipher their language to arrive at a correct alchemical *praxis*. [. . .] But there is another element that the reader was meant to derive from his alchemical sources. This was the aura of authority that a contemporary figurative text acquired by employing the metaphors utilized by older authors (Newman, 1996, 164).

It further offers certain epistemic advantages which is why, for example, in the plaintext decoded by Bean et al. (2022), alchemical texts may still contain their paradigmatic *Decknamen* despite being encrypted in a second layer of ciphertext. This shows that they serve a function complementary to mere encryption.

William Newman has provided descriptions of and names for the most common alchemical stylistic devices such as *Decknamen*, *parathesis*, *syncope* and *dispersio* (Newman, 1996, 159–188). The most important stylistic devices of the alchemical tradition are the 'dispersion of knowledge', that is spreading information over multiple passages or even books (which can be reunited if one pays attention to certain signal words), part of which can be the use of *syncope* ("the elliptical description of an alchemical process" as found in highly abbreviated recipes) or, on the other hand,

[4] They intentionally claimed this German term to avoid having to use an English one which might have a pejorative connotation, even though one might argue that the term 'cover names' entails a notion of intentional concealment in either case. This was actually a connotation Principe and Newman sought to avoid because many *Decknamen* function like simple *termini technici* and it would be misleading to assume that they are always employed with the deliberate intention of concealment.

[5] There is a whole research tradition on mythoalchemy (see Forshaw (2020)).

parathesis ("the heaping-up of synonyms for a given process, substance, or apparatus, again with the intention of bewildering the reader [. . . as] in the profusion of names used" for one single concept) (Newman, 1996, 187). These more complex stylistic devices usually also make use of the most basic element of alchemical style – its *Decknamen*. Table 1 presents a non-exhaustive list of alchemical techniques of secrecy.[6]

Technique	**Example/Reference**
Decknamen, specialist terminology	Newman (1996)
word/name substitution	Principe (1992)
dispersion of knowledge (*dispersio*)	Principe (1992, 65)
parathesis & *syncope*	Newman (1996)
monoalphabetic ciphers	Principe (1992, 67)
polyalbabetic ciphers	Bean et al. (2022)
iconic symbols & codes	Gaede (2017)
alphanumeric knowledge charts	Clucas (2017)
astrological horoscopes	Piorko et al. (2023)
cabbalistic mysticism	Forshaw (2013)
Lullian diagrams	Forshaw (2013)
emblems	Maier (1617)
(mythoalchemical) allegories	Forshaw (2020)
omitted or enciphered publication information	Maier (1609) Piorko (2019)
pseudonomia	Newman (1991)

Table 1: Non-exhaustive list of alchemical techniques of secrecy

11 Alchemical secrecy in practice

Bean et al. (2022) and Piorko et al. (2023) have presented a decrypted alchemical recipe called the *Marrow of Hermetic Philosophy*, found in a medical manuscript (British Library Sloane MS 1902), which contains astrological and alchemical predictions for health and death written by John Dee (1527–1608) and his son, Arthur Dee (1579–1651). Decoding the alchemical cipher within Sloane MS 1902 and tracing its copying in additional manuscripts have shed light on the dissemination of alchemical secrets within Anglo-Scottish knowledge networks of the seventeenth century. This manuscript, previously largely ignored in scholarly literature, came back into the public eye after the cipher contained in it had been decrypted in 2021 (Bean et al., 2021). The cipher used in the manuscript is an early example of a polyalphabetic Bellaso cipher, a strong encryption method, which was historically deemed undecipherable. By exploring the manuscript context in which the cipher was copied and transmitted, we gain invaluable insights into alchemical practices of secrecy and how alchemical secrets were shared. It also provides evidence for the dissemination of this cipher as part of a larger alchemical knowledge network. The same encrypted recipe, along with the Latin plaintext, was subsequently found in a manuscript at the University of Edinburgh (MS DC 1.30), archived with it dated laboratory notes describing the process in practice. An additional reference to the unique passphrase used to decrypt the ciphertext can be found in yet another alchemical manuscript at the Bodleian Libraries (MS Ashmole 1423). Corrections were made to MS Dc 1.30 between rows of ciphertext and in the margins, with some corrections appearing in both manuscripts. The likely author of this manuscript is Patrick Ruthven c. 1629. The evidence from mistakes in both manuscripts supports the argument that neither Ruthven nor Dee was responsible for the original encryption of the recipe.

The key phrase for the ciphertext, "Sic alter Iason aurea felici portabis uellera Colcho," is adapted from the last lines of Giovanni Aurelio Augurello's poem "Chrysopoeia Minor" (Soranzo, 2020, 35–39, 86–88, 110-121) which refers to the myth of Jason and the Golden Fleece in its last lines. Thus, in addition to the strong polyalphabetic encryption and the *Decknamen* present throughout the recipe, obscuring significant parts of the process to readers not deeply familiar with alchemical experiments, a mythoalchemical allegory is added to the mix. Augurello's poem is not a retelling of the traditional myth of Jason and the Argonauts (as Jason is only mentioned in the last lines in the form of name-dropping), but most likely interprets the Golden Fleece as a mythical animal skin that could be used to create a book containing alchemical recipes. The encrypted recipe's key phrase is an exceptionally long one, chosen likely to point readers to Augurello's poem. If one interprets this as a signal for the stylistic device of *dispersio*, omitting relevant information in one place which the reader has to gather from another text, it may also be a clue as to why the encrypted recipe seems to be-

[6] Please note that alchemical symbols are only sometimes secretive when used in codes and ciphers. They most often have specific and fixed meanings or, if not, are used creatively by authors rather than with the intention to conceal necessarily.

gin after a significant part of the experiment has already been completed. *Hermeticae Philosophiae Medulla*, thus, contains an encrypted mythoalchemical allegory that obscures a practical recipe shared via manuscript among a group of physicians seeking alchemical knowledge – or performing before other colleagues that they were in possession of such rare sought-after knowledge.

However, there are practical aspects of the manuscript copies of Hermetiae Philosophiae Medulla that remain mysterious regarding the circulation of this cipher. For instance, the cipher included in Sloane MS 1902 would have been impossible to solve with the incorrect cipher table provided, even with the key phrase. In contrast, a functioning cipher table was included in MS Dc 1.30, but with a partially incorrect key phrase. These two manuscript copies alone are insufficient for cracking the code and would have required external knowledge (likely the original copy) to use this recipe. Additionally, this cipher is unique in that there are hardly any examples of polyalphabetic ciphers from the first half of the seventeenth century, when monoalphabetic ciphering systems were still commonly used despite their cryptographic vulnerability. A second question that arises is the purpose of copying an unsolvable ciphertext and table into a medical manuscript in the first place. In the example of MS Dc 1.30, the Latin plaintext is included with the encrypted ciphertext, so the cipher has already been solved. However, in the medical notebook compiled by Arthur Dee, the broken cipher table and encrypted ciphertext, paired with the key phrase, function as a practice of secrecy independent from the practical recipe (which represents the 'secret as content'). In both cases, the act of copying the cipher into an alchemical medical manuscript is itself a performative allusion to possessing the knowledge encrypted within the cipher, the ultimate alchemical achievement of the Philosophers' Stone. Beyond the theatrics and performance of secrecy in the alchemical compilation networks of *Hermeticae Philosophiae Medulla*, this cipher highlights the importance placed by alchemical adepts on both the recipe obscured within the cipher (secret as content) and the practice of sharing, decoding, and obfuscating secret alchemical knowledge (practice of secrecy).

Another example for a cipher in the context of alchemy is Emperor Rudolf II's "Alchemical Hand Bell" (Bean et al., 2023). The cipher found on the handbell poses a significant challenge in deciphering, as there is little contextual information available. It remains unclear whether the cipher is a genuine code that has yet to be deciphered or serves some other symbolic function meaningful only to its creators. In this case, the next step is to determine whether a solution is even possible given just how short the ciphertext is.

12 Conclusion and future work

Alchemical ciphers remain an understudied field of research with much to be uncovered (Rec, 2014). As is evident from all the different techniques of secrecy described here, alchemical secrecy is a complex conundrum. It would be foolish to just look at alchemical ciphers in isolation without taking all these other phenomena into account. Alchemical secrecy is first and foremost a performance. And while some of the secrets of alchemy can adequately be described using the framework of 'secrets as content', many cannot. Those elements which are better described as 'secrecy as practice' or even 'rhetorics of secrecy' often provide the historical context we need to make sense of alchemical 'secrets as content'. Ciphers are, in many ways, easier to analyze because statistics leave less room for interpretation than qualitative research. However, when the plain text does not yield much in terms of *why* the message was hidden or in which historical context, traces of all the other techniques of secrecy used by alchemists combined might yet provide us with sufficient information to ultimately make sense of the latter.

In order to advance systematic research on alchemical cryptography, it is a desideratum to create an inventory of known alchemical ciphers. In the author's personal experience, many alchemy researchers have encountered enciphered texts in the archives before, yet they did not know what to do with them. They are also often not flagged explicitly in library catalogs: After all, many of them appear in handwritten collections which are often not cataloged in detail because there are simply not enough resources. If the cryptographical community continues to contribute valuable insights and increases the visibility of alchemical ciphers, more researchers might, in turn, come forward and share their findings in outlets such as this conference. If this venture is to be successful, we need to further nurture interdisciplinary collaboration in which cryptologists contribute their skills in cryptanalysis and work together with historians to interpret the results. In fact, historians are needed both before the decryption, for the historical contextualization of the ciphers to narrow down possibilities in the cipher type analysis stage, and after the decryption for the interpretation of the results as well as for helping to evaluate the role of the enciphering method detected in the historical context it belonged to. Promising examples for such collaborations, such as Bean et al. (2022) and Piorko et al. (2023), have shown the gold mine of potential there is in combining the cryptanalysis of alchemical ciphers with a close reading of their historical context. The more alchemical ciphers are known and have been studied in detail, the more conclusions can be drawn about the use of cryptography in alchemy in general. The examples studied thus far have proven themselves to be exceptionally rich historical sources for both the history of alchemy and the history of cryptography.

References

Manuel Bachmann and Thomas Hofmeier. 1999. *Geheimnisse der Alchemie*. Schwabe, Basel.

Richard Bean, Megan Piorko, and Sarah Lang. 2021. Deciphering the philosophers' stone: how we cracked a 400-year-old alchemical cipher. *The Conversation Media Group*.

Richard Bean, Sarah Lang, and Megan Piorko. 2022. Solving an alchemical cipher in a shared notebook of John and Arthur Dee. In *Proceedings of the 5th International Conference on Historical Cryptology HistoCrypt 2022*, number 188, pages 12–21. Linköping University Electronic Press.

Richard Bean, Corinna Gannon, and Sarah Lang. 2023. The cipher of Emperor Rudolf II's "alchemical hand bell". In *Proceedings of the 6th International Conference on Historical Cryptology (HistoCrypt 2023)*. Linköping University Electronic Press.

Donna Bilak. 2020. Chasing Atalanta. Maier, steganography, and the secrets of nature. In *Furnace and Fugue. A Digital Edition of Michael Maier's Atalanta fugiens (1618) with Scholarly Commentary*.

Sissela Bok. 1983. *Secrets. On the Ethics of Concealment and Revelation*. Vintage Books (Random), NY.

Augusto Buonafalce. 2006. Bellaso's reciprocal ciphers. In *Cryptologia*, volume 30, pages 39–51.

Michel Butor. 1990. *Die Alchemie und ihre Sprache. Original in: "Répertoire I", Les Editions de Minuit, Paris 1960*. Fischer, Frankfurt.

Stephen Clucas. 2017. The royal typographer and the alchemist: John Dee, Willem Silvius, and the diagrammatic alchemy of the *Monas Hieroglyphica*. *Ambix*, 64/2:140–156.

Karel Davids. 2005. Craft secrecy in Europe in the early modern period: A comparative view. *Early Science and Medicine. Openness and Secrecy in Early Modern Science*, 10/3:341–348.

William Eamon. 1994. *Science and the Secrets of Nature: Books of Secrets in Medieval and Early Modern Culture*. Princeton University Press, Princeton.

William Eamon. 2006. The 'secrets of nature' and the moral economy of early modern science. *Micrologus. Natura, scienze e società medievali. Nature, Sciences and Medieval Societies XIV. Il Segreto. The Secret*, pages 215–236.

William Eamon. 2013. On the skins of goats and sheep. (Un)masking the secrets of nature in early modern popular culture. In Timothy McCall, Sean Roberts, and Giancarlo Fiorenza, editors, *Visual Cultures of Secrecy in Early Modern Europe*, Early Modern Studies 11, pages 54–75, Kirksville Missouri. Truman State University Press.

William Eamon. 2016. How to read a book of secrets. In Elaine Leong and Alisha Rankin, editors, *Secrets and Knowledge in Medicine and Science, 1500–1800*, pages 23–46, NY. Routledge.

Florian Ebeling. 2001. 'Geheimnis' und 'Geheimhaltung' in den Hermetica der frühen Neuzeit. In Anne-Charlott Trepp and Hartmut Lehmann, editors, *Antike Weisheit und kulturelle Praxis. Hermetismus in der Frühen Neuzeit*, pages 63–80, Göttingen. Vandenhoeck und Ruprecht.

Umberto Eco, 2016. *Il discorso alchemico e il segreto differito*, pages 97–116. La nave di Teseo.

Gerhard Eis. 1965. Von der Rede und dem Schweigen der Alchemisten. In *Vor und nach Paracelsus. Untersuchungen über Hohenheims Traditionsverbundenheit und Nachrichten über seine Anhänger*, pages 51–73, Stuttgart. Gustav Fischer.

Katherine Ellison. 2017. *A Cultural History of Early Modern English Cryptography Manuals*. Routledge.

Peter J. Forshaw. 2013. Cabala chymica or chemia cabalistica—early modern alchemists and cabala. *Ambix*, 60/4:361–389.

Peter J. Forshaw. 2020. Michael Maier and mythoalchemy. *Furnace and Fugue. A Digital Edition of Michael Maier's* Atalanta fugiens *(1618) with Scholarly Commentary*.

Jonathan Gaede. 2017. Zur Verwendung astrologischer und alchemistischer Symbole in frühneuhochdeutschen Fachtexten. In Wolf Peter Klein, Matthias Schulz, Sven Staffelt, and Peter Stahl, editors, *Würzburger elektronische sprachwissenschaftliche Arbeiten (WespA) 19*, Würzburg.

Jan Golinski. 1990. Chemistry in the scientific revolution: Problems of language and communication. In David C. Lindberg and Robert S. Westman, editors, *Reappraisals of the Scientific Revolution*, pages 367–396, Cambridge. Cambridge University Press.

Michael Hunter. 2016. Robert Boyle and Secrecy. In Elaine Leong and Alisha Rankin, editors, *Secrets and Knowledge in Medicine and Science, 1500–1800*, pages 87–104, NY. Routledge.

Daniel Jütte. 2011. *Das Zeitalter des Geheimnisses. Juden, Christen und die Ökonomie des Geheimen (1400–1800).* Vandenhoeck & Ruprecht, Göttingen.

David Kahn. 1996. *The Codebreakers: The comprehensive history of secret communication from ancient times to the internet.* Simon and Schuster.

Sarah Lang and Megan Piorko. 2021. An alchemical cipher in a shared notebook of John and Arthur Dee (Sloane MS 1902) [work in progress]. In *Proceedings of the 4th International Conference on Historical Cryptology HistoCrypt 2021*, number 183, pages 90–93. Linköping University Electronic Press.

Wolfgang Lefèvre. 2018. The *Méthode de nomenclature chimique* (1787): A document of transition. *Ambix*, 65:1:9–29.

Elaine Leong and Alisha Rankin. 2016. Introduction: Secrets and knowledge. In Elaine Leong and Alisha Rankin, editors, *Secrets and Knowledge in Medicine and Science, 1500–1800*, pages 1–22, NY. Routledge.

Karma Lochrie. 1999. *Covert Operations. The Medieval Uses of Secrecy.* Philadelphia.

Pamela O. Long. 2001. *Openness, Secrecy, Authorship. Technical Arts and the Culture of Knowledge from Antiquity to the Renaissance.* The Johns Hopkins University Press, London.

Benedek Láng. 2015. Ciphers in magic: Techniques of revelation and concealment. *Magic Ritual and Witchcraft*, 10/2:125–141.

Benedek Láng. 2018. *Real Life Cryptology. Ciphers and Secrets in Early Modern Hungary.* Amsterdam University Press, Amsterdam.

Fritz Lüdy. 1928. *Alchemistische und chemische Zeichen.* Gesellschaft für Geschichte der Pharmazie, Stuttgart.

Michael Maier. 1609. *De medicina regia et verè heroica,* COELIDONIA. Prag.

Michael Maier. 1614. ARCANA ARCANISSIMA, *hoc est, Hieroglyphica Aegyptio-Graeca.* London.

Michael Maier. 1617. ATALANTA FUGIENS, *hoc est emblemata nova de secretis naturae chymica.* Oppenheim.

Matteo Martelli. 2016. Greco-Egyptian and Byzantine Alchemy. In Georgia L. Irby, editor, *A Companion to Science, Technology, and Medicine in Ancient Greece and Rome. Volume I*, pages 217–231, Oxford. Wiley Blackwell.

Marcos Martinón-Torres. 2011. Some recent developments in the historiography of alchemy. *Ambix*, 58/3:215–37.

Martin Mulsow. 2012. *Prekäres Wissen. Eine andere Ideengeschichte der Frühen Neuzeit.* Suhrkamp, Berlin.

William R. Newman and Lawrence M. Principe. 1998. Alchemy vs. chemistry: the etymological origins of a historiographic mistake. *Early Science and Medicine*, 3/1:32–65.

William R. Newman and Lawrence M. Principe. 2003. The chymical laboratory notebooks of George Starkey. In Frederic L. Holmes, Jürgen Renn, and Hans-Jörg Rheinberger, editors, *Reworking the Bench. Research Notebooks in the History of Science*, Archimedes. New Studies in the History and Philosophy of Science and Technology 7, pages 25–42, Dordrecht. Kluwer Academic Publishers.

William Newman. 1991. *The* Summa Perfectionis *of Pseudo-Geber. A Critical Edition, Translation, and Study.* Collection de Travaux de l'Académie Internationale d'Histoire des Sciences 35. Brill, Leiden.

William R. Newman. 1996. "*Decknamen* or pseudochemical language"? Eirenaeus Philalethes and Carl Jung. In *Revue d'histoire des sciences*, volume 49, pages 159–188.

William R. Newman. 2006. From alchemy to "chymistry". In Katharine Park and Lorraine Daston, editors, *The Cambridge History of Science 3 / Early Modern Science*, pages 497–517, Cambridge. Cambridge University Press.

William R. Newman. 2018. *Newton the Alchemist: Science, Enigma, and the Quest for Nature's "Secret Fire".* Princeton University Press, Princeton.

Tara E. Nummedal. 2007. *Alchemy and Authority in the Holy Roman Empire.* University of Chicago Press, Chicago.

Megan Piorko, Sarah Lang, and Richard Bean. 2023. Deciphering the *Hermeticae Philosophae Medulla*: Textual cultures of alchemical secrecy. *Ambix*, May.

Megan Piorko. 2019. Seventeenth-century chymical collections: A study of unique copies of fasciculus chemicus. *The Papers of the Bibliographical Society of America*, 113:409–445, December.

Lawrence M. Principe and William R. Newman. 2001. Some problems with the historiography of alchemy. In William R. Newman and Anthony Grafton, editors, *Secrets of Nature: Astrology and Alchemy in Early Modern Europe*, pages 385–432, Cambridge/Massachusetts. MIT Press.

Lawrence M. Principe. 1992. Robert Boyle's alchemical secrecy: Codes, ciphers and concealments. *Ambix*, 39/2:63–75.

Lawrence M. Principe. 2000. Daniel Georg Morhof's analysis and defence of transmutational alchemy. In Françoise Waquet, editor, *Maping the World of Learning: The Polyhistor of Daniel Georg Morhof*, Wolfenbütteler Forschungen 91, pages 139–153, Wiesbaden. Harrassowitz Verlag.

Lawrence M. Principe. 2011. *The scientific revolution: A very short introduction*. Oxford University Press, Oxford.

Lawrence M. Principe. 2013. *The Secrets of Alchemy*. The University of Chicago Press, Chicago.

Jennifer M. Rampling. 2020. *The Experimental Fire: Inventing English Alchemy, 1300–1700*. University of Chicago Press, Chicago.

Sara Reardon. 2011. The alchemical revolution. *Science*, 332:914–915.

Agnieszka Rec. 2014. Ciphers and secrecy among the alchemists: A preliminary report. In *Societas Magica Newsletter*, volume 31 (Fall), pages 1–6.

Martin Ruland. 1612. *Lexicon Alchemiae*. Frankfurt am Main.

Julius Ruska and E. Wiedemann. 1924. Alchemistische Decknamen. *Beiträge zur Geschichte der Naturwissenschaften LXVII, Sitzungsberichte der phys.-med. Sozietät Erlangen*, 56:17–36.

Georg Simmel, 1908. *Kapitel V: Das Geheimnis und die geheime Gesellschaft*, pages 256–304. Duncker & Humblot, Berlin.

Pamela Smith. 2016. What is a secret? Secrets and craft knowledge in early modern Europe. In Elaine Leong and Alisha Rankin, editors, *Secrets and Knowledge in Medicine and Science, 1500–1800*, pages 47–68, NY. Routledge.

Johann Christoph Sommerhoff. 1701. *Lexicon pharmaceutico-chymicum Latino-Germanicum et Germanico-Latinum*. Zieger & Lehmann & Froberg, Nuremberg.

Matteo Soranzo. 2020. *Giovanni Aurelio Augurello (1441–1524) and Renaissance Alchemy. A Critical Edition of* Chrysopoeia *and Other Alchemical Poems, with an Introduction, English Translation and Commentary*. Brill, Leiden.

George Starkey. 2004. *Alchemical Laboratory Notebooks and Correspondence*. University of Chicago Press, Chicago.

Joachim Telle. 1978. Alchemie II (historisch). *Theologische Realenzyklopädie 2*.

Hereward Tilton. 2003. *The Quest for the Phoenix. Spiritual Alchemy and Rosicrucianism in the Work of Count Michael Maier (1569–1622)*. De Gruyter, Berlin / NY.

Koen Vermeir and Dániel Margócsy. 2012. States of secrecy: An introduction. In *The British Journal for the History of Science. Special Issue: States of Secrecy*, volume 45/2, pages 153–164.

Koen Vermeir. 2012. Openness versus secrecy? Historical and historiographical remarks. In *The British Journal for the History of Science. Special Issue: States of Secrecy*, volume 45/2, pages 165–188.

Edmund Oskar von Lippmann. 1919. *Entstehung und Ausbreitung der Alchemie. Mit einem Anhange: Zur älteren Geschichte der Metalle. Ein Beitrag zur Kulturgeschichte. Band 1*. Springer, Berlin.

Eugen Wüster. 1991. *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie. 3. Auflage*. Romanistischer Verlag, Bonn.

Armand de Bourbon's

Poly-Homophonic Cipher – 1649

George Lasry
The DECRYPT Project
george.lasry@gmail.com

Abstract

We deciphered two letters from 26 and 27 March 1649, from Armand de Bourbon, Prince de Conti, a leader of the Fronde. Probably addressed to the marquis Louis II de La Trémoille-Noirmoutier, they discuss recent developments in the Parliament of Paris and in French provinces. The cipher is poly-homophonic, a combination of a homophonic cipher, where each plaintext letter may be represented by several cipher symbols, and a polyphonic cipher, where a cipher symbol may represent several plaintext letters. To the best of our knowledge, this is the only documented example of such a cipher.

1 Introduction

In Lachenicht and Braun (2021, p.87), Camille Desenclos mentions a cipher, from Armand de Bourbon, Prince de Conti.[1] Two letters using this cipher have been identified in the Bibliothèque Nationale de France, Français 3584 f.113r, f113v, and f.115.[2] The second letter (f.115) is shown in Figure 1. It contains fragments of French cleartext, such as "A Paris ce 27 Mars 1649", "j'ay cru vous debvoir", or the signature "Armand de Bourbon". The rest is in cipher, using lower-case letters to encipher the original text. Those letters had not been deciphered prior to this present work.

We present, in Section 2, the process of recovering the cipher key and of deciphering the letters. Section 3 provides a short historical background. The deciphered text and its translation are presented in Section 4. Some concluding remarks are given in Section 5.

2 Deciphering the letters

We first transcribed the two documents, then applied to those documents (combined) a computerized codebreaking algorithm developed by the CrypTool 2 team to solve homophonic ciphers, obtaining a tentative decryption and an initial key.[3] With additional manual work, as well as linguistic analysis of tentative decryptions, we were able to reconstruct the cipher key and complete the decipherment of the letters.

2.1 Initial computerized codebreaking

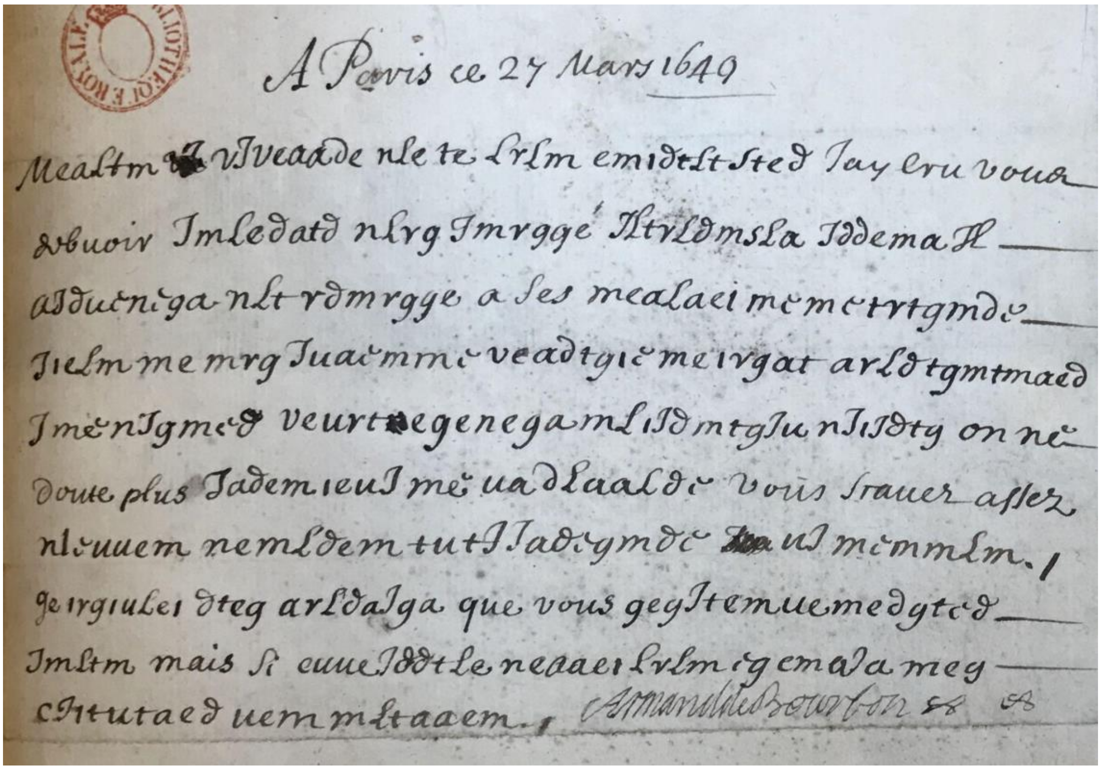
Using the computerized codebreaking algorithm, and assuming that the cipher was homophonic, we obtained the initial key shown in Figure 2. Interestingly, two homophones are assigned to each of the plaintext letters L and R. All the other plaintext letters have only one homophone, which is not typical of contemporary homophonic ciphers, where the vowels and the most frequent letters were usually assigned more than one cipher symbol. With this initial key, we obtained an initial decipherment of the second letter, shown in Figure 3.[4]

[1] Armand de Bourbon, Prince of Conti (1629 – 1666). French nobleman, brother of the Grand Condé. His sister married Henri II d'Orléans-Longueville. He was at time a patron of Molière, the famous French playwright and actor, turning later against him on religious grounds.

[2] Additional letters encoded with this cipher may exist.

[3] The algorithm is described in Kopal (2021).

[4] Due to lack of space, we do not show the decipherment of the first letter.



A Paris ce 27 Mars 1649

Armand de Bourbon

Figure 1 - Second letter - 27 March 1649 (Source: gallica.bnf.fr / BnF fr. 3584 f.115)

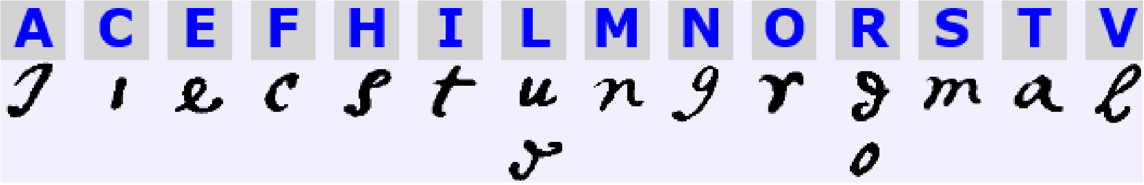
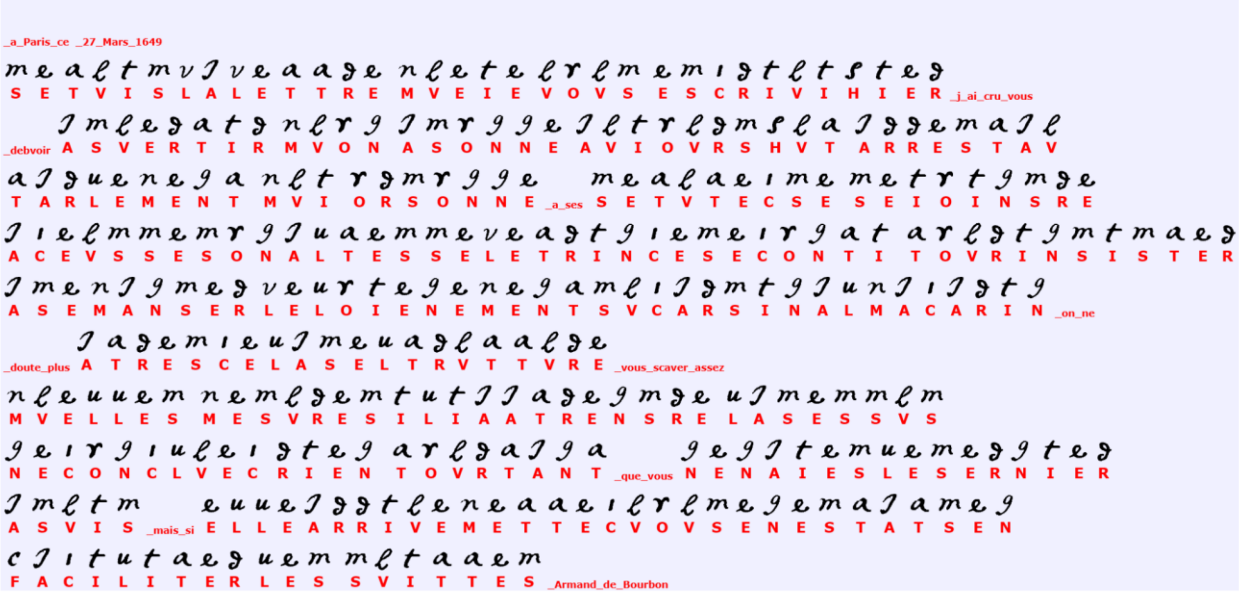

A
C
E
F
H
I
L
M
N
O
R
S
T
V

Figure 2 – Initial key



_a_Paris_ce _27_Mars_1649

S E T V I S L A L E T T R E M V E I E V O V S E S C R I V I H I E R _j_ai_cru_vous

_debvoir A S V E R T I R M V O N A S O N N E A V I O V R S H V T A R R E S T A V

T A R L E M E N T M V I O R S O N N E _a_ses S E T V T E C S E S E I O I N S R E

A C E V S S E S O N A L T E S S E L E T R I N C E S E C O N T I T O V R I N S I S T E R

A S E M A N S E R L E L O I E N E M E N T S V C A R S I N A L M A C A R I N _on_ne

_doute_plus A T R E S C E L A S E L T R V T T V R E _vous_scaver_assez

M V E L L E S M E S V R E S I L I A A T R E N S R E L A S E S S V S

N E C O N C L V E C R I E N T O V R T A N T _que_vous N E N A I E S L E S E R N I E R

A S V I S _mais_si E L L E A R R I V E M E T T E C V O V S E N E S T A T S E N

F A C I L I T E R L E S S V I T T E S _Armand_de_Bourbon

Figure 3 – Initial decryption of the second letter

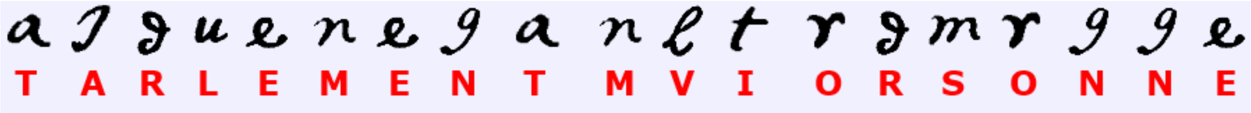


Figure 4 – Fragment of deciphered text with errors – Example 1

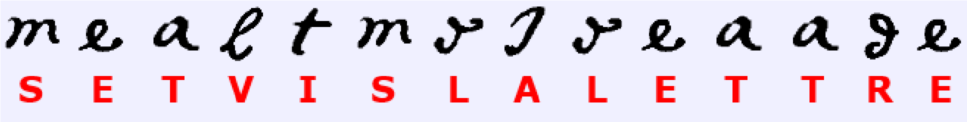


Figure 5 – Fragment of deciphered text with errors – Example 2

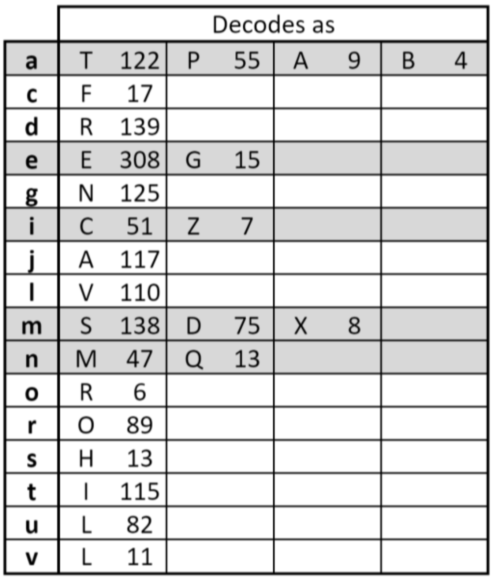
Most of the deciphered text consists of plausible fragments of French, or full French words or expressions. For example, the last line FACILITERLESSVUITES reads as “faciliter les suites” (“to facilitate the follow-up”). But many other fragments seem to contain one or more errors, which cannot be easily resolved. For example, the beginning of the fragment shown in Figure 4, “TARLEMENT” seems to be “PARLEMENT”. But if we simply try to assign the homophone ‘a’ to the letter P instead of the current assignment to T, we obtain “PARLEMENP” which is also wrong.

Similarly, the fragment “SETUISLALETTRE” shown in Figure 5 seems to read “depuis la lettre” (“since the letter”). But if we try to assign the homophone “m” to the letter D instead of to S, and we also try to assign ‘t’ to P instead of to T, we obtain “DEPUIDLALEPPRE” (“depuid la leppre”), which is also wrong.

So obviously, under the assumption that this is a purely homophonic cipher, there is no way to “fix” the key so that all those errors (we spotted over 150 such discrepancies) may be corrected.

2.2 Manual decipherment

To determine the precise structure of the cipher, the next step was to recover the original plaintext, starting from the parts that looked fully or partially plausible. After extensive trial-and-error, we were able to recover most of the original text. We then counted the number of times each plaintext letter is represented by a certain cipher symbol, as shown in Figure 6. For example, the cipher symbol ‘a’ represents the letter T 122 times, the letter P 55 times, the letter A nine times, and the letter B four times.

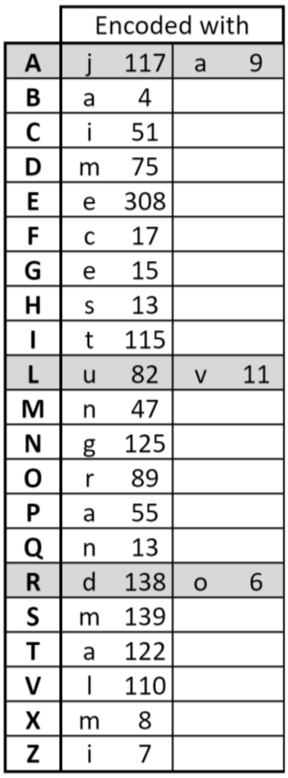


	Decodes as							
a	T	122	P	55	A	9	B	4
c	F	17						
d	R	139						
e	E	308	G	15				
g	N	125						
i	C	51	Z	7				
j	A	117						
l	V	110						
m	S	138	D	75	X	8		
n	M	47	Q	13				
o	R	6						
r	O	89						
s	H	13						
t	I	115						
u	L	82						
v	L	11						

Figure 6 – Evidence for polyphony

From the segment in Figure 4, we had already concluded that ‘a’ could either represent T or P. With the complete data in Figure 6, we can see that ‘a’ may also represent A or B. Similarly, from the segment in Figure 5, we had already concluded that ‘m’ could represent either S or D, and in Figure 6 we see that it may also represent X. In summary, for five symbols (‘a’, ‘e’, ‘i’, ‘m’, and ‘n’), there are two to four possible interpretations, which is consistent with a polyphonic cipher.

We also produced the reversed analysis, showing the breakdown of which cipher symbols are used to encode a certain plaintext letter, as shown in Figure 7.



	Encoded with			
A	j	117	a	9
B	a	4		
C	i	51		
D	m	75		
E	e	308		
F	c	17		
G	e	15		
H	s	13		
I	t	115		
L	u	82	v	11
M	n	47		
N	g	125		
O	r	89		
P	a	55		
Q	n	13		
R	d	138	o	6
S	m	139		
T	a	122		
V	l	110		
X	m	8		
Z	i	7		

Figure 7 – Evidence for homophony

Figure 7 shows that each of the three plaintext letters, A, L, and R, has two homophones.

While the evidence for polyphony is somehow stronger than the evidence for homophony,[5] there is enough evidence to establish that we have here a new type of cipher, a poly-homophonic cipher, which combines the two types.

We show in Figure 8 the final decipherment of the second letter. There are still a few errors, such as “AVIOURDVT” which should be “AVIOURDVI” (“aujourd’hui”, today) on the second line, but those kinds of sporadic errors are expected in any enciphered document.

The complete decipherment, after correcting the remaining errors, and formatting the text, is shown in Section 4.

2.3 The cipher key

We show the final key in Figure 9. As described in Lachenicht and Braun (2021), most French ciphers in the 16th and 17th centuries were homophonic. Polyphonic ciphers were less prevalent, and those documented were primarily used by papal nuncios in the 16th century (Meister 1906). An example of a polyphonic French cipher, used by the Duc de Mayenne, is given in Tomokiyo (2019). In addition, most contemporary ciphers also had a nomenclature, with additional symbols to encode entire words, names, and parts of words.

In contrast with most contemporary ciphers, the cipher used by Armand de Bourbon has no nomenclature, and it combines polyphony with homophony. The author is not aware of any other example of such a poly-homophonic cipher. While more secure, such a cipher would also have been difficult to employ when enciphering a text, and even more difficult when deciphering a ciphertext, even if the key was known to both parties.

One of the reviewers of this paper has noticed that the second column of Figure 7 may be read as a keyword (or key expression), as follows: *jaimecestungrandmal* (*J’aime c’est un grand mal*).[6] This happens to be the title of an *air de cour* composed by Antoine Boesset in 1642.[7] The year is too close to the time the letters were sent and the key expression too long to be a coincidence. So, it is very likely that the two correspondents relied on this shared expression to exchange or remember the key.

[5] One may argue that this cipher is only weakly homophonic. On the one hand, there are only three letters of the alphabet represented by more than one cipher symbol, and among them, the use of the homophone ‘a’ to encipher the letter A might be due to confusion – the same letter being wrongly enciphered with itself. On the other hand, this may also indicate a “lazy” use of the homophones, the person enciphering the letter almost always using only one of the two homophones assigned to each letter.

[6] Loosely translated as *I love this hurts badly*.

[7] According to Wikipedia, the *air de cour* was a popular type of secular vocal music in France in the late Renaissance and early Baroque period, from about 1570 until around 1650.

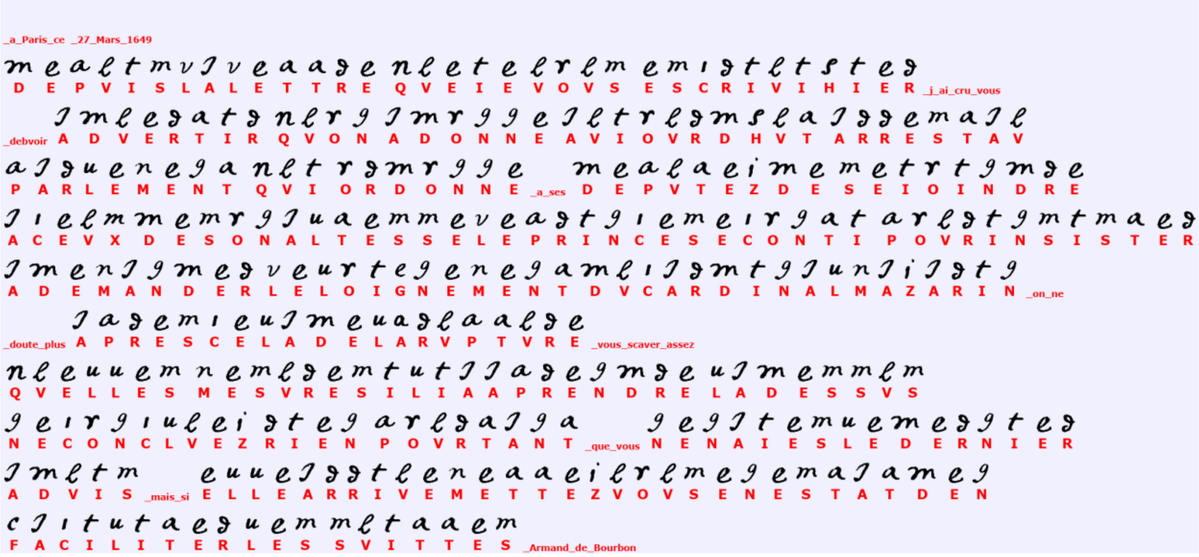

_a_Paris_ce _27_Mars_1649
DEPVISLALETTREQVEIEVOVSESCRIVIHIER _j_ai_cru_vous
_debvoir ADVERTIRQVONADONNEAVIOVRDHVTARRESTAV
PARLEMENTQVIORDONNE _a_ses DEPVTEZDESEIOINDRE
ACEVXDESONALTESSELEPRINCESECONTIPOVRINSISTER
ADEMANDERLELOIGNEMENTDVCARDINALMAZARIN _on_ne
_doute_plus APRESCELADELARVPTVRE _vous_scaver_assez
QVELLESMESVRESILIAAPRENDRELADESSVS
NECONCLVEZRIENPOVRTANT _que_vous NENAIESLEDERNIER
ADVIS _mais_si ELLEARRIVEMETTEZVOVSENESTATDEN
FACILITERLESSVITTES _Armand_de_Bourbon

Figure 8 – Final decryption of the second letter

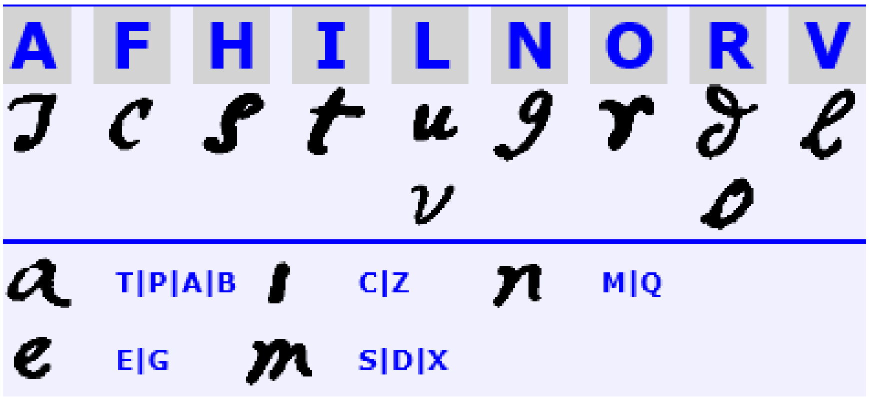


Figure 9 – Final key

3 Historical background

Before presenting the full deciphered text, we provide here a short historical background about the Fronde, and the Parliamentary Fronde (the First Fronde) in particular.[8]

3.1 The Fronde

The Fronde was a series of civil wars and insurrections in France between 1648 and 1653, occurring during the Franco-Spanish War, which had begun in 1635, and right after the Thirty-Year War, which ended in 1648. Minor King Louis XIV (1638-1715), his regent mother Anne of Austria,[9] and Cardinal Mazarin[10] confronted the combined opposition of the princes, the nobility, and the Parliaments, but eventually managed to subdue them all. The dispute started when the government of France issued fiscal edicts to increase taxation. The Parliament of Paris resisted and sought to check the King's powers. The Fronde was divided into three phases, the Parliamentary Fronde (1648-49), the Fronde of the Princes (1650-1651), and the Fronde of Condé (1651-1652). Cardinal Mazarin blundered into the crisis but came out well ahead at the end.

3.2 The First Fronde – the Parliamentary Fronde

The First Fronde – the Parliamentary Fronde, started in May 1648 when a tax levied on judicial officers of the Parliament of Paris provoked not

[8] The information in this background section was mostly taken from Wikipedia.

[9] Anne of Austria (1601 – 1666) was an infanta of Spain who became Queen of France as the wife of King Louis XIII from their marriage in 1615. When Louis XIII died in 1643, Anne became regent to her son Louis XIV, during his minority, until 1651. During her regency, Cardinal Mazarin served as France's chief minister.

[10] Cardinal Jules Mazarin (1602 – 1661), born Giulio Raimondo Mazzarino or Mazarini, was an Italian cardinal, diplomat and politician who served as the chief minister to the Kings of France Louis XIII and Louis XIV from 1642 to his death.

merely a refusal to pay but also a condemnation of earlier financial edicts.[11] In August 1648, Mazarin suddenly arrested the leaders of the Parliament, whereupon Paris broke into insurrection and barricaded the streets. The noble faction demanded the calling of an assembly of the Estates General. The royal faction, having no army at its immediate disposal, had to release the prisoners and to promise reforms. However, France's signing of the Peace of Westphalia allowed the French army to return from the frontiers,[12] and by January 1649, Mazarin’s ally the prince de Condé had put Paris under siege.[13] The Parliament’s legalist faction led by the first president Mathieu Molé and the president Henri de Mesmes pushed for negotiations. The two warring parties signed the Peace of Rueil (11 March 1649) after little blood had been shed, followed by the Peace of Saint-Germain (1 April 1649). The Parisians, under the military leadership of Armand de Bourbon, Condé’s brother, having refused an offer of help from Spain, but with no prospect of military success without such external aid, eventually submitted to the government while receiving some concessions.

The two letters we deciphered were written after the Peace of Reuil, and a few days before the signature of the Peace of Saint-Germain.

4 The deciphered letters

The parts in cipher are in *italics*.

4.1 First letter

A Paris ce 26 mars 1649

Affin que Monsieur *Le Monsieur de Noirmo[u]stier*[14] *sache l'estat* des choses *positivement* comme *elles se passent à St Germain,*[15] il sera *averti* que *l'on insiste* fort dans *la conférence sur les intérêts* du *parlement de Normandie* avec lesquels le *parlement de Paris [est?] tellement joint. Que* hier *le premier président* et *le président de Mesmes* déclarèrent que si *le parlement de Normandie* n'estoit *content l'on romproit.* Les intérêts *du dit parlement sont grants* ce qui faict *croire la rupture. Que l'on insistera* fort *encore sur l'exclusion du Cardinal les députez* de Messieurs *les généraux* en ayant *eu ordre exprès*. Que l'on *demande encor positivement* et sans *qu'on en puisse relascher* que *Monsieur de Longueville*[16] *traittera* avec *Monsieur de Tenerande(?).* Que sur ces *articles on attend la rupture de laquelle on parleroit p[lu]s positivement si son Altesse* n’agessoit pas avec quelque *dépendance. Que véritablement le parlement a bien accordée la t[r]êve* mais *avec ordre positif de ne continuer* pas la *conférence après quatre jours espiréz* qui est le *temps que doit durrer la prolongation* et que *les députez s’en reviendront* dans lequel *temps les choses pouvant* pas vraisemblablement *s‘ajuster, cela donne pour lieu* et croire que *la conférence se rompra.* Que l’on a esté *fort surpris de la retraite de* Monsieur *l'archiduc*[17] *du Pont-a-Vert* et l’abandonnement *des passages de la rivière*, Monsieur *le maréchal du Plessis*[18] *n'ayant que de médiocres forces et Erlac*[19] *s'avançant avec si peu de monde.* Que *cette retraite a fort mal réussi* et faict un fort *grand tort aux affaires* mais que n'y *ayant nul danger de reprendre un passage sur la rivière* cela redonneroit *chaleur aux affa[i]res et remettroit*

[11] In January 1648, Mazarin had issued those fiscal edicts, but the Parliament of Paris decided to ignore them. To convince this parliament to withdraw its opposition, Mazarin exempted it from paying for the renewal of the “Paulette”. In an exceptional display of solidarity, the sovereign courts, which included the various parliaments, decided on 13 May 1649 to convene in the Palace of Justice, where the Parliament of Paris resided, starting the Parliamentary Fronde.

[12] Treaty of Münster, 24 October 1648.

[13] Louis de Bourbon, Prince de Condé (1621 – 1686), known as le Grand Condé for his military exploits, was a French general and a member of the Condé branch of the House of Bourbon, and the brother of Armand de Bourbon. Having fought on the side of the French court during the First Fronde, we rebelled against Louis XIV as the leader of the last Fronde in 1651, leading to his exile from France until 1659 when he was rehabilitated.

[14] Louis II de La Trémoille, marquis, later Duc de Noirmoutier (1612-1666), often simply called « Noirmoutier », was a nobleman and general who joined Armand de Bourbon and Longueville on the Parliament side during the First Fronde.

[15] In January 1649, the Queen Mother and King Louis XIV had fled from Paris to St. Germain, ordering that Paris be put under siege.

[16] Henri II d'Orléans, Duc de Longueville or Henri de Valois-Longueville (1595 – 1663), a prince of France of royal descent, was a major figure during the Fronde, and served as governor of Picardy, then of Normandy. He married Armand de Bourbon’s sister.

[17] Archduke Leopold Wilhelm of Austria (1614 – 1662), younger brother of Emperor Ferdinand III, was an Austrian soldier, administrator, and patron of the arts. He served as Governor of the Spanish Netherlands and offered his help to the First Fronde leaders in Paris. Despite being nominated as Holy Roman Emperor after Ferdinand's death in 1657, he stood aside in favor of his nephew Leopold I.

[18] César, Duc de Choiseul, Comte du Plessis-Praslin (1602 – 1675) was a Marshal of France and French diplomat, loyal to the French court.

[19] Jean Louis d'Erlach (1595–1650) was a Swiss general and politician, who supported the French court during the First Fronde.

tout en *bon estat* ce que l'on attend *absolument* de son *altesse impér[iale].* Que *dimanche ou lundi* on ne manquera *de vous donner advis de la dernière résolution des choses* qu'on espère telle qu'on la peut désirer y ayant *peu de lumière à l'accomodement* mais qu'au cas *qu'il se fist ce ne sera point sans faire tous les effors pour tirer* les assurances *de la p[a]ix générale.* Les *provinces sont en cet estat*: Monsieur *de Longueville a dix(six?)*[20] *mil hommes.* Que *Thoulouse* a donné *l'arrest*. Que *Bordeaux va le donner.* Que la *Provence* est en *armes avec le Poitou le Périgueux de Q[u]ercy de Limosin et* quantité *de lieus e[n] B[r]etaigne et le peuple se pris* mieux intentionné que jamais.

Paris, 26 March 1649

For Monsieur, Monsieur de Noirmoustier to positively know the state of affairs as they took place in St Germain, he should be informed that we are strongly insisting in the conference on the interests of the Parliament of Normandy, which the Parliament of Paris is joining. That yesterday, the First President and de Mesmes the President have declared that unless the Parliament of Normandy is satisfied, there will be a rupture. As the interests of the said Parliament have been granted, this rupture is expected. That we will strongly insist on the exclusion of the Cardinal, the deputies of Messieurs, the generals having been given an explicit order. That we are still positively demanding, without being able to give up, that Monsieur de Longueville deal with Monsieur de Tenerande(?). That because of those articles, a rupture is expected, which is seen more positively, unless his Highness acts with some dependence. That the Parliament has indeed accorded a truce, but with a positive order not to continue the conference for more than four additional days (which is the time of the prolongation), and the deputies will return to it, and as things are not likely to be settled by then, it is expected that the conference will adjourn. That there has been a great surprise about Monsieur the Archduke's retreat in Pont-a-Vert, and the abandonment of the crossings of the river, Monsieur the Marechal de Plessis having only some mediocre forces at his disposition, and Erlach advancing with so few men. That this retreat has badly failed, and has strongly harmed the affairs, but as there is no danger of attempting to cross the river, this would enable the affairs to be brought back to a good state, which is expected from his Imperial Highness. That Sunday or Monday, we will not fail to notify you of the latest resolution on the matters, hoping to have a better understanding about the accommodation, but in case this happens, this will not be without making every effort to obtain the assurances of the general peace. The provinces are in the following situation: Monsieur de Longueville has ten (or six?) thousand men. That Thoulouse has given the ruling. That Bordeaux is about to give it. That the Provence has taken arms with the Poitou, the Périgueux de Quercy, the Limosin, and many places in Brittany, and the people is better intentioned than ever.

4.2 Second letter

A Paris, le 27 mars 1649

Depuis la lettre que je vous escrivi hier j'ay creu vous debvoir *advertir qu'on a donné aujourd'hui arrest au parlement qui ordonne* à ses *députez de se joindre à ceux de son Altesse le Prince de Conti à demander l'éloignement du cardinal Mazarin.* On ne doute plus *après cela de la rupture.* Vous scavez assez *quelles mesures il y a à prendre là-dessus. Ne concluez rien pourtant* que vous *n'ayes le dernier advis* mais si *elle arrive mettez-vous en estat d'en faciliter les suittes.*

Armand de Bourbon

Paris, 27 March 1649

Since the letter I wrote to you yesterday, I ought to notify you that today, a ruling has been given in Parliament, which orders its deputies to join her Highness the Prince de Conti in demanding that Cardinal Mazarin be removed. There is no doubt that this will result in the rupture. You know well enough which measures should be taken about this. Nevertheless, do not conclude anything before you get the latest notice, but if this happens, be ready to facilitate what follows.

Armand de Bourbon

[20] Ambiguous, because the symbol for S is the same as the symbol for D. An example of the challenge of deciphering a polyphonic cipher, even when the key is known.

5 Conclusion

This cipher illustrates an interesting combination of two schemes of substitution ciphers, and there are no other known examples of such a poly-homophonic cipher. Also, while the letters of Armand de Bourbon are of special interest for research on the development of ciphers, their contents may also be of interest to historians.

Acknowledgments

The author would like to thank Camille Desenclos from the University of Picardie – Jules Verne for bringing this interesting cipher to light, providing copies of the documents, assisting in the interpretation of the text, and reviewing an early draft of this paper.

The author would also like to thank the anonymous reviewer for the astounding discovery of the key expression in the homophonic key table.

Funding

This work has been supported by the Swedish Research Council, grant 2018-06074, DECRYPT – Decryption of Historical Manuscripts.

References

Alois Meister. 1906. *Die Geheimschrift im Dienste der Päpstlichen Kurie von Ihren Anfängen bis zum Ende des XVI. Jahrhunderts*, vol. 11. Paderborn: F. Schoningh.

Susanne Lachenicht and Guido Braun (ed.). 2021. *Spies, Espionage and Secret Diplomacy in the Early Modern Period*. Kohlhammer Verlag.

Nils Kopal. 2019. “Cryptanalysis of homophonic substitution ciphers using simulated annealing with fixed temperature." *Proceedings of the 2nd International Conference on Historical Cryptology*, HistoCrypt.

Tomokiyo, S. 2019-2022. *A Polyphonic Substitution Cipher of the Catholic League (1592-1593)*, Cryptiana. Accessed December 22, 2022. http://cryptiana.web.fc2.com/code/mayenne.htm

International Conference on the Voynich Manuscript 2022

Colin Layfield
University of Malta
colin.layfield@um.edu.mt

John Abela
University of Malta
john.abela@um.edu.mt

René Zandbergen
Independent Researcher
rene.zandbergen.rz@gmail.com

Claire Bowern
Yale University
claire.bowern@yale.edu

Lisa Fagin Davis
Medieval Academy of America
lfd@themedievalacademy.org

Michael Rosner
University of Malta
mike.rosner@um.edu.mt

Lonneke van der Plas
Idiap Research Institute
lonneke.vanderplas@idiap.ch

Abstract

The International Conference on the Voynich Manuscript, which took place 30 November and 1 December 2022, was the first peer-reviewed conference that was dedicated entirely to the Voynich MS. The only similar event took place ten years earlier at Villa Mondragone, Frascati, Italy, with invited presentations but without published proceedings (Schmeh, 2013). This paper summarises the event, its preparation and organisation, with a summary of the papers that were presented and potential avenues for further research.

1 Introduction

The Voynich Manuscript is an enigmatic medieval mystery. It is (widely believed) to be a 15th century tome which is written in a unknown script, in an unknown language whose contents, despite the efforts of many experts over the last century, has yet to be deciphered. Within its pages lie fantastical figures and illustrations of plants (many unidentified as of yet, if they exit at all) and people, in addition to drawings of an astronomical nature, that baffle interpretation. This conference, and the research presented, is a reflection of the continued interest in this manuscript and the ongoing efforts to unravel its mysteries (or, determine if there is a mystery at all, it may also be a hoax). A good starting point for anyone interested in gaining more knowledge on the manuscript in general is René Zandbergen's informative website on the topic, http://www.voynich.nu.

The informal group of researchers that make up the "Voynich Research Group" had their first meeting in May 2019. At that point participants were Michael Rosner, Lonneke van der Plais and Colin Layfield at the University of Malta: a group of academics who had a common interest in the mystery that is the Voynich Manuscript. In 2020 John Abela joined our group and it grew further, as the need for additional outside expertise was apparent. The initial members were from the fields of Artificial Intelligence and Computational Linguistics, but proper Voynich research would need historical and mediaeval studies expertise too. René Zandbergen, Lisa Fagin Davis and Claire Bowern joined the group in the period 2020–2021.

The idea of a conference on the Voynich Manuscript was first proposed in September of 2021. It was agreed this would be an interesting venture and an opportunity to see what the larger Voynich community has been working on, and that an online format was the most appropriate due to the uncertainty around the COVID-19 global situation at the time.[1] The review process used a 2 phase double blind process where 3 reviewers (who would not know who the submitting author(s) were) would be assigned to each paper, initially reviewing an abstract. If this passed the first review, the author(s) would be invited to submit a full paper where the same reviewers would

[1] As an aside, the VM has a tenuous link with Malta due to the fact it is believed that Father Strickland (a Jesuit priest from Malta) acted as some sort of intermediary between the Jesuits of Villa Mondragone and Voynich himself resulting in the purchase of the VM.

subject it to a second review and, if successful, the author(s) would be invited to participate in the conference to present their work.

The call for papers generated a total of 32 submissions. At the end of the process 16 papers were selected for admission to the conference. They covered a wide spectrum of topics and approaches to investigating the Voynich, which we will explore below. The conference also had two well respected keynote speakers: René Zandbergen opening the proceedings and Lisa Fagin Davis closing the conference, in addition to the welcome presence of Ray Clemens, curator at the Beinecke Library at Yale where the Voynich is preserved as MS 408, who opened the conference with a few words.

The conference covered two half days, allowing convenient attendance from East Asia (evening times) to the US West Coast (early morning times). Times listed in this paper are in Central European Time. Each day was split into two sessions of four presentations each. All selected presentations were pre-recorded on video, thus ensuring that the programme was followed precisely according to the planned timeline.

The following sections will explore the presentations at the conference and the insights they have offered. The papers are all available online at the conference proceedings website[2] (Layfield and Abela, 2022). All of the presentations were prerecord and the videos of the paper presentations can be found on our YouTube channel[3].

2 Conference Presentations

The following sections provide a summary of the papers presented at the conference, organized by session.

2.1 The Keynote Presentations

In his opening keynote, *Transliteration of the Voynich MS text*, René Zandbergen outlined both old and new aspects of the data that is used as input in all studies of the MS text. After a brief historical introduction, he indicated issues with the accuracy of existing transliterations, and introduced new representation formats and tools that facilitate the handling of these data. This included in particular a new transliteration alphabet that combines all features of the existing alphabets, and allows direct comparison of all existing transliterations (Zandbergen, 2022).

In her closing keynote, *Voynich Paleography*, Lisa Fagin Davis demonstrated how the principles of Latin Paleography may be profitably applied to the Voynichese writing system, resulting in the identification of distinctive features of five scribes. In addition, she suggested that several symbols may be abbreviations or ligatures and proposed how to interpret these as such (Davis, 2022).

2.2 Session One, 30 November 2022, 14:15 – 16:15

The first session of the conference had an eclectic mix of papers. The topics discussed included; a comparison of the Voynich to Sloane MS 3188 (Enochian constructed language), the interpretation of the tent-like illustrations in the Voynich, the role that the gynaecological and sexological content in the Voynich played in the (possible) encipherment of the manuscript, and whether it can be argued that the Voynich is too non-random to be gibberish (Boxer, 2022).

The first paper of the session, *Fingerprinting Gibberish: A Quantitative Comparison of the Voynich and Sloane MS 3188*, by Alexander Boxer, examines the Enochian test of the Sloane MS 3188 to the VM. Sloane MS 3188 is an important corpus of gibberish text that is just over a hundred years younger than the VM. In the course of his work, the author also created, and made freely available, a new transcription of the Voynich and, using the new transcription, compared the VM to the Sloane MS 3188. The author concluded that, although there are a number of qualitative similarities between the two manuscripts there are also substantial statistical differences and even if both manuscripts are gibberish they do not belong to the same type of gibberish (Gheuens and Rapaport, 2022).

The second paper, *Above and Beyond Voynich Canopies: Tents as a Recurring Motif in Beinecke MS 408*, by Koen Gheuens and Cary Rapaport, investigated the presence, and possible interpretation, of a recurring motif of tent-like structures (or canopies). The authors compared two groups of tent-like structures from different sections of the manuscript and compared them to those in contemporary mediaeval images. The authors hypothesised that the tent-like structures in the VM were likely inspired by tensile architecture from the pe-

[2] http://ceur-ws.org/Vol-3313/

[3] http://youtube.com/@VoynichResearchGroup

riod. They also argued that Voynich tent images incorporate the visual metaphor of the sky as a tent and emphasised that this symbolic aspect is important for understanding what these images represented to a mediaeval audience.

In the third paper, *'I beg your grace that you suppress this chapter or else allow it to be written in secret letters': The emotions of encipherment in late-medieval gynaecology*, by Keagan Brewer investigated the emotions involved in late-medieval gynaecology and sexology. In particular, Brewer argued that many of the illustrations in the VM cross mediaeval lines of taboo and that concerns about certain taboo subjects may have served as motivation for the encipherment of the VM. The paper makes several references to examples of encipherment, erasure, and self-censorship in gynaecological and sexological texts of the period (Brewer, 2022).

The final paper of Session One, *Gibberish after all? Voynichese is statistically similar to human-produced samples of meaningless text*, by Daniel E. Gaskell and Claire L. Bowern, argued that gibberish text does not have to be (statistically) random. If the Voynich manuscript indeed contains only gibberish then it may not have been produced by a random process but by some methods that created meaningless text that was meant to look like natural language. The authors recruited 42 volunteers to write gibberish text and compared the resulting text against the VM and linguistically meaningful texts. The authors argued that the results obtained refute the idea that the low-level structure of the VM is too non-random to be meaningless (Gaskell and Bowern, 2022).

2.3 Session Two, 30 November 2022, 16:15 – 18:15

Two of the four papers in session two addressed the possibility that the VM is an encrypted text whose statistical properties may be different from the original plain text. The other two respectively focused on whether computational linguistic techniques applied to VM can validly reveal insights concerning, its authorship, or its linguistic structure. The order of presentation in the programme did not exactly follow these thematic groupings, as indicated explicitly below.

Opinions differ about whether VM is an enciphered natural language or whether it is gibberish. Its predictability at character level is often cited as support for the latter argument. However, this argument assumes that unusual predictability at character level implies that the text as a whole is gibberish because it lacks the kind of higher-level structure that characterises meaning-bearing language. *Enciphered after all? Word-level text metrics are compatible with some types of encipherment* by Claire Bowern and Daniel Gaskell (second paper of session) pours some doubt on this argument by asking whether ciphers exist that produce the textual characteristics that make Voynichese unusual at the character level, whilst preserving higher level topic structure across larger segments of text. To investigate this, 22 methods of textual manipulation were unleashed on samples of genuine historical and contemporary NL text and the results were compared, using a similarity metric based on statistical properties, with VM, and with the output of a gibberish construction method. The textual manipulations indeed produced a range of effects on the similarity measurements. Several produced outcomes that are similar to Voynich text according to at least some metrics, whilst showing differences according to others. The authors concluded that the unusual word-level predictability highlighted in previous work is not conclusive evidence that the Voynich manuscript is gibberish (Bowern and Gaskell, 2022).

Polygraphia III: The cipher that pretends to be an artificial language by Jürgen Hermes (third paper of session) provided additional evidence in this vein by examining a historical example. A cipher where individual letters are replaced with invented words that are very similar to each other has the potential to generate text whose statistical properties (using joint entropy, distribution of word length and similar words), are very similar to those of Voynichese. This kind of cipher is found in the first printed book on cryptology, namely the Polygraphia written by Johannes Trithemius (Trithemius, 1518), who described procedures based on the fact that the letters of the plain text occur at specified positions of the cipher text, the rest being filled with nulls. The creator of the message must generate an extremely large amount of inconspicuous text, and Polygraphia III was the third of a series of refinements to the original idea which reduced the effort by exploiting the fact that many similar words have the same stem (Hermes, 2022).

Demystifying the scribes behind the Voynich Manuscript using Computational Linguistic Techniques by Kevin Farrugia, Colin Layfield and Lonneke van der Plas (first paper of the session) attempted to provide a computational validation of Lisa Fagin Davis' hypothesis, based on established palaeographic techniques, that the VM is the work of five different scribes. This is achieved using machine learning to train several classifiers based on character sequences using Davis' original classification as a gold standard. Training was performed on 90% of the corpus, and testing on the remaining 10%, repeating the exercise for different splits of the corpus (10-fold cross validation). The authors concluded that there was a reasonable overlap between the classifier predictions and the ground truth taken from palaeographic work. However there were also some anomalies (e.g. cases where "all classifiers agreed with one another but not with Dr Davis") suggesting that further research would be necessary to perfect the choice of data upon which classification is made (Farrugia et al., 2022).

In *An Analysis of the Relationship between Words within the Voynich Manuscript* Andrew Caruana, Colin Layfield and John Abela (fourth paper of the session) investigated the presence of linguistic structure within VM by analysing various properties of word-pairs found in the manuscript as well as in other works written in natural languages such as the Bible, Dante's *La Divina Commedia*, and Shakespeare's *Macbeth* and *Julius Caesar*. An analysis of the order of words in the word-pairs indicated many 'skewed pairs' whose words were more likely to appear in one order than the other. The ratio of the number of skewed pairs to all pairs in each work was plotted, along with the same ratio for random shuffles of each work. The results indicated that there was a substantial difference in all natural language documents between their normal and shuffled counterparts. The difference was not as large within the Voynich Manuscript but the word-pair occurrence ratio of the original was still considerably higher than the ratio of the shuffled manuscript. The authors concluded that this could indicate that the Voynich Manuscript is not random text but may be a language or a cipher (Caruana et al., 2022).

An overarching theme underlying all but the last paper is the enduring debate on whether VM is gibberish. Unsurprisingly, the question was not conclusively answered but during the session some light was shed on the extent to which, paradoxically, ciphers of meaningful text can display surface properties that resemble gibberish.

2.4 Session Three, 1 December 2022, 13:00 – 15:00

Session Three had a mix of papers focusing on word-level statistical characteristics, transliteration alphabets, analysis of the script and the positional distribution of glyphs.

The first paper of the session was entitled *Crux of the MATTR: Voynichese Morphological Complexity* by Luke Lindemann. It uses two carefully validated word distribution statistics, the Moving Average Type-Token Ratio (MATTR) and the Most Common Words percentage (MCW) to determine the morphological complexity of the VM compared to an extensive set of languages from different families: 311 languages from 38 families. The results suggest that the VM is more complex than the average for Germanic and Romance languages and less complex than Semitic and Slavic (Lindemann, 2022).

The second paper in the session had the title *A new transliteration alphabet brings new evidence of word structure and multiple languages in the Voynich manuscript* by Massimiliano Zattera. This paper focuses on regularities that can be found in sequences of VM glyphs and proposes a so-called slot alphabet that is subject to a number of constraints. The majority of tokens can be decomposed in such slots. An algorithm is used to create a formal grammar for the word types in the VM which is subsequently used to classify sections of the text successfully (Zattera, 2022).

Examining the history of VM glyphs using phylogenetic methods by Katie Painter and Claire Bowern proposed a method based on phylogenetic networks on paleographic features (ten glyphs) to identify manuscript hand clusters. The method is first validated on known manuscript traditions. All the VM hands cluster closely together. The VM groups closest to the Uncial tradition, because of the absence of serifs and the relative lack of use of ligatures. However, in shape characteristics it is closest to Beneventan hands (Painter and Bowern, 2022).

The paper *Rightward and Downward Grapheme Distributions in the Voynich Manuscript* by Patrick Feaster proposed a systematic approach for detecting positional distributions of words and glyphs within lines and paragraphs in terms of "rightwardness" (distance towards the right end of a line) and "downwardness" (distance towards the bottom of a paragraph). The paper provides three examples of possible graphemic minimal word pairs, for example those containing [k] and [t], to show the nature of these patterns (Feaster, 2022).

2.5 Session Four, 1 December 2022, 15:00 – 17:00

The four papers in session 4 are divided between properties of the text, how the Voynich Manuscript fits in a typology of encrypted book-length works, and two papers examining ownership of the manuscript.

Seven Habits of Highly Eccentric Paragraphs by Tavi Stafford focuses on the ways in which the gallows glyphs[4] are distributed across words, lines, and paragraphs in ways that do not resemble letters in alphabetic text. Stafford's observations, along with Feaster's and Zattera's in other sessions, provide some insight into the structure of Voynich text and how word structure may help us understand the document's composition (Stafford, 2022).

Comparison of composition is also the focus of *The Voynich Manuscript Compared with Other Encrypted Books* by Klaus Schmeh and Elonka Dunin; they report on preliminary results of a corpus of enciphered books, examining their purpose of creation, authorship, and other features. Based on comparison with 118 encrypted books created between the 15th Century and the present, they argue that the Voynich manuscript is not a diary, and most resembles a book of knowledge. They also point out that the great majority of similar books have a single author. However, few conclusions can be drawn with certainty (Schmeh and Dunin, 2022).

From Voynich to the Beinecke, the Trail of Ownership by Farley Katz presents newly found archival documents and traces the path of ownership of the Voynich manuscript from the time of Voynich's death (in 1930) to its donation by H.P. Kraus to Yale's Beinecke Library in 1969. Working from Wilfrid and Ethel Voynich's wills, through documents related to Anne Nill's sale to Kraus' donation, he is able to correct existing online and printed summaries of these events (Katz, 2022).

Book transactions of Emperor Rudolf II 1576-1612. New findings on the earliest ownership of the Voynich manuscript by Stefan Guzy traces the opposite end of the Voynich chronology. Following up on a lead originally suggested[5] by René Zandbergen, Guzy presents his newly discovered evidence for how the Voynich manuscript arrived into the collection of Emperor Rudolf II. The paper provides evidence for the bill of sale of a small number of 'unusual books' from the Augsburg physician Karl Widemann. This sale took place in 1599 and the amount paid by Rudolf was 600 florins. Some suggestions are provided for where Widemann may have obtained the manuscript (Guzy, 2022).

3 Conclusions and outlook

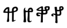
A survey taken after the conference clearly indicates that it was a successful event. There were 75 registered attendees for the conference who actively participated in the question periods following each talk.

The main takeaways from the presentations may be summarised as follows:

1. New and better information about the physical manuscript: its provenance in the 20th Century and details about its immediate history pre-Yale, as well as more information about the circumstances under which it may have ended up in Rudolph II's court. Tracing the sale (as well as the likely source) provides new clues for further establishing the location of the Voynich Manuscript between its creation and its arrival in Prague.

2. New work on the features of the text and script, although some conclusions are more subjective than others. A recurring theme throughout the conference was the identification of patterns above the level of individual words and what they might suggest about the circumstances of composition.

3. New—but still not entirely conclusive—work on the question of whether the contents are enciphered or meaningless. Both arguments have merit and make testable predictions, particularly in light of the focus that several

[4] 

[5] http://www.voynich.nu/history.html

papers had on the page structure's correlation with glyph distribution.

4. Some work on imagery was presented, although this is an area in which more research and comparative work would be welcome, especially in the light of the increasing availability of digital manuscripts.

There is an interest in repeating this event, possibly every two or three years.

Acknowledgments

The authors wish to acknowledge the generous support from the University of Malta in facilitating the online event, and hosting the Conference Web page.

We are indebted to the Programme Committee members for their support. These include, beside the authors of this paper: Dr. George Azzopardi (University of Groningen), Prof. Albert Gatt (University of Malta), Dr. Kristian Guillaumier (University of Malta), Dr. Chris Porter (University of Malta), Dr. Rafał Prinke (Emeritus, Eugeniusz Piasecki University), Prof. Jennifer Rampling (Princeton University) and Dr. Marc Tanti (University of Malta),

References

Claire Bowern and Daniel Gaskell. 2022. Enciphered after all? Word-level Text Metrics are Compatible with some Types of Encipherment. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Alexander Boxer. 2022. Fingerprinting Gibberish: A Quantitative Comparison of the Voynich and Sloane MS 3188. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Keagan Brewer. 2022. 'I beg your grace to suppress this chapter or else to have it written in secret letters': The Emotions of Encipherment in Late-Medieval Gynaecology. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Andrew Caruana, Colin Layfield, and John Abela. 2022. An Analysis of the Relationship between Words within the Voynich Manuscript. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Lisa Fagin Davis. 2022. Voynich Paleography. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Kevin Farrugia, Colin Layfield, and Lonneke van der Plas. 2022. Demystifying the Scribes behind the Voynich Manuscript using Computational Linguistic Techniques. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Patrick Feaster. 2022. Rightward and Downward Grapheme Distributions in the Voynich Manuscript. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Daniel Gaskell and Claire Bowern. 2022. Gibberish after all? Voynichese is Statistically Similar to Human-Produced Samples of Meaningless Text. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Koen Gheuens and Cary Rapaport. 2022. Above and Beyond Voynich Canopies: Tents as a Recurring Motif in Beinecke MS 408. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Stefan Guzy. 2022. Book Transactions of Emperor Rudolf II 1576-1612. New Findings on the Earliest Ownership of the Voynich Manuscript. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Jürgen Hermes. 2022. Polygraphia III: The Cipher that Pretends to be an Artificial Language. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Farley Katz. 2022. From Voynich to the Beinecke, the Trail of Ownership. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Colin Layfield and John Abela, editors. 2022. *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*, volume 3313. CEUR Workshop Proceedings http://ceur-ws.org/Vol-3313/.

Luke Lindemann. 2022. Crux of the MATTR: Voynichese Morphological Complexity. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Katie Painter and Claire Bowern. 2022. Examining the history of Voynich Glyphs using Phylogenetic Methods. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Klaus Schmeh and Elonka Dunin. 2022. The Voynich Manuscript Compared with Other Encrypted Books. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Klaus Schmeh. 2013. A milestone in voynich manuscript research: Voynich 100 conference in monte porzio catone, italy. *Cryptologica*, 37(3):193–203.

Tavi Stafford. 2022. Seven Habits of Highly Eccentric Paragraphs. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Johannes Trithemius. 1518. *Polygraphia libri sex*. Haselberg, Basel.

René Zandbergen. 2022. Transliteration of the Voynich MS Text. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Massimiliano Zattera. 2022. A new Transliteration Alphabet brings new Evidence of Word Structure and Multiple ”languages” in the Voynich Manuscript. In *Proceedings of the 1st International Conference on the Voynich Manuscript 2022 (VOY2022) (Online [Zoom], Qrendi, Malta, Nov 30-Dec 1 2022)*.

Historical Language Models in Cryptanalysis: Case Studies on English and German

Beáta Megyesi and Justyna Sikora
Uppsala University
Sweden

Filip Fornmark and Michelle Waldispühl
University of Gothenburg
Sweden

Nils Kopal and Vasily Mikhalev
University of Siegen
Germany

Abstract

In this paper, we study the impact of language models (LM) on decipherment of historical homophonic substitution ciphers. In particular, we investigate if decipherment by using hill-climbing and simulated annealing can benefit from LMs generated from historical texts in general and century-specific texts in particular. We carry out experiments on homophonic substitution ciphers with English and German as plaintext languages. We take into account ciphertext length as well as n-gram size of the LMs. We compare the results on decipherment based on historical LMs with large LMs generated from modern texts. The results show that using historical LMs in decipherment of homophonic substitution ciphers leads to significantly better performance on ciphertext produced in the 17th century or earlier, and century-specific language models yield better results on longer and older ciphertexts.

1 Introduction

One of the main components in cryptanalysis of historical ciphertexts is language models of the underlying plaintext. Oftentimes the choice of texts on which the language model is based is opportunistic, i.e. the cryptanalysts choose what texts or models are available. Since collections of (more or less) contemporary texts, such as the collection of the Gutenberg project or the Wikipedia articles are accessible and freely available for many languages, these are often used by cryptanalysts for the generation of LMs, e.g. by Lasry (2018) and Bean (2020). However, using language models derived from contemporary languages might not be optimal for the decipherment of historical ciphertexts since language changes over time.

Before spelling and grammar were normalized for many European written languages in the course of the 18th and 19th century, we find a large variation in spelling in historical texts. The variation can be found not only across regions and writers but the same author could also spell the same word in the same document differently. The use of punctuation marks such as dots and commas was rarer than in modern texts. Further, writers used abbreviations to save space of the expensive paper or parchment. Another important aspect when dealing with historical texts is the fact that language changes over time; new words enter the language while others disappear. Not only words, but also the grammatical structure of the language changes with respect to word order and the internal structure of words (so called morphology). Given the above mentioned reasons we can expect that the usage of LMs generated from contemporary or historical texts have an impact on decipherment accuracy.

In this paper we aim to investigate the role of historical LMs in the decipherment process of historical ciphertexts. In particular, we are interested in finding the answer to the following research questions:

- Do historical LMs have a positive impact on the decipherment of historical ciphers?
- Do historical LMs created from the same century as the cipher originates from lead to an increase in decipherment performance?
- Does increasing the n-gram size of a model result in an increase of accuracy in decipherment of historical manuscripts?

We present a pilot study on English and German ciphertexts and LMs generated from texts originating from the 14th to the 20th centuries of various n-gram sizes. We focus on homophonic substitution ciphers of various lengths. As historical homophonic substitution ciphertexts have been successfully deciphered by using hill-climbing and simulated annealing (see e.g. Lasry et al. (2020)), we chose to apply the same decipherment method in our experiments.

In Section 2, we present previous studies on using historical LMs in decipherment. In Section 3, we give an overview of the data sets used for the experiments on English and German. In Section 4, we describe the method with the experimental setup. In Section 5, the results from the various experiments are presented, and discussed in Section 6. Finally, in Section 7, we conclude our paper.

2 Background

The usage of language models generated from the underlying plaintext language of the cipher is inevitable for successful decipherment. Already in the 9th century, the Arab philosopher and mathematician Al-Kindi described the value of frequency analysis derived from the plaintext and the ciphertext in cryptanalysis. Since then, cryptographers used LMs with information about frequencies of letters and letter co-occurrences to create models of the underlying plaintext. One of the most simplest, efficient and also commonly occurring language models are n-grams that use a statistical approach to predict the probability of a word or character given its context. N-gram models are built by analyzing a sequence of n words or characters in a text corpus and building a probability distribution of the next word or character based on the occurrence of each n-gram in the training data. The distribution of various n-gram orders (e.g. unigrams, bigrams, trigrams) is also reflected on the ciphertext and can thus give more clues into how the text was encrypted (Kahn, 1996) and (Dooley, 2018).

In order to train or generate LMs for decipherment purposes, a wide range of text collections are used. These might include the translation of the Human Rights, texts from Wikipedia, or Google books. When dealing with historical ciphers, the most commonly used corpus is the Gutenberg collection from Project Gutenberg[1]. The Gutenberg project is a digital library of free e-books, collected since 1971. The collection consists of more than 60.000 books in a large number of languages, for which U.S. copyright has been expired; the great majority originating from the 19th century. The collection is publicly available and copyright-free, which explains its popularity to use when building historical language models.

Another collection of historical texts is the HistCorp corpus (Pettersson and Megyesi, 2018) which contains sixteen European languages including Czech, Dutch, English, French, German, Greek, Hungarian, Icelandic, Italian, Latin, Polish, Portuguese, Russian, Slovene, Spanish, and Swedish. The transcriptions of the original manuscripts are diplomatic editions, i.e. the orthography of the original text is kept and mistakes in the original are preserved. The texts are released in a uniform format. Noteworthy is that the number of texts and the data size included for the various languages vary greatly in the collection depending on what kind of historical text corpora are available for the particular language.

To generate language models, character as well as word-based models are common: unigram, bigram, trigram up to sixgram models are used for the purpose of decipherment. Naturally, the higher the order of the n-grams, the more texts are needed to generate suitable models, and the bigger the models become.

Surprisingly, even though historical cryptologists agree on the importance of LMs in cryptanalysis, the role of historical LMs has not been studied before, neither extensively, nor systematically.

A few studies report, however, on the evaluation of various n-gram sizes in decipherment using modern texts. Ravi and Knight (2008) present a method for solving substitution ciphers using low-order character-based n-gram models and show how decipherment accuracy varies as a function of cipher length and n-gram order.

The same authors (Ravi and Knight, 2011) present a Bayesian approach for deciphering complex substitution ciphers and evaluate with different setups of LMs including character-based bigrams and trigrams, as well as word-based trigrams. They conclude that the best decipherment results are achieved with trigram models and a word list.

Nuhn et al. (2014) apply a method for solving substitution ciphers, including the Zodiac-408 cipher, and evaluate n-gram models of orders four, five and six, where the sixgram models performed the best with lowest error rate.

[1] www.gutenberg.org

Hauer et al. (2014) presents an approach for deciphering monoalphabetic substitution ciphers that combines both character-level and word-level based LMs. In the above mentioned studies the language models were generated from contemporary texts, and evaluated in the light of specific approaches.

Indications for using historical sources instead of contemporary texts for the decipherment of historical ciphertexts have been given in the study by Pettersson and Megyesi (2019) for the automatic language identification task in three types of historical ciphertexts. The results showed that historical LMs perform considerably better on the tested languages (German and Italian) and that using models based on historical texts enables to capture old word forms that are not present in modern corpora, despite their larger size.

Historical LMs were also used to identify cleartext and its language in historical ciphertexts (Gambardella et al., 2022). The authors conducted a series of experiments on 214 documents in 8 languages including Dutch, French, Hungarian, Italian, Latin, Portuguese, and Spanish, and tested the ability of the models in various n-gram settings; both character and word based, trained on historical corpora from the HistCorp collection.

More extensive studies on the impact of historical LMs in decipherment are based on two thesis works: one bachelor's thesis in linguistics carried out by Fornmark (2022) on English, and one master's thesis in language technology by Sikora (2022) on German. The thesis works are based on the same method designed by the authors of this paper. In this study, we built upon the two theses and compare the similarities and differences of the decipherments using modern vs historical LMs for the two languages.

3 Text Collections

To carry out experiments on the impact of historical texts in decipherment for English and German, we use the HistCorp collection for the generation of LMs from historical texts, and contemporary texts from the Gutenberg project.

In English, the earliest texts are mostly Biblical sources, for example from The EDGeS Diachronic Bible Corpus (EDGes) (Bouma et al., 2020), whereas the later texts are of different genres and styles, including sources from the Corpus of Late Modern English Texts (DeSmet, 2006) and the Lampeter Corpus of Early Modern English Tracts (Schmied et al., 1999).

The German data consists of texts from the HistCorp collection including material from the Deutsches TextArchiv (DTA) (Textarchiv, 2010), the EDGeS Diachronic Bible Corpus (Bouma et al., 2020), the Nottingham Corpus of Early Modern German Midwifery and Women's Medicine (GeMi) (Whitt, 2016), GerManC (Durrell et al., 2012), Reference Corpus of Middle High German (ReM) (Klein et al., 2016), Reference Corpus of Middle Low German/Low Rhenish (ReN) (Schröder, 2018), and Register in Diachronic German Science (Ridges) (Lüdeling et al., 2016).

The texts from all corpora for both languages were then sorted into centuries to create subdomains per time interval of 100 years to serve as basis for the creation of century-specific LMs.

For English, texts from between the 11th and 13th centuries as well as the 15th century are missing from the source material, and could thus not be included. For German only one time period, namely the 20th century, could not be represented due to data sparseness. More detailed information about the data source and creation is given by Fornmark (2022) for English and by Sikora (2022) for German.

4 Method

To investigate whether LMs generated on historical corpora can lead to a better performance on decipherment of historical ciphers, compared to LMs generated on large modern corpora, we choose to experiment with English and German as the underlying plaintext languages. We use plaintexts from the 11th to the 20th centuries. We first describe the features that might have effect on the decipherment results. Then, we present the experimental design with a walk-through of the various stages.

4.1 Features

For historical ciphers, we select one of the most commonly occurring types, namely homophonic substitution ciphers with a mixture of the number of homophones per plaintext alphabet letter. The reason behind varying the number of homophones for each plaintext letter is to create more authentic ciphers which are similar to original ciphers, as retrieved from European archives and libraries. Previous studies showed that in homophonic substitution ciphers we can find variable number of code elements for different plaintext letters depending on their frequency in the particular language; frequently occurring plaintext letters usually receive two or three code elements, while less frequent letters are assigned one code element (Kahn, 1996) and (Megyesi et al., 2022).

To measure the correlation between decipherment accuracy and the order of LMs on the length of ciphertexts, we apply ciphertexts consisting of 200 and 500 characters. The decision was determined by the assumption that generating shorter ciphers would significantly increase the level of decipherment difficulty. For the investigation of the impact of various n-gram sizes on decipherment, we experiment with trigram, fourgram, and fivegram character-based LMs. The experimental setup of the features and their values that our study is built upon is summarized in Table 1.

Feature	Values
Language:	German, English
Time period (cent.):	11th-20th
N-gram size:	3, 4, 5
Ciphertext length (chars):	200, 500

Table 1: Experimental setup with features and their values.

4.2 Experiment design

Our point of departure is a ciphertext and its corresponding plaintext for evaluation without any access to the cipher key. First, we collect and preprocess the plaintexts in English and German to generate the plaintext alphabet for various centuries for both languages. Since we do not have access to the original ciphertexts (with their corresponding plaintext) we need to generate them to be able to evaluate the decipherment results. We then create the LMs of various order sizes given the plaintexts for the various centuries and languages. Finally, we run cryptanalysis using the various LMs and evaluate the output. The entire process is illustrated in Figure 1.

4.2.1 Alphabet generation

To create century specific LMs, we generate plaintext alphabets for the specific time periods by counting unigram frequencies on the basis of texts in the HistCorp collection for both languages. It turns out that the size of the alphabet varies greatly across centuries, which causes memory problems in the generation of LMs. Therefore, we set several threshold values for letter frequencies: 0.001, 0.01 and 0.05, and linguistically analyze the alphabet output. We decide to use characters more frequent than 0.01% for English and 0.05% for German.

We then normalize the output of the alphabets; all letters are upper-cased, with the exception of the German *SS* letter[2], while punctuation marks, digits and white space are removed. Furthermore, we expand abbreviations and transform characters consisting of a regular letter and a superscript letter into two separate characters.

4.2.2 Key, ciphertext and plaintext generation

To be able to test the impact of historical and contemporary LMs on decipherment, the keys, ciphertexts and their corresponding plaintexts are automatically generated. To generate timely typical historical cipher keys, we studied original cipher keys collected from the DECODE database (Megyesi et al., 2019) with certain characteristics. We decided to derive homophonic substitution keys containing English or German plaintext or cleartext languages with available transcription. Since language identification in cipher keys might be challenging, we studied more carefully the words in the nomenclature list to decide whether the cipher key was created for English or German plaintexts.

To generate plaintext we choose the HistCorp collection. For each language, we split the language specific data set in HistCorp into a training set and a test set in the portion of 80%–20%, respectively. Table 2 presents the number of texts in the training and test sets for English and German. The training set serves for the generation of LMs in the subsequent step while the test set is used for the gold-standard decrypted text to be used for evaluation. To be able to make automatic comparisons between the automatically generated decrypted text and their corresponding gold-standard, the test set is preprocessed; normalized to capitalize all characters, and double spaces, punctuation, and non-letter characters are removed.

The text files for each century are randomly

[2] We keep the letter *SS* in lowercase, since Python returns the upper-cased strings, and transforms *SS* into double *S* - “SS”, which causes problems in frequency analysis.

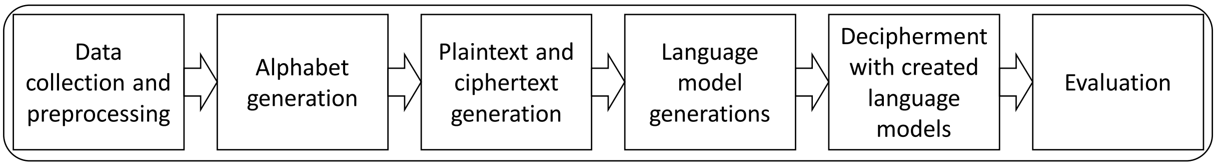

Data collection and preprocessing
Alphabet generation
Plaintext and ciphertext generation
Language model generations
Decipherment with created language models
Evaluation

Figure 1: Method.

Cent.	Training		Test	
	English	German	English	German
11th	–	22	–	6
12th	–	155	–	38
13th	–	98	–	24
14th	60	70	15	18
15th	–	74	–	19
16th	64	39	16	10
17th	122	275	30	69
18th	184	582	46	145
19th	179	529	45	132
20th	46	–	12	–

Table 2: Number of documents in training and test sets in English and German per century.

shuffled to create a random selection of texts. Then, 200 vs 500 characters of each plaintext file are extracted to create plaintexts of various lengths.

Given the keys and the plaintexts, we then generate ciphertexts. For each time period, ciphertexts are created, both of length 200 and 500.

Furthermore, for each ciphertext of a given length and for a given time period, homophonic ciphers with a mixed number of homophones are built. The keys use the percentage value of the measured character-based unigram frequencies described above. A number of homophones between one and five is assigned to each plaintext letter. Between three and five homophones for the most common 14 letters, one homophone for the least common, and between two and three homophones for letters in between. Null characters and nomenclature elements are not considered.

In the experiments, the keys and ciphertexts are all numeric, and use a fixed, uniform length for each ciphertext letter. If the key needs less than 100 homophones, all plaintext letters map to two numbers (00, 01, . . . , 99) in the ciphertext. Otherwise, the plaintext letters map to three numbers (000, 001, . . . , 999) in the ciphertext. An example of a generated cipher key is illustrated in Figure 2.

```
A:[99|18|12|68]     H:[77|08|67]     O:[86|35|41|75]     V:[65]
B:[60]              I:[00|38|34|97]  P:[52]              W:[54]
C:[11]              J:[47]           Q:[30]              X:[20]
D:[29|61]           K:[33]           R:[98|56|24]        Y:[62]
E:[92|13|40|17|19]  L:[88|42]        S:[14|22|79]        Z:[23]
F:[93]              M:[28]           T:[27|84|51|80|96]
G:[72]              N [89|21|74]     U:[45]
```

Figure 2: A generated key; homophones per plaintext letter based on unigram frequencies.

4.2.3 Language model generation

Before generating the models themselves, duplicated texts – the same texts appearing in several centuries – are removed. Then, character-based n-gram models of order 3, 4, and 5 are created from the training set for each century, and a more generic model with all texts available for each language from the Gutenberg collection.

The model format consists of a data and a metadata section. The data section is an array with n-grams and their respective frequencies. The number of occurrences of any particular n-gram is stored as the logarithm of the frequency of that n-gram relative to the full body of text data. The metadata section, which is located in the beginning of the LM file starts with a file identifier “CTLS” (CrypTool Language Statistics), followed by the language code (“EN” for English and ”DE” for German), an integer describing the “n” value of the gram, and the model alphabet.

From the training sets, various character-based models for the three n-gram sizes are created for the different centuries, as well as combined models are generated from all texts.

Lastly, word-based LMs are generated from the texts. The data is cleaned up by removing residual punctuation for both languages. For English, the diacritics were also removed while for German they were kept. The resulting format is a single word per line, with words occurring in the source material. A combined dictionary from the included English and German HistCorp material is also created along with a general German and English dictionary generated from the Gutenberg data.

4.2.4 Decipherment

The cryptanalysis is performed using CrypTool 2[3] (CT2), a freely available open source tool[4] which allows the automatic decipherment of historical and modern ciphers (Kopal, 2018). CT2 contains a component for the cryptanalysis of homophonic substitution ciphers, the so-called Homophonic Substitution Analyzer, see Figure 3. To ease the use of the cryptanalytic algorithm implemented in the component, we extracted the core cryptanalysis algorithm. Thus, it could be used without the need of starting a full-blown CT2 instance. This furthermore speed up the cryptanalysis and allowed us to perform several hundreds of cryptanalysis runs per model needed for our evaluations.

To decipher a given ciphertext, CT2's Homophonic Substitution Analyzer component implemented with hill climbing with simulated annealing (Kopal, 2019) was used. Additionally, a dictionary of common words was given to the algorithm. During the cryptanalysis process, the dictionary is used to already "lock" partially correct decipherments to improve and speed up the further analysis process.

4.2.5 Evaluation

To evaluate the effect the LMs have on decipherment, we calculate decipherment accuracy as defined in the equation below.

$$\mathbf{Correct} = \sum_{\mathbf{i=0}}^{\mathbf{Length-1}} \mathbf{n} \begin{cases} 1, & \text{if D[i] = P[i]} \\ 0, & \text{otherwise} \end{cases}$$

where Length is the length of the ciphertext, D is the deciphered ciphertext, and P is the real plaintext.

We compute the percentage of the correctly deciphered letters of a deciphered ciphertext as defined below.

$$\mathbf{Accuracy} = (Correct/Length) \cdot 100$$

Finally, a LM receives a "point" if it was able to decipher a given ciphertext with $Accuracy \geq 80\%$. Here, only the best of all models received a point. But if more than one model lead to the same Accuracy, all these models obtained a point since they performed equally good. To evaluate the different LMs, we compare the number of points each LM received in our evaluation. The graphs in the Results section show how many points each LM received. We evaluated each LM 500 times for each time period by cryptanalyzing each generated ciphertext using each of the LMs.

5 Results

We provide the results for English in Figures 4 and 5 and for German in Figure 6.

For the English texts we found that texts composed in a certain time period were, in general, best analyzed by the model trained on texts from the same century. This trend was especially clear for the earliest texts, i.e. for texts from the 14th, 16th and 17th centuries. For later texts it became less clear which model performed the best, but there was a marked drop in the performance of the early models. These results can be linked to the development of English orthography, which with the spread of dictionaries in the 18th century became more standardized.

Given the choice of the n-gram order, the results showed that 5-gram models achieved the highest decipherment accuracy on more modern texts, and 4-gram models led to highest performance for older ones, produced in the 17th century or earlier.

Not surprisingly, the longer ciphertexts (500 characters) achieved higher decipherment accuracy in general.

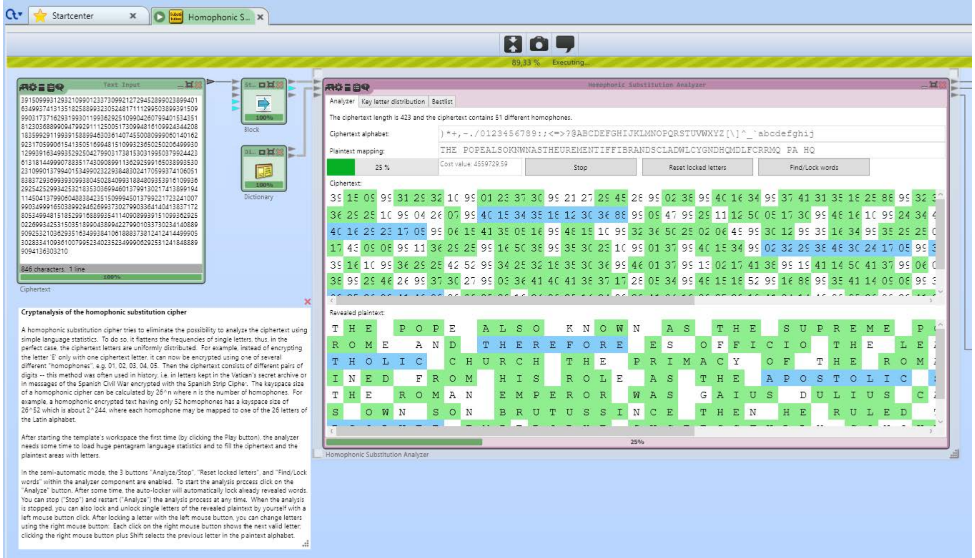
The order size of LMs has also impact on the decipherment performance of the cipherlength. We found that shorter texts require higher order n-grams; the 5-gram models performed better compared to the 3-gram and 4-gram models on the 200 character long text.

Interestingly, using the Gutenberg model compared to the model generated from the combination of all text material (1350-1999), the results are diverse. While Gutenberg 4-and 5-gram models perform best on longer and more modern texts (18th-20th century), the historical 4- and 5-gram merged model yields best result for shorter texts. However, noteworthy is that for all 3-gram models, the historical merged model leads to best decipherment performance.

For German, the results show similar trends, albeit not as pure and straightforward as for English.

[3] https://www.cryptool.org/en/ct2/

[4] https://github.com/CrypToolProject/CrypTool-2



```
3915099931293210990123373099212729452899023899401
6349937413135182588993230524817111299503899391509
9903173716293199301199362925109904260799401534351
8123036889909479929111250051730994816109924344208
1835992911993915889946303614074550080999060140162
9231705990615413505169948151099323650250206499930
1299391634993529250427990317381530319950379924423
6131814499907883517430908991136292599165038993530
2310990137994015349902322938483024170599374106051
8383729369939309938045028409931884809353916109936
2925425299342532183530369946013799130217413899194
1145041379906048838423515099945013799221723241007
9903499916503899294626993730279903364140413837172
8053499481518529916889935411409089939151099362925
0226993425315035189904389942279901033730234140889
9092532103629351634993841061888373812412414499905
3028334109361007995234023523499906292531241848889
9094136303210
846 characters, 1 line
```

Cryptanalysis of the homophonic substitution cipher

A homophonic substitution cipher tries to eliminate the possibility to analyze the ciphertext using simple language statistics. To do so, it flattens the frequencies of single letters, thus, in the perfect case, the ciphertext letters are uniformly distributed. For example, instead of encrypting the letter 'E' only with one ciphertext letter, it can now be encrypted using one of several different "homophones", e.g. 01, 02, 03, 04, 05. Then the ciphertext consists of different pairs of digits -- this method was often used in history, i.e. in letters kept in the Vatican's secret archive or in messages of the Spanish Civil War encrypted with the Spanish Strip Cipher. The keyspace size of a homophonic cipher can be calculated by 26^n where n is the number of homophones. For example, a homophonic encrypted text having only 52 homophones has a keyspace size of 26^52 which is about 2^244, where each homophone may be mapped to one of the 26 letters of the Latin alphabet.

After starting the template's workspace the first time (by clicking the Play button), the analyzer needs some time to load huge pentagram language statistics and to fill the ciphertext and the plaintext areas with letters.

In the semi-automatic mode, the 3 buttons "Analyze/Stop", "Reset locked letters", and "Find/Lock words" within the analyzer component are enabled. To start the analysis process click on the "Analyze" button. After some time, the auto-locker will automatically lock already revealed words. You can stop ("Stop") and restart ("Analyze") the analysis process at any time. When the analysis is stopped, you can also lock and unlock single letters of the revealed plaintext by yourself with a left mouse button click. After locking a letter with the left mouse button, you can change letters using the right mouse button: Each click on the right mouse button shows the next valid letter; clicking the right mouse button plus Shift selects the previous letter in the paintext alphabet.

The ciphertext length is 423 and the ciphertext contains 51 different homophones.

Ciphertext alphabet:)*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefghij

Plaintext mapping: THE POPEALSOKNWNASTHEUREMENTIFFIBRANDSCLADWLCYGNDHQMDLFCRRMQ PA HQ

Cost value: 4559729.59

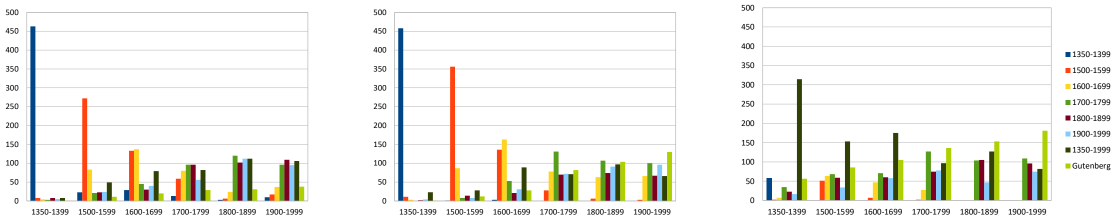
Ciphertext:

```
39 15 09 99 31 29 32 10 99 01 23 37 30 99 21 27 29 45 28 99 02 38 99 40 16 34 99 37 41 31 35 18 25 88 99 32
36 29 25 10 99 04 26 07 99 40 15 34 35 18 12 30 36 88 99 09 47 99 29 11 12 50 05 17 30 99 48 16 10 99 24 34
40 16 29 23 17 05 99 06 15 41 35 05 16 99 48 15 10 99 32 36 50 25 02 06 49 99 30 12 99 39 16 34 99 35 29 25
17 43 09 08 99 11 36 29 25 99 16 50 38 99 35 30 23 10 99 01 37 99 40 15 34 99 02 32 29 38 48 30 24 17 05 99
39 16 10 99 36 29 25 42 52 99 34 25 32 18 35 30 36 99 46 01 37 99 13 02 17 41 38 99 19 41 14 50 41 37 99 06
38 99 29 46 26 99 37 30 27 99 03 36 41 40 41 38 37 17 28 05 34 99 48 15 18 52 99 16 88 99 35 41 14 09 08 99
```

Revealed plaintext:

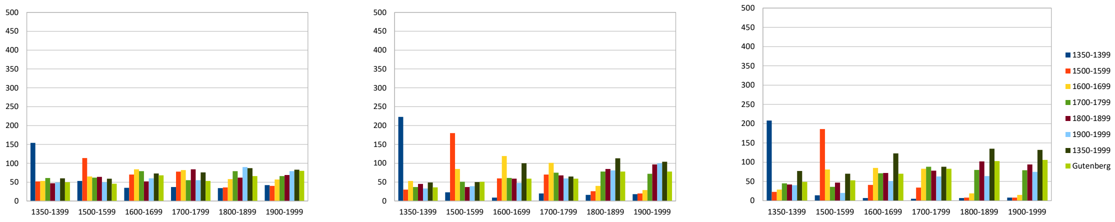
```
T H E   P O P E   A L S O   K N O W N   A S   T H E   S U P R E M E   P
R O M E   A N D   T H E R E F O R E   E S   O F F I C I O   T H E   L E
T H O L I C   C H U R C H   T H E   P R I M A C Y   O F   T H E   R O M
I N E D   F R O M   H I S   R O L E   A S   T H E   A P O S T O L I C
T H E   R O M A N   E M P E R O R   W A S   G A I U S   D U L I U S   C
S   O W N   S O N   B R U T U S S I N C E   T H E N   H E   R U L E D
```

Figure 3: Breaking homophonic substitution ciphers in CrypTool2 .



(a) 3-grams. (b) 4-grams. (c) 5-grams.

Figure 4: Best result for English: 500 character long ciphertexts.



(a) 3-grams. (b) 4-grams. (c) 5-grams.

Figure 5: Best result for English: 200 character long ciphertexts.

Similar to English, historical century-specific LMs achieve best decipherment performance, but for German, that applies to all historical texts produced earlier than the 19th century. The 16th century model yielded most impressive results for ciphertexts produced up to the 18th century.

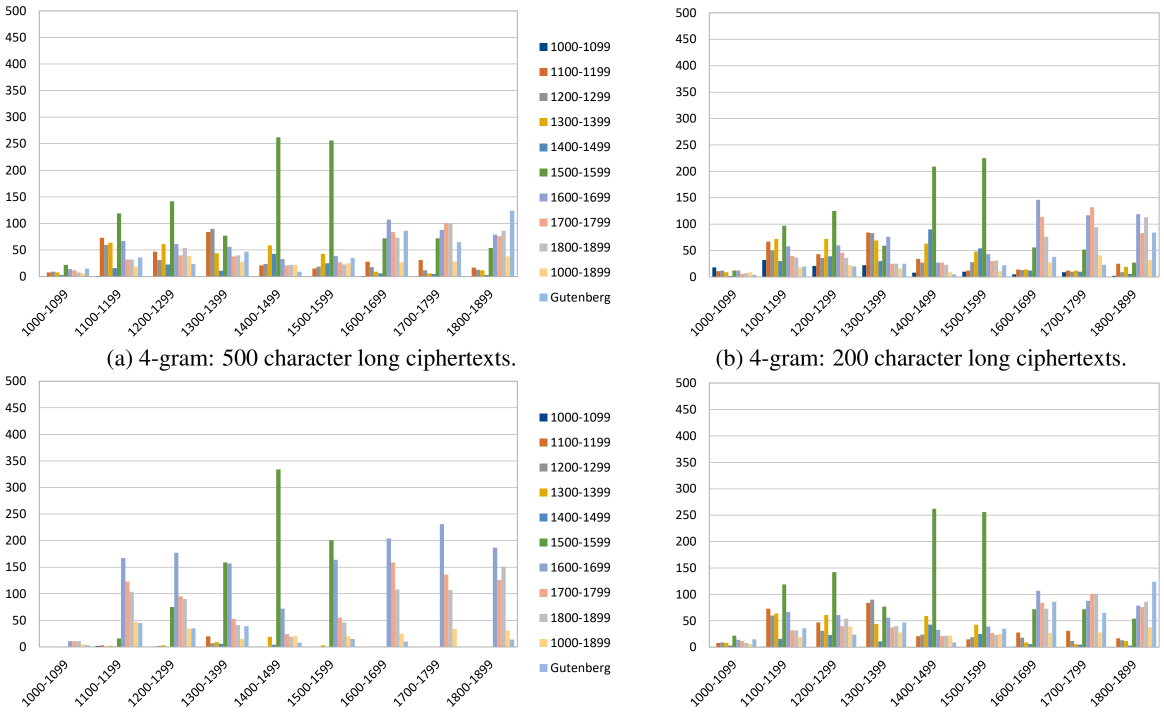
Interestingly, the Gutenberg model outperformed a merged model generated from all historical texts from the HistCorp collection (1000-1899) with the exception of a few cases, see 4-gram and 5-gram models on 500 character long texts from 1200-1299 and 1400-1499, and 200 character long ciphertexts from 1700-1799.

Like English, the longer the ciphertext, the higher decipherment accuracy is. 5-gram models fit best for shorter as well as long ciphertexts if these are century-specific. However, if no century-specific data is available for LM generation, 4-gram models seem to be optimal for longer texts.

6 Discussion

The overall best results for English and German are achieved by applying 4- and 5-gram models, historical 4-grams for texts produced in the 17th century or earlier for English, and 19th century or earlier for German, while 5-gram models are preferable for longer and more modern texts.

The results are not surprising. Shorter cipher-



(a) 4-gram: 500 character long ciphertexts.

(b) 4-gram: 200 character long ciphertexts.

(c) 5-gram: 500 character long ciphertexts.

(d) 5-gram: 200 character long ciphertexts.

Figure 6: Best result for German with 4- and 5-gram models.

texts require more reliable LMs for successful decipherment, which in turn require more specific and larger amounts of input data to achieve cryptanalysis with higher performance. Thus, the results are highly dependent on the amount of available training data, since the larger models tend to perform better in these cases. The conclusion should therefore be considered carefully, and be regarded as potential trends, and further validation with the use of different models, source data, and languages is needed.

Noteworthy is also that accuracies reported in this study might have become higher by applying more restarts and other parameters in the decipherment process. The goal of this study, however, is not to reach the highest decipherment accuracy, but to evaluate the general performance of the models given language data from various time periods.

Considering previous studies discussed in Section 2, such as Nuhn and Knight (2014) and Bean (2020) in which better results are reported on using higher order n-gram models, the data sizes required are significantly larger than for most of the models used in our study. The ciphers themselves are, furthermore, different from the ones analysed here, why a full comparison of the results is not possible.

Our next step is to carry out evaluation of the language models on various cipher types with underlying plaintext of many European languages. While we only investigated two languages, both belonging to the Germanic language family with rather similar structure, we would like to include other, more dissimilar languages of different types, such as the Indo-European Romance and Slavic, as well as Finno-Ugric languages such as Finnish and Hungarian.

Another plan is to investigate the impact of the size and coverage of LMs and their impact on decipherment. The texts which the models were generated from cannot be said to be balanced, and not of the same size with respect to centuries. In addition, for comparison between historical LMs and those generated from other sources, several language models could be used: from unigram up to 6-gram character- as well as word-based models. Apart from the Gutenberg collection, we aim to use the recently released google n-gram models for a wide range of languages from 1 to 5 character-based n-grams generated from printed books of different genres and time periods (Google, 2022)[5].

[5] https://storage.googleapis.com/books/ngrams/books/datasetsv2.html

7 Conclusion

In this paper, we investigated the influence and impact of language models on decipherment of historical ciphertexts. We conducted experiments on English and German to find out if language models generated from historical texts or modern text fit best for decipherment. We ran experiments on texts from the 11th to the 19th centuries. We investigated character-based n-gram models of size 3-, 4-, and 5-grams. We focused on homophonic substitution ciphers of various lengths. For ciphertext length, we experimented with 200 and 500 characters long ciphertext messages. We conducted experiments on homophonic substitution ciphers with a mixture of 1, 2, 3, 4, and 5 code elements to imitate the nature of original ciphers from early modern times. For comparison we generated LMs from contemporary texts derived from the Gutenberg project, and a merged model of all historical century-specific texts.

The experiments clearly indicate that the age and the length of the ciphertext have great influence on the results, and that ciphertext characteristics shall be taken into account when choosing suitable language models for cryptanalysis. Likewise, the amount of available plaintext data serving for the generation of the language models should also be considered when choosing a suitable n-gram order.

The results show that decipherment by hill-climbing and simulated annealing using historical n-gram models perform better on ciphertext produced in the 17th century or earlier for English and in the 19th century or earlier for German, and century-specific language models perform better on longer, older, and less complex ciphertexts. The larger LMs generated from the Gutenberg collection are preferable on ciphers from 18th-19th centuries. Further, experiments on n-gram size show that LMs based on 4-grams and 5-grams achieve highest performance on both English and German; 5-gram models for longer text and more modern texts and 4-gram models for historical ones.

Acknowledgments

This work has been supported by the Swedish Research Council, grant 2018-06074, DECRYPT – Decryption of Historical Manuscripts.

References

Richard Bean. The Use of Project Gutenberg and Hexagram Statistics to Help Solve Famous Unsolved Ciphers. In *Proceedings of the 3rd International Conference on Historical Cryptology HistoCrypt 2020*, number 171, pages 31–35. Linköping University Electronic Press, 2020.

Gerlof Bouma, Evie Coussé, Trude Dijkstra, and Nicoline van der Sijs. The EDGeS Diachronic Bible Corpus. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, Marseille, France, May 2020. European Language Resources Association.

Hendrik DeSmet. The Corpus of Late Modern English Texts (extended version), 2006.

John F Dooley. *History of Cryptography and Cryptanalysis*. Springer, 2018.

Martin Durrell, Paul Bennett, Silke Scheible, and Richard J. Whitt. GerManC, 2012. URL `http://hdl.handle.net/20.500.12024/2544`. Oxford Text Archive.

Filip Fornmark. Models, Keys and Cryptanalysis - Evaluating Historical Statistical Language Models in Cryptanalysis of Homophonic Substitution Ciphers, 2022. Bachelor thesis in Linguistics, Gothenburg University, Sweden.

Maria-Elena Gambardella, Beáta Megyesi, and Eva Pettersson. Identifying Cleartext in Historical Ciphers. In *Proceedings of the Workshop on Language Technologies for Historical and Ancient Languages. LT4HALA 2022*, 2022.

Google. The Google Books Ngram Viewer Dataset, link: https://pypi.org/project/google-ngram-downloader/, 2022.

Bradley Hauer, Ryan Hayward, and Grzegorz Kondrak. Solving Substitution Ciphers with Combined Language Models. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, Dublin, Ireland, August 2014.

David Kahn. *The Codebreakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet*. Scribner, New York, NY, 1996.

Thomas Klein, Klaus-Peter Wegera, Stefanie Dipper, and Claudia Wich-Reif. Reference Corpus Middle High German (1050–1350) Referenzkorpus Mittelhochdeutsch (1050–1350), version

1.0. https://www.linguistics.rub.de/rem/, 2016. Accessed: 2022-07-30.

Nils Kopal. Solving Classical Ciphers with CrypTool 2. In *Proceedings of the 1st International Conference on Historical Cryptology HistoCrypt 2018*, number 149, pages 29–38. Linköping University Electronic Press, 2018.

Nils Kopal. Cryptanalysis of Homophonic Substitution Ciphers using Simulated Annealing with Fixed Temperature. In *Proceedings of the 2nd International Conference on Historical Cryptology, HistoCrypt*, pages 107–16. Linköping University Electronic Press, 2019.

George Lasry. *A Methodology for the Cryptanalysis of Classical Ciphers with Search Metaheuristics.* kassel university press GmbH, 2018.

George Lasry, Beáta Megyesi, and Nils Kopal. Deciphering Papal Ciphers from the 16th to the 18th Century. *Cryptologia*, pages 479–540, 2020. URL https://www.tandfonline.com/doi/full/10.1080/01611194.2020.1755915.

Anke Lüdeling, Carolin Odebrecht, Thomas Krause, Gohar Schnelle, and Catharina Fischer. Ridges-herbology (version 9.0). https://www.deutschestextarchiv.de/, 2016. Accessed: 2022-07-30.

Beáta Megyesi, Nils Blomqvist, and Eva Pettersson. The DECODE Database: Collection of Ciphers and Keys. In *Proceedings of the 2nd International Conference on Historical Cryptology, HistoCrypt19*, Mons, Belgium, June 2019.

Beáta Megyesi, Crina Tudor, Benedek Láng, Anna Lehofer, Nils Kopal, Karl de Leeuw, and Michelle Waldispühl. Keys with Nomenclatures in the Early Modern Europe. *Cryptologia*, 0 (0):1–43, 2022. doi: 10.1080/01611194.2022.2113185.

Malte Nuhn and Kevin Knight. Cipher Type Detection. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1769–1773. Association for Computational Linguistics, 01 2014.

Malte Nuhn, Julian Schamper, and Hermann Ney. Improved Decipherment of Homophonic Ciphers. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, October 2014. Association for Computational Linguistics.

Eva Pettersson and Beáta Megyesi. The HistCorp Collection of Historical Corpora and Resources. In *Proceedings of the Digital Humanities in the Nordic Countries 3rd Conference*, Helsinki, Finland, March 2018.

Eva Pettersson and Beata Megyesi. Matching Keys and Encrypted Manuscript. In *Proceedings of the 22nd Nordic Conference on Computational Linguistics*, pages 253–261, Turku, Finland, 30 September – 2 October 2019. Linköping University Electronic Press.

Sujith Ravi and Kevin Knight. Attacking Decipherment Problems Optimally with Low-Order N-gram Models. In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, pages 812–819. Association for Computational Linguistics, 01 2008.

Sujith Ravi and Kevin Knight. Bayesian Inference for Zodiac and Other Homophonic Ciphers. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, page 239–247. Association for Computational Linguistics, 06 2011.

Josef Schmied, Claudia Claridge, and Rainer Siemund. The Lampeter Corpus of Early Modern English Tracts, 1999. URL http://hdl.handle.net/20.500.12024/2400. ICAME, Oxford Text Archive.

Ingrid Schröder. Reference Corpus of Middle Low German/Low Rhenish (1200–1650). https://corpora.uni-hamburg.de/hzsk/de/islandora/object/text-corpus:ren-1.0, 2018. Accessed: 2022-07-30.

Justyna Sikora. The Influence of Language Models on Decryption of German Historical Ciphers. Master's thesis, Uppsala University, 2022. Master thesis in Language Technology.

Deutsches Textarchiv. Grundlage für Ein Referenzkorpus der Neuhochdeutschen Sprache. Herausgegeben von der Berlin-Brandenburgischen Akademie der Wissenschaften. https://www.deutschestextarchiv.de/, 2010. Accessed: 2022-07-30.

Richard J Whitt. The Nottingham Corpus of Early Modern German Midwifery and Women's Medicine (ca. 1500-1700). 2016.

What is the Code for the Code?
Historical Cryptology Terminology

Vasily Mikhalev[1], Nils Kopal[1], Bernhard Esslinger[1], Michelle Waldispühl[2], Benedek Láng[3], Beáta Megyesi[4]
[1]University of Siegen, Germany
[2]University of Gothenburg, Sweden
[3]ELTE, Hungary
[4]Uppsala University, Sweden

Abstract

The cross-disciplinary nature of historical cryptology involves the challenge to find a terminology that is both consistent and accepted across the different disciplines and applicable in the single fields. In this paper, we propose a terminology based on concise principles developed by an interdisciplinary group of researchers. We present terms prominent in the study of historical cryptology, define them, and illustrate their usage. Our goal is to initiate and/or continue the discussion of how we use various terms for different types of historical encrypted sources, and their study. Our hope is that this paper will contribute to consistent and systematic usage of terms in the HistoCrypt community.

1 Introduction

Historical cryptology, the study of codemaking and codebreaking of historical ciphers, is a cross-disciplinary field engaging not only cryptographers and cryptanalysts but historians, linguists, computational linguists, computer scientists, computer vision specialists, codicologists, paleographers, archivists, and librarians, to name a few. Each field has its own angle and methodology to find the answers to the research questions of their interest. This might include the study and the interpretation of the ciphers, ciphertexts, codes, keys, nomenclators, nomenclatures, or codebooks. The task is not easy given the wide range of time periods, geographic areas and languages covered by the sources.

The encrypted material evolved over the centuries; many types of linguistic entities have been encrypted from letters, syllables and morphemes to named entities, words and phrases, with different code structures including various alphabets, digits, or graphic signs with fixed and/or variable length of codes.

Over the years, we have seen numerous studies dealing with historical encrypted sources, many with their own usage of specific terminology, defined or left to be interpreted by the readers. The problem is further complicated by the fact that the meaning of terms has changed over time and some terms have multiple meanings not only across but also within the same study. Additionally, variation in British and American English might also create confusion.

In light of the above mentioned reasons and challenges, we present hereby a proposal of terms and their definitions for describing the most common concepts related to ciphertexts on one hand, and cipher keys on the other. Our long-term goal is to create a consistent terminology for historical cryptology which fits various scientific fields involved and which covers and allows for expressing the most common concepts in our field.

While there have been previous attempts to introduce more or less consistent terminology for historical cryptology, such as (Meister, 1906; Friedman, 1959; Employees of Bletchley Park, 1945; Kahn, 1996; Schmeh, 2018; Dunin and Schmeh, 2020), we believe our proposal is unique in its actuality, and its well-defined structure grown out to be a compromise between experts from various scientific disciplines. Our aim was not to reconstruct the historical actors' categories, i.e. how they referred to the various elements of the encryption process. Nor was it to create a terminology that is primarily applicable to modern ciphers. Rather, we aimed at introducing consistent and modern terminology that is applicable to the historical ciphers. However, our aim is not only to make historical-cryptology terminology consistent, but also adequate, unambiguous, and simple to use in order to be able to become a standard. To achieve our goals, we tried to be specific without being too complex so that people without a background in the field can read and hopefully also apply the terminology suggested in our work.

Last but not least, we would like to encourage the community to continue the discussion about terminology issues. Our hope is that the community will adapt the proposed terms systematically in the future whenever suitable and appropriate. Needless to mention, we are open to changes and welcome feedback and suggestions for improvements. After all, standards are not given from scratch but emerge by systematic usage by many people.

In the following, we start by presenting previous attempts to describe terminology related to historical cryptology. In Section 3, we describe the principles behind our proposal, followed by a description of the usage of the terminology. In Section 4, we introduce the terms with their definition, and in Section 5 we discuss our reasoning, problems and some shortcomings of our approach. Lastly, in Section 6, we conclude the paper and give some directions for future work.

2 Related work

In this section, we present related work in the field of terminology for historical cryptology. Various researchers, authors, and cryptanalysts faced the same problem we did. They were writing about historical cryptographic topics or are part of cryptologic history themselves (e.g. because they worked in Bletchley Park). A metalanguage was needed for the description and study of historical documents, and many developed their own terminology or even applied terms without explicitly defining them. In this section, we briefly present the most prominent and important examples to the best of our knowledge.

In 1906, Aloys Meister wrote the most comprehensive collection and analysis of Papal ciphers in his German work ”Die Geheimschrift im Dienste der päpstlichen Kurie von ihren Anfängen bis zum Ende des XVI. Jahrhunderts” (Engl. ”The secret writing in the service of the papal curia from its beginnings to the end of the XVI century”) (Meister, 1906). Mainly focusing on papal ciphers, he used terms like ”Geheimschrift” (Engl. ”secret writing” or simply cipher), ”Nomenklator” (Engl. nomenclator), and ”Trugbuchstaben” (Engl. letters of deception) which we today know as ”nulls”.

The ”Bletchley Park Cryptographic Dictionary” (Employees of Bletchley Park, 1945) from 1944 is another work that introduces terminology of cryptology. A text reproduction of this dictionary can be found on Tony Sale’s webpage[1]. The dictionary features a broad set of terms used by Bletchley Park employees in their daily work, e.g. Bombe (a cryptanalyical machine to break daily keys used with the German Enigma) or Tunny (a ”German electric letter-subtractor, or virtual letter-subtractor, cipher machine using the teleprinter alphabet”). As can be seen, many terms were developed and listed quite specifically for analyzing World War II cipher machines.

William Friedman was one of the first who considered cryptology as a scientific field in its own right. He developed the idea that cryptology consists of two (main) parts: cryptography, which is the making of ciphers, and cryptanalysis which is the breaking of ciphers. Furthermore, he defined other important terms as part of cryptology, such as traffic analysis which is the analysis of communication flows. He wrote two book series: ”Military Cryptanalysis” (Friedman, 1959) as single author, and ”Military Cryptanalytics” (Friedman and Callimahos, 1985) which he co-authored with Lambros D. Callimahos. The books of both series were classified and only published for government use in the past. The ”Military Cryptanalysis” series as well as the first two books of the ”Military Cryptanalytics” series have been declassified. In ”Military Cryptanalytics”, a comprehensive glossary of the terms used in cryptology is presented. Their terminology was created for training of NSA and military cryptanalysts for the cryptanalysis of military ciphers.

David Kahn’s ”The Codebreakers” (Kahn, 1996) first published in 1967 is one of the most famous standard works on historical cryptology. It has inspired many researchers to become passionate about historical cryptology. In the introductory chapter of his book, Kahn introduces his terminology. He states that ”Cryptology is the science that embraces cryptography and cryptanalysis, but the term ”cryptology” sometimes loosely designates the entire dual field of both rendering signals secure and extracting information from them” (Kahn, 1996). Kahn introduces many, nowadays standard, and widely used cryptologic terms, e.g.

[1] The 1944 Bletchley Park Cryptographic Dictionary: `https://www.codesandciphers.org.uk/documents/cryptdict/`

plaintext, ciphertext, and cipher. Moreover, he introduces different types of ciphers such as monoalphabetic and polyalphabetic ciphers. His terminology is, of course, mainly needed to describe the historical cryptologic methods and practices presented in his book.

Two authors who started to use mathematical terms describing classical ciphers are Alan G Konheim (Konheim, 1981) and F L Bauer (Bauer, 1997).

Recently, Klaus Schmeh presented an overview of relevant terms and definitions in his blog (Schmeh, 2018) and later in the glossary of his and Elonka Dunin's book *Codebreaking: A Practical Guide* (Dunin and Schmeh, 2020). Schmeh was one of the first to point out the lack of consistency in the terminology of historical cryptology and made significant contributions to raise awareness for terminological issues and to initiate a discussion of standardization of terminology in the field.

Another glossary for Historical cryptology is available on the "Portal of Historical Ciphers" (Antal, 2018), which is a website developed and maintained by Eugen Antal, where a database of historical ciphers and keys, as well as tools for document analysis are provided. The terms presented in the glossary (Antal and Zajac, 2020) are taken from the aforementioned works by Schmeh (Schmeh, 2018) and Friedman (Friedman and Callimahos, 1985).

The release of the DECODE database (Megyesi et al., 2019) including a large collection of historical ciphers and keys with a description of the metadata about their origin, source, and characteristics led to the introduction of new terms, and the refinement of some others. For example, the distinction between plaintext (the underlying non-encrypted text) and cleartext (a non-encrypted part in the encrypted document) was suggested which is established today.

The transcription guidelines developed for ciphertexts and keys (Megyesi, 2020) and (Megyesi and Tudor, 2021) provide a further attempt to explain important terms of the field, but from a visual and paleographic (i.e. the study of handwriting) point of view to suggest consistent transcription of images of encrypted sources.

Important terms used in practice for explaining the structure of keys and the cryptanalysis of ciphertexts in the Papal correspondence during the 16th and 18th centuries in the Vatican have been introduced and explained in more or less detail in the paper by (Lasry et al., 2020).

Another recent article dealing with the study of the evolution of cipher keys presented an extensive description of the content of cipher keys originating from early modern times. The paper describes the plaintext as well as the code structure providing detailed descriptions of the content of keys (Megyesi et al., 2022).

Chapter 2 of the book (Esslinger, 2023), written by the same authors as this paper, describes historical cryptology and discusses the corresponding terminology.

The work and the terminology presented above serve as the basis for our suggestion for terminology for historical cryptology, which we describe next.

3 Creating terminology

Introducing and defining terms to create a nomenclature or terminology for scientific fields requires expert knowledge. Identifying frequently used terms in various contexts and interpretations as well as knowing the uncommon terms are indispensable. In order to succeed in acceptance by the public, adapting the terms to readers of various backgrounds and scientific fields is as important. Below, we reveal our reasoning and considerations that finally led to the principles we applied when developing the terminology for historical cryptology.

3.1 Principles

The main documents in consideration of our study are the plaintext, the ciphertext, and the cipher key. When we describe the work related to encrypted sources, we think of two sides of the coin: the ciphertext side with the code structure, and the plaintext side with the underlying text message consisting of linguistic entities. Our suggestion for terminology and the following five basic principles behind are based on the two different parts.

Symmetry Historical cryptology primarily focuses on ciphers, which are algorithms that convert plaintext to ciphertext by applying encryption and back to plaintext through decryption. One can easily see that this process is somewhat symmetric, which means that most of the concepts discussed in the scope of historical cryptology usually have their related expression on the other side

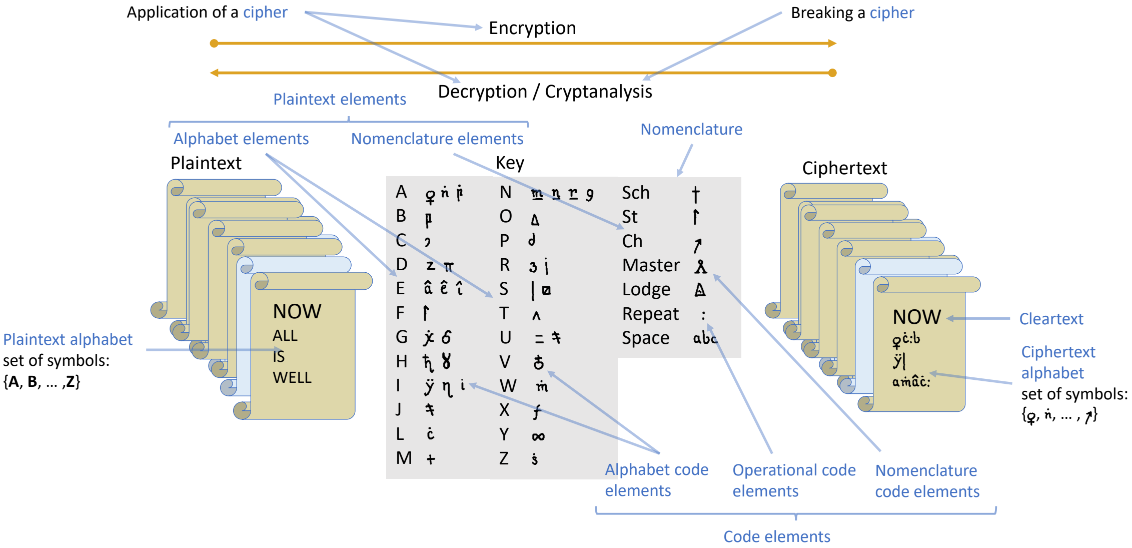


Figure 1: Most of the terms indicated together

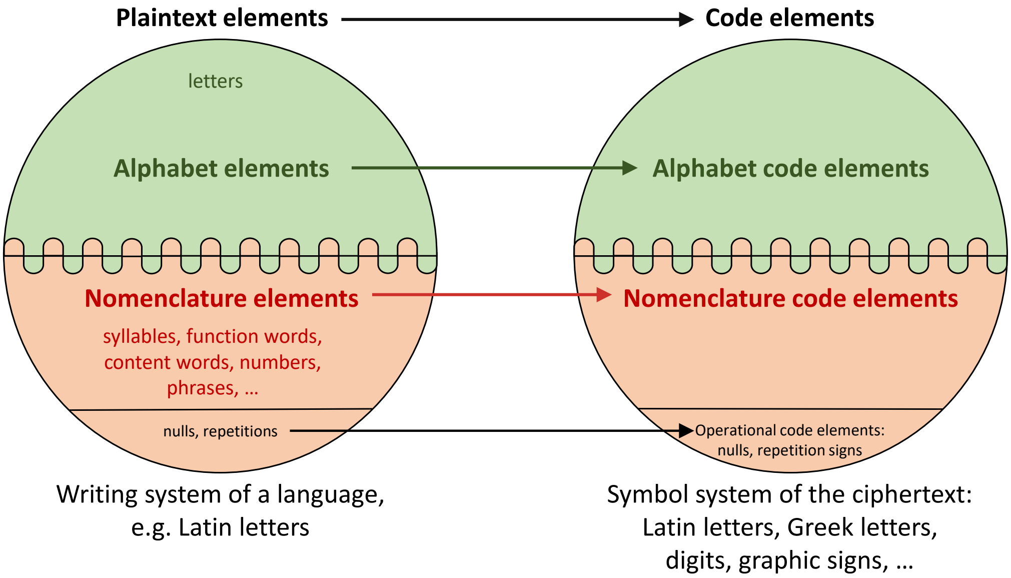
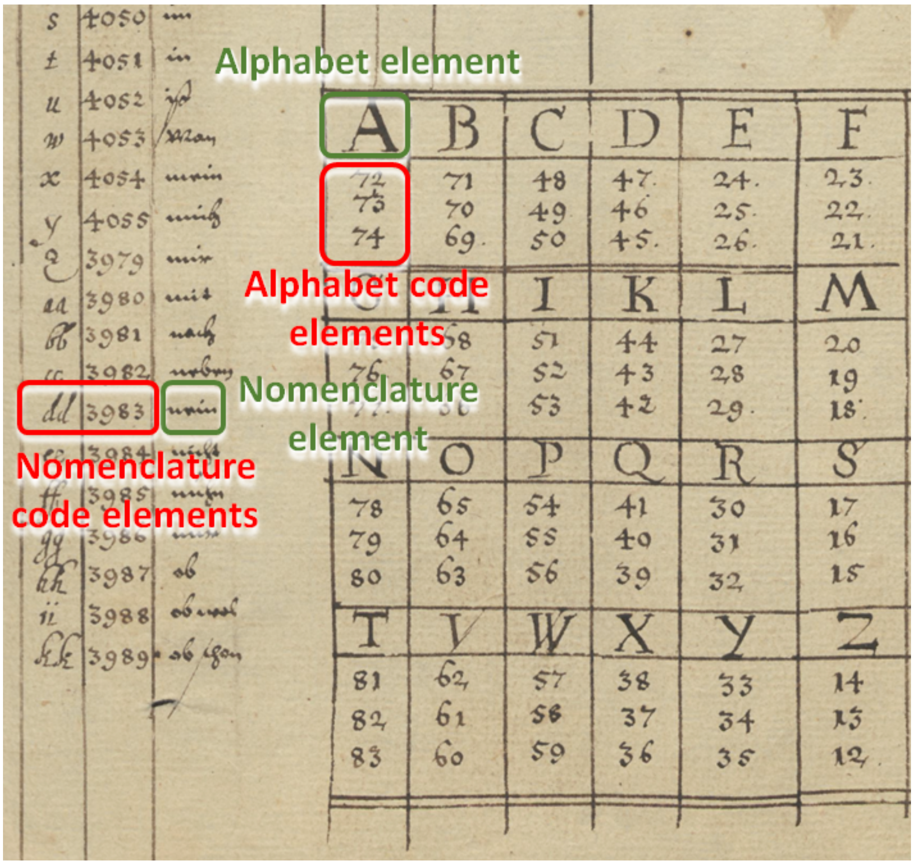

Plaintext elements
Code elements
letters
Alphabet elements
Alphabet code elements
Nomenclature elements
syllables, function words, content words, numbers, phrases, …
Nomenclature code elements
nulls, repetitions
Operational code elements: nulls, repetition signs
Writing system of a language, e.g. Latin letters
Symbol system of the ciphertext: Latin letters, Greek letters, digits, graphic signs, …

Figure 2: Mapping of the corresponding terms

of the encryption/decryption process. The goal is to ensure that the proposed terms adhere to this principle, which helps to create a well-defined structure for the terminology, as illustrated in Figures 1 and 2. In other words, we tried to think of terms in terms of pairs: on the ciphertext side, and on the plaintext side.

Explicitness Despite making sure that the related terms for each of the sides are proposed, it's also helpful if the one who sees the term for the first time immediately has a clear understanding to which side it refers. We use the term ”code” specifically when referring to elements that are present in the ciphertext side, in order to provide immediate clarity and understanding for readers.

Hierarchy While developing the proposed terminology, we aimed to create a more organized system by clearly indicating when a group of elements is a subgroup of a larger group. Whenever possible we try to show the relationship between the terms of the same side, as represented in Figures 1 and 2.

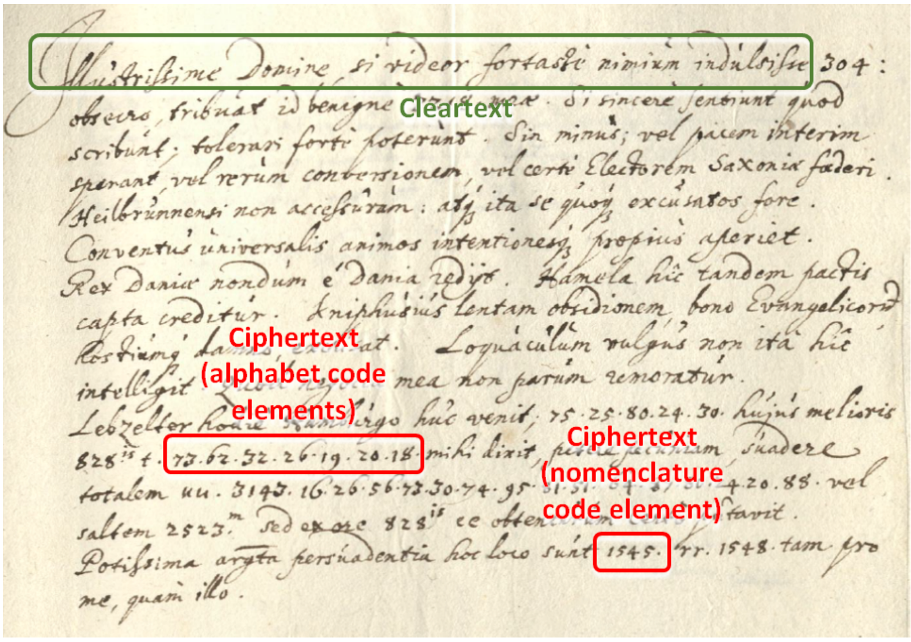


Alphabet element
Alphabet code elements
Nomenclature element
Nomenclature code elements

s	4050	[illegible]
t	4051	[illegible]
u	4052	[illegible]
w	4053	[illegible]
x	4054	[illegible]
y	4055	[illegible]
z	3979	[illegible]
aa	3980	[illegible]
bb	3981	[illegible]
cc	3982	[illegible]
dd	3983	[illegible]
ee	3984	[illegible]
ff	3985	[illegible]
gg	3986	[illegible]
hh	3987	ob
ii	3988	[illegible]
kk	3989	[illegible]

A	B	C	D	E	F
72 73 74	71 70 69.	48 49 50	47. 46 45.	24. 25. 26.	23. 22. 21.
G	**H**	**I**	**K**	**L**	**M**
[illegible] 76 77	68 67 [illegible]	51 52 53	44 43 42	27 28 29.	20 19 18.
N	**O**	**P**	**Q**	**R**	**S**
78 79 80	65 64 63	54 55 56	41 40 39	30 31 32	17 16 15
T	**V**	**W**	**X**	**Y**	**Z**
81 82 83	62 61 60	57 58 59	38 37 36	33 34 35	14 13 12.

Figure 3: Example of a key: key used in Swedish diplomatic correspondence in the 1630ies, Riksarkivet Sweden, Chifferklaver II:24.



Cleartext
Ciphertext (alphabet code elements)
Ciphertext (nomenclature code element)

Illustrissime Domine, si videor fortasse nimium indulsisse 304:
obsecro, tribuat id benigne [illegible] mea. Si sincere sentiunt quod
scribunt; tolerari forte poterunt. Sin minus; vel pacem interim
sperant, vel rerum conversionem, vel certe Electorem Saxoniæ fœderi
Heilbrunnensi non accessuram: atque ita se quoque excusatos fore.
Conventus universalis animos intentionesque propius aperiet.
Rex Daniæ nondum e Dania rediit. Hamela hic tandem pactis
capta creditur. Kniphusius lentam obsidionem bono Evangelicorum
hostiumque damnis, [illegible]at. Loquaculum vulgus non ita hic
intelligit [illegible] mea non parum remoratur.
Lebzelter [illegible] Hamburgo huc venit; 75.25.80.24.30. hujus melioris
828is t. 73.62.32.26.19.20.18. mihi dixit, petere pecuniam suadere
totalem uu. 3143. 16.26.56.73.30.74.95 [illegible] 4.20.88. vel
saltem 2523m. Sed ex ore 828is ce obten[illegible]tavit.
Potissima argta persuadentia hoc loco sunt 1545. rr. 1548. tam pro
me, quam illo.

Figure 4: Example of a ciphertext: letter written by Adler Salvius to Axel Oxenstierna in 1633 using the key in Figure 3, Riksarkivet Sweden, Oxenstierna samlingen E 708:28.

Unambiguity Some terms, e.g. the word "code" have numerous meanings in various scientific fields, making it a source of confusion for readers. It is therefore recommended to avoid using it as a standalone term and instead provide more specific context or terminology to avoid misunderstandings. Thus, we try to avoid using terms that have various meanings in different disciplines.

Simplicity Our last, but nonetheless important goal was to make sure the text written using the proposed terminology is easily readable by people with various backgrounds.

3.2 Terminology usage

The full list of the proposed terms with their definitions is given in the Section 4.

To enable easier understanding of our proposal, most of these terms are illustrated in Figure 1. We will now explain how these terms are applied and how they relate to each other. A *cipher* is the algorithm used for encryption or decryption of information. The text which is meant to be encrypted is called a *plaintext*. The resulting encrypted text is known as *ciphertext*, which is made up of symbols from a *ciphertext alphabet*. Sometimes, an encrypted document may also contain non-encrypted text, known as a *cleartext*.

The process of encryption is controlled by the cipher *key*, and when the key is known, the ciphertext can be easily decrypted. Without the key, the process of analyzing the ciphertext to reveal the original plaintext is known as *cryptanalysis*.

Historical keys are typically composed of *plaintext elements* and their corresponding *code elements*. The plaintext elements are divided into two categories: alphabet elements (single letters) and the nomenclature elements (representing entities above the alphabet level). Similarly the code elements are composed of *alphabet code elements* and *nomenclature code elements*. The *nomenclature* is a part of a key which contains the nomenclature elements and their corresponding code elements.

Some keys also contain *empty code elements*, which are placeholders that can be filled in later, and *operational code elements*, which have special functions to carry out an operation on the revealed plaintext. Examples of operational code elements include *nulls*, which are fake code elements that encode an empty string in the plaintext, and *cancellation signs*, which mark the removal of a certain sequence of ciphertext. The relationship between different plaintext elements and code elements is shown in the Figure 2.

4 Proposed terms and their definitions

We propose to use the following terms:

Plaintext
The text intended for encryption and/or the decrypted text.

Cleartext
Intentionally unencrypted text in an encrypted document.

Ciphertext
The encrypted text.

Encryption
The process of transforming plaintext into ciphertext using a key.

Decryption
The process of transforming ciphertext into plaintext using a key.

Cipher
A set of rules (algorithm) describing the process of encryption/decryption.

Key
A piece of information needed for encryption and decryption. A key has to be kept secret for security.

Cryptanalysis/Codebreaking
The process of analyzing a ciphertext without knowing or partially knowing a key to reveal the original plaintext (and key).

Plaintext alphabet
The set of elements used in the plaintext, e.g. letters, digits, punctuation marks, and spaces.

Ciphertext alphabet
The set of symbols used in the ciphertext (e.g. digits, Latin and Greek letters, alchemical or Zodiac signs). We find these symbols not only in the ciphertext but also in the manuscript containing the key.

Plaintext element
Any type of plaintext entity that has a corresponding code element assigned to it. It can represent a letter, double letter, syllable, name, function (e.g. preposition), or content word (e.g. noun, verb) as well as a phrase. The set of plaintext elements includes the alphabet and nomenclature elements.

Alphabet element
Any letter in the alphabet of the writing system

that has a corresponding code element assigned to it. Alphabet elements constitute a subset of plaintext elements.

Nomenclature element
A plaintext element which is above the alphabet level. A nomenclature element can be a syllable, a name, a function and a content word as well as a phrase.

Code element
A symbol or a concatenation of symbols of the ciphertext alphabet used during the encryption for substitution of the corresponding plaintext element or to indicate that an operation on the revealed plaintext is needed. We distinguish between the following types of code elements: alphabet code elements, nomenclature code elements, and operational code elements.

Alphabet-code element
Code element used for encryption of one or several alphabet elements.

Nomenclature-code element
Code element used for encryption of a nomenclature element. Nomenclature elements are often encrypted using a different symbol type or of a different length than used for the alphabet code elements.

Nomenclature
A part of the key with a list of nomenclature elements and the corresponding nomenclature code elements.

Empty code element
Code element presented in the nomenclature which doesn't have any plaintext element assigned to it and is treated as a placeholder to be filled in later.

Operational code element
A code element that has a special function to carry out an operation on the revealed plaintext. Examples are repetition signs, cancellation signs, and nulls.

Repetition sign
An operational code element which indicates that the preceding letter in the revealed plaintext has to be repeated.

Cancellation sign/Nullifier
An operational code element which indicates that a certain sequence of a ciphertext (and hence the corresponding revealed plaintext) is to be removed.

Null/Nullity/Nullity sign/Blender
An operational code element which represents an empty string in the plaintext. Their purpose is to confuse the codebreaker or to mark the start and/or the end of the nomenclature elements.

Code separator / Token separator
A symbol or a concatenation of symbols that separates code elements or groups of code elements from each other. The main intention is to help the receiver to tokenize the ciphertext. In the case of cryptanalysis, it can help to break the cipher more easily.

5 Discussion

In this work, we discuss the terms which refer to cryptographic concepts that were actual before the 20th century when the widespread application of cipher machines began. Moreover, we are focused on the elements that are found in the ciphertexts and keys, which is only a part of the entire historical cryptography terminology.

We start by explaining some of the issues that we faced while designing our solution and provide the reasons for our decisions.

While working on the proposed terminology, we had to deal with trade-offs between perfect structure and simplicity. For instance, when referring to the elements on the ciphertext side it would be logical to use the term ”ciphertext elements”. Nevertheless, we use the term ”code elements” which is commonly used in the area of historical cryptology. Moreover, it is shorter and easier to remember.

Another example is that we recommend using ”nulls/nullities” without the word ”sign” as it is also an already well-established term.

We also point out that ”empty code elements” are not included in the set of ”operational code elements”. In fact, they do not indicate any operation, but rather refer to nomenclature code elements without a concrete plaintext element assigned yet.

Finally, we would like to mention that there is no strict border between the ”alphabet elements” and ”nomenclature elements.” Often some nomenclature elements were presented in the same table as the alphabet elements. This is the reason why there is a curvy line between these two sets in Figure 2.

We now describe certain common cases where the inconsistent usage of terms may lead to confusion.

One of the most frequent examples is that in everyday life the word ”cipher” is used in the meaning of the ”ciphertext”. However, in scientific works, mixing these terms can lead to inconsistency or even misunderstanding. Hence, we would like to see this tradition be stopped.

Other terms that sometimes are used differently while other times as synonymous expressions are the ”nomenclature” vs ”nomenclator”. They might indicate a shorter or longer list of words with code elements, or the entire cipher key containing such a list.

The relation between the terms ”key” and ”cipher” may also become a source of confusion. In the scope of historical cryptology, these two words may sometimes be used with relatively close meanings. Given that for substitution ciphers, the key completely defines the concrete cipher, the evolution of key types also resulted in the parallel development of the corresponding ciphers. Nevertheless, the terms ”key” and ”cipher” have different meanings and should not be mixed.

Finally, we find the terms ”encipherment” and ”decipherment” problematic due to their ambiguous interpretations. For the term decipherment, the range of possible meanings varies from the synonym of ”decryption” which assumes the straightforward application of the cipher and the knowledge of the key, to ”cryptanalysis” where the plaintext and the key are revealed from the ciphertext. It may also mean an umbrella term for converting the ciphertext to the plaintext either using the key or by applying cryptanalysis. Additionally, the pair ”encipherment/decipherment” does not follow the symmetric principle, unless the ”decryption” is meant by ”decipherment”. Thus we believe that their usage may lead to confusion and the more specific terms decryption vs cryptanalysis would be preferable instead.

6 Summary

In this paper we presented a terminology for historical cryptology, focusing on the plaintext and the ciphertext sides of the encrypted sources. Our intention has been to contribute to a more systematic and consistent use of various terms, which we believe is especially important given the cross-disciplinary nature of our field. The terminology presented here is based on five principles: symmetry, expliciteness, hierarchy, unambiguity, and simplicity.

Our work presented in this paper shall be treated as just is: a work-in-progress and a point of departure to continue the discussion that started in 2018 in the HistoCrypt community. Noteworthy is that we only considered elements in the ciphertext and plaintext side, based on early modern ciphers. We have not suggested terminology for various types of operations, nor of cipher types, which would be the next step forward.

Acknowledgments

This work has been supported by the Swedish Research Council, grant 2018-06074, DECRYPT – Decryption of Historical Manuscripts.

References

Eugen Antal and Pavol Zajac. 2020. Hcportal overview. In *In Proceedings of the 3rd International Conference on Historical Cryptology. (HistoCrypt 2020)*, pages 18–20.

Eugen Antal. 2018. HC Portal. Portal of Historical Ciphers. `https://hcportal.eu/`.

Friedrich L Bauer. 1997. *Decrypted Secrets – Methods and Maxims of Cryptology*. Springer, 2 edition.

Elonka Dunin and Klaus Schmeh. 2020. *Codebreaking: A practical guide*. Hachette UK, London.

Employees of Bletchley Park. 1945. *A Cryptographic Dictionary*. NR 4559, Historic Cryptographic Collection, Pre-World War I Through World War II, Record Group 457, The National Archives and Records Administration (NARA) 8601 Adelphi Road, College Park, Maryland.

Bernhard Esslinger. 2023. *The CrypTool Book*. ArtechHouse.

William Frederick Friedman and Lambros D Callimahos. 1985. *Military Cryptanalytics*. Aegean Park Press.

William F Friedman. 1959. Military Cryptanalysis. *Aegean Park Press*.

David Kahn. 1996. *The Codebreakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet*. Scribner, New York, NY.

Alan G Konheim. 1981. *Cryptography, a primer*. John Wiley & Sons, Inc.

George Lasry, Beáta Megyesi, and Nils Kopal. 2020. Deciphering Papal Ciphers from the 16th to the 18th Century. *Cryptologia*, pages 479–540.

Beáta Megyesi and Crina Tudor. 2021. Transcription of Historical Ciphers and Keys: Guidelines, version 2.0. https://cl.lingfil.uu.se/~bea/publ/transcription-guidelines-v2.pdf. Version: March 30, 2021.

Beáta Megyesi, Nils Blomqvist, and Eva Pettersson. 2019. The DECODE Database: Collection of Ciphers and Keys. In *Proceedings of the 2nd International Conference on Historical Cryptology, HistoCrypt19*, Mons, Belgium.

Beáta Megyesi, Crina Tudor, Benedek Láng, Anna Lehofer, Nils Kopal, Karl deLeeuw, and Michelle Waldispühl. 2022. Keys with nomenclatures in the early modern europe. *Cryptologia*.

Beáta Megyesi. 2020. Transcription of Historical Ciphers and Keys. In *Proceedings of the 3rd International Conference on Historical Cryptology, HistoCrypt20*, Budapest, Hungary.

Aloys Meister. 1906. *Die Geheimschrift im Dienste der Päpstlichen Kurie von Ihren Anfängen bis zum Ende des XVI. Jahrhunderts*, volume 11. F. Schöningh.

Klaus Schmeh. 2018. Revisited: A terminology for codes and nomenclators. https://scienceblogs.de/klausis-krypto-kolumne/2018/10/07/revisited-a-terminology-for-codes-and-nomenclators.

Encrypted epigraphy - the case of a mysterious inscription in the Neapolitan church of Santa Maria La Nova

Cosimo Palma
University of Naples ”L'Orientale” / Via Duomo, 219, 80139 Napoli, Italy
University of Pisa / Lungarno Antonio Pacinotti, 43, 56126 Pisa, Italy
cosimo.palma@phd.unipi.it

Abstract

This paper contains all steps regarded as necessary and relevant towards the decryption of the epigraph placed in the “Turbolo Chapel” of the Neapolitan Church of Santa Maria La Nova. The inscription has been processed by means of Python- written procedures in combination with the decryption software *AZdecrypt*, thus displaying linguistic features converging unequivocally to the hypothesis that its clear-text be encrypted by monoalphabetic substitution from a natural language, possibly alongside with transposition and polyalphabetism. A preliminary analysis on characters n-grams and vocals-consonants combinations has not succeeded yet in individuating any clear-text language among all the analysed corpora.

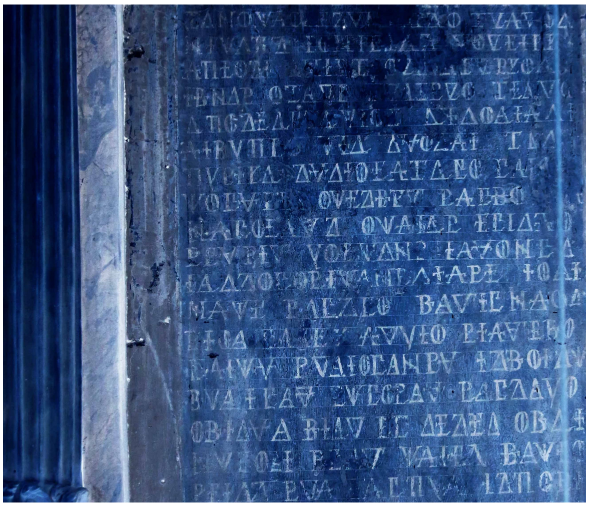


Figure 1: A snapshot of the encrypted epigraph of the Turbolo Chapel in the Neapolitan Church of Santa Maria La Nova (courtesy of the cultural association ”Oltre il chiostro”, leader manager of the monumental complex).

1 Introduction

Inside the Turbolo Chapel of Santa Maria la Nova (Naples, Italy) two epigraphs are to be found. The one on the observer's left side contains an indulgence statement from Pope Gregory XIII related to holy masses celebrated therein, in dedication to Miss Turbolo, the noblewoman Giovanna De Rosa. On the right side is placed an epigraph written in an unknown alphabet[1]. This immediately recognizable feature does not let us understand that we behold an encrypted text yet: as it often happens in these cases, the only decisive proof that we are coping with an encryption would be finding out its decryption, or at least any other external element which could let us pointing at it as such without further doubts.

2 Related Work

In few sparse newspaper articles, it emerges that serious attempts at decryption have been indeed undertaken, although related scientific publications seem impossible to find. The only two published works explicitly mentioning the inscription of the Turbolo chapel are either focused on the the artistic and architectural features of the whole church (Rocco, 1928), or on a specific set of historical events and aristocratic genealogy displayed in order to ascertain the reliability of the theory, according to which in the tomb placed outside, on the other side of the very same epigraph's wall, the count Vlad III of Wallachia (popularly know as Dracula) be buried (Miriello, 2021).

In the former work, the cultivated and generally

[1] The inscription is constituted by approximately six hundred glyphs, against the one thousand seven hundred fifty of the papal indulgence. Beside the obvious space reasons, the insertion of the whole photography would have been superfluous because of the many deletions and erosions mostly present on the top and bottom areas. Refer to section ”Historical alphabets and coeval codes review” for further details.

esteemed author assesses that the inscription is the Greek translation of the adjacent Latin inscription (and yet, even a profane would easily acknowledge that the first epigraph is too long, compared to the second, which in turns displays many glyphs obviously not belonging to the common Greek alphabet). In the latter, three conjectures about the inscription's origins are brought forth, among which I report only the most relevant one: the inscription shall be nothing else than a table used by the Franciscan monks for didactic purposes.

The church was considered in fact, at the end of the XVI. century, as one of the most flourishing university poles in Southern Italy where, among others, also oriental languages and calligraphy were taught. The same author reports that a radiocarbon-dating performed on the epigraph places it at around the XVI-XVII century and that a preliminary attempt for decryption, consisting in substituting each glyph by probable phonemes related to similar glyphs in other alphabets, has lead to no satisfactory results. It is not unusual to find mysterious inscriptions in western as well as eastern churches, although in these cases they are mostly present in the form of tetragrams (Moutafov, 2006), a text-length which cannot be compared to the one analysed in this seat.

Another article (almost homonym to the present one) first formulates the epistemological conundrum of an encrypted object situated in plain sight (Rosenmeyer, 2019). According to the author, the magical or mystic power attributed to the letters of the alphabet within oriental and Egyptian culture explains the use of acrostics in religious contexts in Egypt. Furthermore, she states that the encoder may want to tease the observer with a riddle, as an intellectual game, and not with the real purpose to conceal a message. A further scenario, as subtle as important, is the case where something is by law or enforcement compelled to be written, implicitly assuming its understandability, without expressing it, thus opening the possibility for the reluctant engraver to perform a sort of Isolt's oath, whereby the requirement is fulfilled in its form, but not in its substance. Rosenmeyer's study bears also the merit for describing a rare encryption method, based on the art of *isopsephy*. As in the rabbinic *gematria*, it associates to every letter of the Greek alphabet a numerical value, ranging from 1 to 900. By the most common encryption strategy, every sign translates in the one resulting from the difference between the numeral order immediately superior, and the numeric value itself. For instance, if Ψ equals to *700*, it shall be read as *1000* minus *700*, hence *300*, the value of the letter τ. In accordance to this calculation, all glyphs corresponding to numbers with a 5, such as 50 and 500, do not change. Despite its interestingness, I have regarded the inclusion of this approach to the present paper not relevant, since it still falls under the umbrella of a monoalphabetic substitution, which could be eventually easily investigated to assess whether it has been applied by following isopsephic principles.

3 Problem Statement

Plenty of superficial observations, such as the amount of different characters, in the same range of standard alphabets, are sufficient to assume that the epigraph is most likely the encryption of a text written in a natural language. However, the fact itself that the artefact is located in plain sight, allows us to imagine that whatever information contained therein must not be exceptionally secret, and is required to be unlocked by a group of people possessing the suitable key or a sufficient amount of time. The dense conglomerate of history and mystery the epigraph is entangled with induces first at taking a step back and deciding from which angle the decryption shall be approached. The very same is undoubtedly, at the same time, a source of motivation for undertaking this challenge.

4 Methodology

Because of the abundant literature available for the historical aspect, and of the lack of a serious quantitative analysis of the object, I decided to focus only on the tasks strictly related to decryption, in the hope that my leads, once joined with the others from different domains, could finally bring to a holistic solution of the puzzle. The major steps I undertook to attempt the epigraph decryption are:

- Review of historical alphabets and coeval codes;
- Statistical analysis on single characters;
  - Candidate-languages corpora harvesting and pre-processing;
  - Transposition of the inscription's glyphs into Latin characters;

- Preliminary analysis on vowels-consonants intertwining;
- Index of Coincidence calculation for text-snippets extracted from each corpus;
- Shannon Information Entropy calculation for text-snippets extracted from each corpus;
- Friedman's test calculation for every candidate language;

• Statistical analysis on N-grams;

- Generation of a N-grams file suitable for the *AZdecrypt* software;
- Creation of a ".ini" file for initialization of the software calculation;
- Copying the main epigraph's transliteration into the software's input window;
- Running the solver for every relevant decryption mode;
- Output files storage and analysis.

In the following, only the non-trivial steps among the above listed ones will be described in deeper detail.

5 Review of historical alphabets and coeval codes

Before diving into the canonical decryption procedure, for which I followed the path laid down in (Knight et al., 2011; Hauer and Kondrak, 2016) on the basis of (Pommerening, 2021), I first pursued a review of all existing alphabets, which substantially confirmed the already mentioned range of alphabets participating in the epigraph's composition, such as old Slavonic, Greek, Latin and Coptic (Miriello, 2021). In addition to those, the Carian alphabet[2] was found to be the alphabet which, singularly, contains the highest amount of the inscription glyphs.

Not only existing alphabets, but also invented ones have been taken in duly consideration[3], as listed in (King, 2001; Della Porta, 1563; Trithemius, 1518; Somogyi, 1906; Schöning, 2014; Meister, 1902; Kranz and Oberschelp, 2009; Rous and Mulsow, 2015). The quoted works also contain the state of the art of XVI century cryptology, which at the time was still in its dawn. If the inscription's radio-carbon dating were to be considered reliable (which is highly probable, since it matches the dates engraved on the Turbolo's tomb), the hypothesis for polyalphabetic or homophonic substitution ciphers cannot be safely discarded yet (see section "Statistical Analysis on single characters" for an empirical assessment of this question), because they first originated exactly in those years (Della Porta, 1563; Trithemius, 1518). The survey in this domain can be regarded as accomplished only if yet another set of symbols is taken in consideration, namely those which could be mapped to concepts, instead of usual alphabet characters, as in the case of alchemical signs[4].

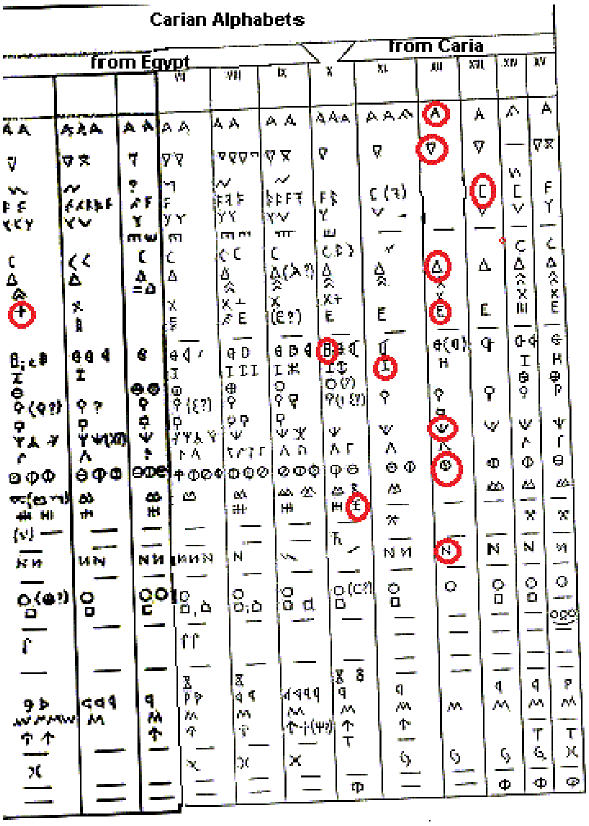

Carian Alphabets
from Egypt
from Caria
VI
VII
IX
X
XI
XII
XIII
XIV
XV

Figure 2: A diachronic display of Carian alphabets. Glyphs resembling the epigraph's ones are red-circled.

Despite the remarkable similarity with some of the symbols occurring in the epigraph, an unambiguous way to match them with single characters could not be found. For instance, taking only

[2] Carian is an extinct language spoken until the first century B.C. in Caria, a region of western Anatolia, and in Egypt (Adeigo, 2006).

[3] The two most remarkable examples, are the *Lingua Chaldeorum* of Rudolph IV Duke of Austria, and the *Lingua Ignota* of Saint Hildegard of Bingen. to which even a glossary of more than one thousand words is linked.

[4] The combination alchemy-cryptology characterises most of the Middle Ages- and Renaissance occultism.

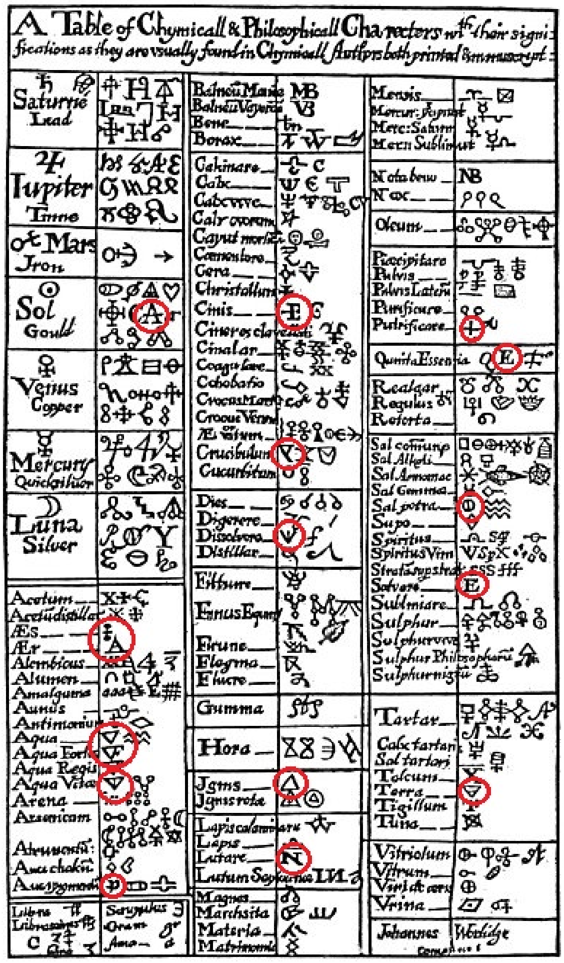
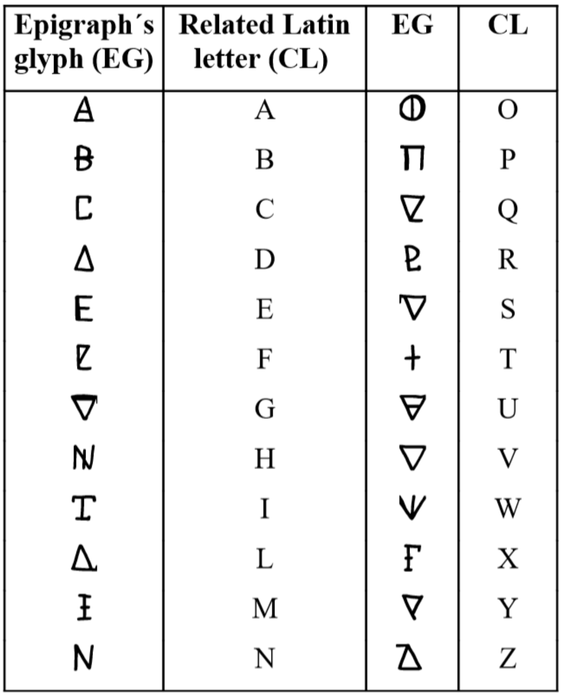


Figure 3: Table of alchemical symbols taken from "The last will and testament of Basil Valentine, monke of the Order of St. Bennet" (1671). Symbols resembling the epigraph's ones are red-circled. Source: *Wikimedia*.

the initial letter of the represented element, would result in a heavy homophony, which once practically transposed would yield no meaningful word or sentence.

6 Statistical analysis on single characters

The epigraph clearly underwent major deletions, mostly in the lower part: the entirety of text at our disposal is therefore not suited for a satisfactory statistical analysis, which usually delivers best results when performed over larger samples. However, the preliminary manual attack have been unsuccessful, thus leading to a computer-aided one[5]. The first step towards the decryption was to establish an arbitrary letter-glyph mapping in order to produce a machine-readable document. Afterwards, a handful of candidate languages have been selected according to various factors, such as their prestige at the period to which the inscription itself belongs, as well as their diversity[6], and harvested online, mostly through the online corpus database HistCorp (HistCorp, 2023)[7].



Epigraph´s glyph (EG)	Related Latin letter (CL)	EG	CL
A	A	Φ	O
Ƀ	B	Π	P
C	C	∇	Q
Δ	D	Ꝛ	R
E	E	∇	S
Ł	F	+	T
∇	G	∀	U
N	H	∇	V
I	I	Ѱ	W
Δ	L	Ϝ	X
Ɨ	M	∇	Y
N	N	Δ	Z

Figure 4: Mapping between Latin and epigraph's glyphs.

6.1 Index of Coincidence

The uncertainty drawn by these preliminary observations has induced the necessity of a statistical analysis. The related literature reports about different methods for clear-text language assessment for an encrypted text, among which the Index of Coincidence and the Shannon's Information Entropy represent the most known and efficient ones. The Index of Coincidence (IC) is a measure which reflects the probability in a given text that two randomly selected letters coincide. It is used in cryptology for the identification of the clear-text language, since every language has a relatively constant IC. In the equation of the IC, *N* represents

[5] The project's repository can be accessed at https://github.com/Glottocrisio/MariaLaNova .

[6] Considering different linguistic families knowingly accelerates the process of the clear-text language identification.

[7] The corpus for Old Luxemburgish has been obtained manually, by scanning a copy of the *Codex Mariendalensis*, a manuscript from the beginning of the XIV century about the life of Yolanda of Vianden consisting of 5,963 lines of rhyming couplets in the Moselle-Franconian German dialect.

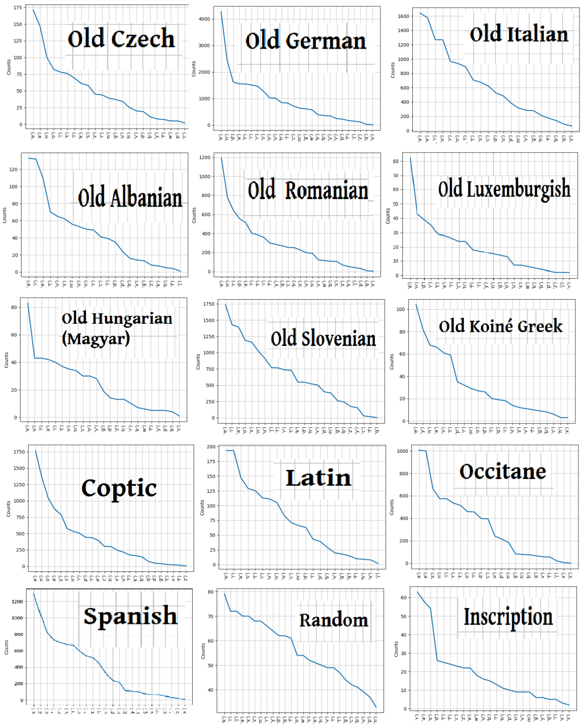


Figure 5: A synoptic view of character frequency distribution for all candidate languages, performed on random snippets with similar length to the inscription's one, extracted by related corpora.

the length of the text, and n_1 through n_c are the frequencies (as integers) of the c letters of the alphabet.

$$\mathbf{IC} = \frac{\sum_{i=1}^{c} n_i(n_i - 1)}{N(N-1)/c}$$

Nonetheless, it shall be considered, as observable from the adjacent epigraph as well, that the writing style and conventions for epigraphs is very different from the normal ones: the suppression of all doubles, as well as the use of abbreviations, automatically result in a lower IC. On the other side, the *scriptio continua*[8] enhances the same probability, since the final letter of a word can match the initial letter of the following ones. In order to parametrize these discrepancies in the IC equation, it would be needed to analyse these changes on a larger corpus of epigraphs, written in both styles, a task which is not necessary to undertake for this work. The IC has been automatically calculated by means of the project's ”getIOC” function, which implements the above mentioned mathematical formula.

Candidate language	IC
Inscription	1.72
Random text	1.03
Magyar (Old Hungarian)	1.61
Old Luxemburgish	1.87
Old Romanian	1.71
Magyar (Old Hungarian) 2	1.75
Coptic	1.87
Koiné Greek	1.88
Latin	1.94
Old Albanian	1.73
Old Slovenian	1.52
Old Spanish	1.81
Old Italian	1.53
Old German	1.95
Old Czech	1.30

Table 1: Normalized Index of Coincidence values for every candidate language.

6.2 Other statistical tests

Another measure through which a whole language can be captured is the Shannon Information Entropy (SIE), defined as follows (Shannon, 1951):

$$H(x) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i).$$

The formula calculates the average level of *information*, inherent to the variable's possible outcomes (in our case, the characters), whereby a more *informative* outcome is intended to be the less expected one.

Other methods for the assessment of the cleartext language are statistical similarity tests, such as the χ^2 , the Kolmogorov-Smirnov or the Kullback-Leibler divergence tests, conceived to measure how similar two samples are, i.e. how probable it is that the they are generated by the same distribution function. Once again, the limited length of the inscription does not allow to take this method in serious consideration. Condition for the comparison is that the samples, algorithmically rendered as two vectors or arrays, be of the same length. Moreover, their values shall be normalized, since the character frequency distributions

[8] From Latin, literally: continuous writing, i.e. space between words is omitted. This writing style is typical of Greek epigraphy as well.

Candidate language	SIE
Inscription	~ 4.1
Random text	~ 4.7
Magyar (Old Hungarian)	~ 4.1
Old Luxemburgish	~ 4.2
Old Romanian	~ 4.3
Coptic	~ 4.2
Koiné Greek	~ 5.0
Latin	~ 4.1
Old Czech	~ 4.3
Old Albanian	~ 4.4
Old Slovenian	~ 4.4
Old Spanish	~ 4.2
Old Italian	~ 4.2
Old German	~ 4.1

Table 2: Shannon Information Entropy-value for every candidate language.

may be originated from corpora of different size[9]. After proceeding with the visualization of the letter frequency distribution for every candidate language, a Friedman's Test has been performed by means of the *CrypTool* software (CrypTool, 2023) to assess the probability of monoalphabetic substitution, which as expected yielded the same result for every language (except for English, where a slight possibility for homophony is still contemplated). It has to be highlighted that homophonic ciphers usually display more than twenty-four letters, while polyalphabetism evens out characters frequency towards distributions typical of random texts. Magyar, Old Albanian and Old Romanian are the languages whose IC more resemble the inscription's.

The same calculation performed over text samples of different length, for any language, shows that the IC may vary even significantly, which compels us to consider this result as an indication, more than a prediction. The most important takeaway of this kind of analysis is that a randomly generated text shows a completely different value. The same point can not be made for the SIE, whereby the value related to the Greek language results even higher than a random text. Moreover, the lack of significant discrepancies among the entropy values suggests that this kind of measure, at least in our case, can not be effectively exploited alone in discerning between randomness and meaningfulness of a clear-text language.

7 Statistical analysis on N-grams: a brief excursus on the AZdecrypt software

The cryptologic attack based on N-grams was supported by *AZdecrypt* (AZdecrypt, 2023), the same software used to decrypt the famed Zodiac cypher[10]. Although it is not the most user-friendly and advanced software at disposal in terms of supporting community and releases periodicity, it is extremely flexible to incorporate new corpora.

AZdecrypt is conceived for modern cryptanalysis, especially for homophone ciphers, nevertheless it is easily adaptable to historical ciphers[11]. The n-grams vocabularies included in AZdecrypt are formatted in binary, yet it is possible to include them through a textual file. In the project *MariaLaNova* the function which generates a N-grams file suitable to be processed by *AZdecrypt* can be run by the command:

ngramsAZ(file, 5, case = ”lower”)

, whereby the first parameter is a corpus, the second one the desired *N*, and the third, optional, for rendering the output in lower- or uppercase. The output file will enlist each N-gram immediately followed by its log value, a number between 0 and 255 obtained by:

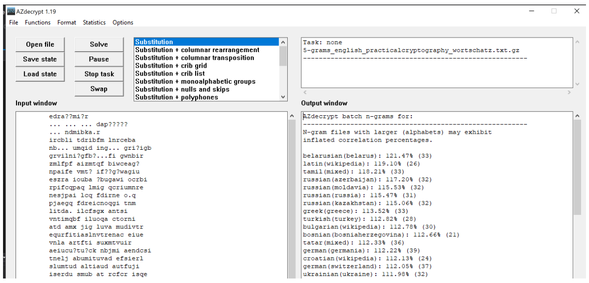
$$log_{10}(ngramfrequency_{corpus}) * 10$$

All N-grams followed by ”000” could be removed. The GUI library used for AZdecrypt does not support Unicode. Hence, only languages that can be represented in ASCII are visually supported. A workaround for this problem is substituting Unicode with ASCII and then providing a ASCII to Unicode mapping table in the n-gram *.ini* file. The *.ini* format is used for simple text files containing initialization parameters. In *AZdecrypt*, it accompains in the ”Ngrams” folder every N-gram file. Its appearance for Persian, a language not supported by Unicode, is:

[9] The procedures yielding these statistical tests' results are to be found in the project *MariaLaNova*, under the file ”Graphs.py”.

[10] https://www.iflscience.com/fbi-confirms-zodiac-killers-infamous-340-cipher-has-been-decoded-and-his-message-finally-revealed-62044 .

[11] The repository containing all the outputs, sorted by language and decryption mode can be accessed at https://github.com/Glottocrisio/AZDecryptMariaLaNova .



```
Task: none
5-grams_english_practicalcryptography_wortschatz.txt.gz
------------------------------------------------------------
```

```
        edra??mi?r
        ... ... ... dap?????
        ... ndmibka.r
        ircbli tdribfm lnrceba
        nb... umqid ing... gri?igb
        grvilni?gfb?...fi gwnbir
        zmlfpf aizmtqf biwceag?
        npaife vmt? if??g?wagiu
        eszra iouba ?bugawi ocrbi
        rpifcqpaq lmig qcriumnre
        nesjpai lcq fdirne o.q
        pjaegq fdreicnoqgi tnm
        litda. ilcfsgx antsi
        vntimqbf iluoqa ctorni
        atd amx jig luva mudivtr
        equrfitiaslnvtrenac eiue
        vnla artfti suxmtvuir
        aeiucu?tu?ck nbjmi aendcsi
        tnelj abumituvad efsierl
        slumtud altiaud autfuji
        iserdu smub at rcfcr isqe
```

```
AZdecrypt batch n-grams for:
------------------------------------------------------------
N-gram files with larger (alphabets) may exhibit
inflated correlation percentages.

belarusian(belarus): 121.47% (33)
latin(wikipedia): 119.10% (26)
tamil(mixed): 118.21% (33)
russian(azerbaijan): 117.20% (32)
russian(moldavia): 115.53% (32)
russian(russia): 115.47% (31)
russian(kazakhstan): 115.06% (32)
greek(greece): 113.52% (33)
turkish(turkey): 112.82% (28)
bulgarian(wikipedia): 112.78% (30)
bosnian(bosniaherzegovina): 112.66% (21)
tatar(mixed): 112.33% (36)
german(germania): 112.22% (39)
croatian(wikipedia): 112.13% (24)
german(switzerland): 112.05% (37)
ukrainian(ukraine): 111.98% (32)
```

Figure 6: The AZdecrypt "Languages" function.

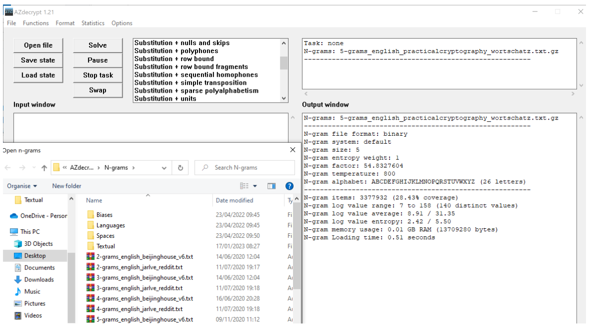
```
N-gram size=b5
N-gram factor=90.11
Entropy weight=1
Alphabet=#<*)576%4$
,3:-+?1;0(2&"!8'/.>9=
Temperature=700
```

, whereby in the first line the "b" stands for "binary"[12]. It should be deleted for all non-binary formatted N-grams files. The alphabet line shall contain all characters present in the related N-grams file. The *temperature* variable refers to the probability of accepting a modification with a lower fitness. It continuously decreases, emulating the process of annealing in metallurgy, therefore the name.

The strategy adopted in my study to avoid unsupported characters is transposing the corpus into Latin characters *before* generating the N-grams file. This is achieved by the functions contained into the file "Replace.py", and at the state of the art are available for Greek, Coptic and Cyrillic. Other alphabets can be mapped easily following the same model used for the other "Replace" functions.

Before attempting at the decryption via all modes available in the software sound protocol expects to use the function "Languages" to determine which in which clear-text language a given input cipher could be by selecting "File", then "Batch n-grams (substitution)", and by opening "Languages.azd" under "Languages". Nevertheless, it has been my previous statistical analysis to suggest me which languages should have been prioritized in my software-supported attack, namely Latin, Hungarian, Czech, Romanian, Albanian and Church Slavonic.

By clicking on "File" then "Load N-grams" the folder with all N-grams is accessed. Before running the decryption in one of the modes selected in the list above[13], on the window's right side the ".ini" file content, as well as statistical observations on the uploaded N-grams file are displayed.



```
Task: none
N-grams: 5-grams_english_practicalcryptography_wortschatz.txt.gz
------------------------------------------------------------
```

```
Biases                              23/04/2022 09:45   Fi
Languages                           23/04/2022 09:45   Fi
Spaces                              23/04/2022 09:50   Fi
Textual                             17/01/2023 08:27   Fi
2-grams_english_beijinghouse_v6.txt 14/06/2020 12:04   A
2-grams_english_jarlve_reddit.txt   11/07/2020 19:17   A
3-grams_english_beijinghouse_v6.txt 14/06/2020 12:04   A
3-grams_english_jarlve_reddit.txt   11/07/2020 19:18   A
4-grams_english_beijinghouse_v6.txt 16/06/2020 20:28   A
4-grams_english_jarlve_reddit.txt   11/07/2020 19:18   A
5-grams_english_beijinghouse_v6.txt 09/11/2020 11:12   A
```

```
N-grams: 5-grams_english_practicalcryptography_wortschatz.txt.gz
------------------------------------------------------------
N-gram file format: binary
N-gram system: default
N-gram size: 5
N-gram entropy weight: 1
N-gram factor: 54.8327604
N-gram temperature: 800
N-gram alphabet: ABCDEFGHIJKLMNOPQRSTUVWXYZ (26 letters)
------------------------------------------------------------
N-gram items: 3377932 (28.43% coverage)
N-gram log value range: 7 to 158 (140 distinct values)
N-gram log value average: 8.91 / 31.35
N-gram log value entropy: 2.42 / 5.50
N-gram memory usage: 0.01 GB RAM (13709280 bytes)
N-gram Loading time: 0.51 seconds
```

Figure 7: A typical view of AZdecrypt's working environment .

The N-grams analysis has embraced also an experiment, for which the software was not required. Moving from the observation that all analysed languages display all their vowels within their first eight letters ranked over frequency, a vocals-consonants intertwining can be investigated even if we are still unaware of the clear-text. By replacing in the epigraph *all possible* vowels (all first eight ranked glyphs) with a *V*, we can observe if the behaviour of possible vowels with consonants reflects the one of the other languages. This experiment is based on the assumption that while not all *V*s are surely vowels, all not- *V*s are consonants needs. After performing this operation, plenty of all-consonants 3-, -4 and 5-grams are to be found. Usually, three consonants in sequence are already pretty seldom, five almost impossible, in almost any language, which would constitute a strong case either for transposition, or even for randomness.

8 Results

By following the above mentioned steps, plenty of text files have been generated for every considered plain-text language, currently still under examination. Feeding *Google Translate* with some of the outputs' snippets from Latin or Hungarian, we are faced with fascinating translations of outstanding coherence. In spite of this, they cannot be correct at the same time; furthermore, *Google Translate*

[12] For a binary 4-gram file, for example, the first byte would represent the 0 to 255 (log) value of the AAAA n-gram, the second byte would be AAAB... etc. up to ZZZZ. All possible n-grams have to be included in this order.

[13] Among all, I have used only the functions for "Substitution", "Substitution + Nulls & Skips", "Substitution + simple transposition", "Substitution + sparse polyalphabetism".

showed to be extremely unreliable because meaningless words are therein often rendered with their most similar meaningful term, and translated accordingly. This experience has let me understand that a real awareness of the candidate languages is an indispensable condition for consistent improvements towards decryption. I have done my best with the languages I sufficiently know, such as Latin, Greek, Italian, Spanish and German, although even they cannot be confidently sorted out from our quest. For all the other ones, it is unavoidable gathering help from the scientific community, because only a synergistic effort can finally break this centuries-lasting riddle.

The statistical approach seems to fails at a first attempt, but there are plenty of strategies which can still be taken in account. Besides transposition and substitution, there are indeed other gimmicks which could allow to the ciphertext to maintain its statistical features, as occuring in the seventh challenge of Bellaso (Bellaso, 1553), where the characters frequency distribution resembles our inscription's, although it has been encrypted mainly by the method of *scrumbled alphabets*.

9 Concluding remarks and future work

Differently from other case studies, where the encrypted text is usually far longer and cleaner, we have been put against a cryptological challenge for which probabilistic solving heuristics may not be enough, eliciting the need for semantics-aware decryption. Nevertheless, there are still some options to be taken into consideration, still coherent with the monoalphabetic substitution, like the conjecture according to which the epigraph may be edited in more than one language. This path could by walked by creating many bilingual corpora, from which to extract the N-grams. Yet another possible scenario could be the absence of vowels, i.e. the inscription may contain only consonants.

An option which cannot be explored realistically is the use of *steganography* (Trithemius, 1518): even in that case we would fall back in the great obstacle of the used esoteric glyphs, which is even more aggravated by the established practice among coeval cryptographers to willingly insert mistakes or abbreviation in the clear text, in order to impede decryption. In case that insights from collateral elements, maybe discovered inside the church, will suggest that the inscription may contain one or more words, the *AZdecrypt* function *Substitution + Crib grid* can be used to check whether in the epigraph be contained combination of glyphs allowing that very word. This way has been already tried with terms such as ”Holy Mary” and ”Jesus Christ”, translated into the main candidate languages.

A further idea which shall not be discarded, may be the creation of a corpus entirely constituted of magical/esoteric words, harvested from grimories, glossaries of conlangs such as Hildegard's *Lingua Ignota*, magical papyri, gnostic libraries and other books about occultism and alchemy: if the alphabet was invented, no one forbids that also the related vocabulary was.

10 Acknowledgments

This work has benefited from the support of many scholars, providing me various suggestions and material for improvement. Therefore I thankfully acknowledge my supervisors Dr. Maria Pia Di Buono and Prof. Dr. Johanna Monti, as well as by Prof. Dr. Beata Megyesi and Dr. Carola Dahlke, for providing me precious critics related to the first draft's flaws in structure and content;

Dr. Richard Bean and Dr. Ivan Parisi, who have bestowed me valuable insights to substantially change the first version of the paper, in particular by shifting my attention to a more thorough statistical analysis;

Francesco Pastore and Francesco Afro Di Falco, for having shared with me very interesting considerations regarding the possible sense of the epigraph, particularly from historical and the esoteric standpoint;

Prof. Dr. Giuseppe Reale, head of the association ”Oltre il Chiostro”, which manages the museum complex of Santa Maria La Nova, for presenting me very detailed descriptions of the church (in particular concerning the architectonic elements strictly related to the analysed artefact), for granting me free entrance in the museum as well as high quality digital pictures of the inscription;

Dr. Jarl Van Eycke and David Oranchak have been very patient and generous in showing me the approach to integrate the *AZdecrypt* software with the multilingual historic corpus I needed for my inquiry.

References

I. J. Adeigo. 2006. *The Carian Language*. Brill.

AZdecrypt. 2023. Azdecrypt. `https://github.com/doranchak/azdecrypt`. Accessed: 2022-11-30.

G.B. Bellaso. 1553. *Il vero modo di scrivere in cifra*.

CrypTool. 2023. Cryptool portal - cryptography for everybody. `https://www.cryptool.org/de//`. Accessed: 2022-11-30.

G . B. Della Porta. 1563. *De Furtivis Literarum Notis*.

B. Hauer and G. Kondrak. 2016. Decoding anagrammed texts written in an unknown language and script. *Transactions of the Association for Computational Linguistics 4*, page 75–86.

HistCorp. 2023. Histcorp-historical corpora. `https://cl.lingfil.uu.se/histcorp/index.html`. Accessed: 2022-10-23.

D. A. King. 2001. *The Ciphers of the Monks: A Forgotten Number-notation of the Middle Ages*. Franz Steiner Verlag, Wiesbaden.

K. Knight, B. Megyesi, and C. Schaefer. 2011. The copiale cipher. In *4th Workshop on Building and Using Comparable Corpora:Comparable Corpora and the Web BUCC*, pages 12–19.

H. Kranz and W. Oberschelp. 2009. *Mechanisches Memorieren und Chiffrieren um 1430, Johannes Fontanas "Tractatus de instrumentis artis memorie"*. Boethius Band 59. Franz Steiner Verlag.

A. Meister. 1902. *Die Anfänge der moderne diplomatische Geheimschriften- Beiträge zur Geschichte der italienischen kryptographie des XV Jahrhunderts*.

L. Miriello. 2021. *Sulla presunta tomba di Dracula a Napoli*. I Polifemi. Stamperia del Valentino.

E. Moutafov. 2006. Translating encrypted messages: greek and slavonic tetragrams as a mixture of languages or as a universal code.

K. Pommerening. 2021. *Cryptology Part I: Classic Ciphers (Mathematical Version)*.

G. Rocco. 1928. *Il convento e la chiesa di Santa Maria La Nova*. Tipografia Pontificia degli Artigianelli, Napoli.

Patricia A. Rosenmeyer, 2019. *Encrypted Inscriptions: a Paradoxical Practice*, pages 373–392. Brill.

Anne-Simone (Ed.) Rous and Martin (Ed.) Mulsow. 2015. *Geheime Post-Kryptologie und Steganographie der diplomatischen Korrespondenz europäischer Höfe während der Frühen Neuzeit*, volume 106 of *Historische Forschungen (HF)*. Duncker & Humblot, Berlin.

F. Schöning. 2014. *Geheimschrift im deinste der päpstlichen Kurie von ihren Anfängen bis zum Ende des XVI Jahrhunderts*. Nabu Press.

C. E. Shannon. 1951. Prediction and entropy of printed english. *Bell System Technical Journal*, 30(1):50–64.

J.W. Somogyi, 1906. *Caratteristiche strutturali di cifrari monoalfabetici italiani nei secoli XIV e XV*, pages 195–213.

J. Trithemius. 1518. *Polygraphiae libri sex*.

11 Appendix

List of relevant functions from *MariaLaNova*

Freqvoc: this function takes a ".txt" file as input and returns a frequency dictionary of the words contained therein.

letterngramsfilerank: this function generates a n-grams file of a given input file ranked by frequency.

generateioccorpus: the input corpus of "ngramsaz" is created by selecting files whose Index of Coincidence is most similar to the epigraph's one by means of this function.

ngramsaz: this function takes a corpus as input (as the one generated in the previous function) and generates a ".txt" file containing the logarithmic values for all n-grams in the format requested by azdecrypt (5-grams): "exampl123xampl009ample007".

patternsearch: this function loops through a particular word list seeking for a pattern. It takes the number of characters and the pattern as parameters. 'k' means consonant, 'y' vowel and 'u' all different letters.

matchwordscrco: this function loops through a ".txt" file to find a word matching with the input word as parameter i.e.: *word*="polo" and *file*=" bebe; marmelade; mouse; poster; fata", the result will be "fata". In this version, we assume that the script contains no space (see **matchword** otherwise).

Deciphering Charles Quint (A diplomatic letter from 1547)

Cécile Pierrot
Université de Lorraine
CNRS, Inria, LORIA
F-54000 Nancy, France
cecile.pierrot@inria.fr

Camille Desenclos
Université de Picardie Jules-Verne
Centre d'histoire des sociétés,
des sciences et des conflits
F-80025 Amiens Cedex 1
camille.desenclos@u-picardie.fr

Pierrick Gaudry
Université de Lorraine
CNRS, Inria, LORIA
F-54000 Nancy, France
pierrick.gaudry@loria.fr

Paul Zimmermann
Université de Lorraine
CNRS, Inria, LORIA
F-54000 Nancy, France
paul.zimmermann@inria.fr

Abstract

An unknown and almost fully encrypted letter written in 1547 by Emperor Charles V to his ambassador at the French Court, Jean de Saint-Mauris, was identified in a public library, the *Bibliothèque Stanislas* (Nancy, France). As no decryption of this letter was previously published or even known, a team of cryptographers and historians gathered together to study the letter and its encryption system. First, multiple approaches and methods were tested in order to decipher the letter without any other specimen. Then, the letter has now been inserted within the whole correspondence between Charles and Saint-Mauris, and the key has been consolidated thanks to previous key reconstructions. Finally, the decryption effort enabled us to uncover the content of the letter and investigate more deeply both cryptanalysis challenges and encryption methods.

1 Introduction

What is required to decipher an encrypted letter which was composed five centuries ago? Modern cryptographic knowledge would normally be more than sufficient to decipher a 3-page letter. That was the first guess while discovering, in a public library, the *Bibliothèque Stanislas* (Nancy, France), an isolated encrypted letter that was written on 22nd February 1547 by Emperor Charles V to Jean de Saint-Mauris, his ambassador at the French Court. But, due to too many symbols, brute force attacks were hopeless and cleverer methods were unsuccessful. This initial failure reveals the mutual benefits for cryptographers and historians to work together in order to uncover the content of an almost fully encrypted letter and, above all, to better understand the encryption methods, first by working from scratch, then by comparing with other deciphered letters and finally by merging with former partially recovered keys.

2 General context

2.1 The story of the project

The existence of an encrypted letter of Charles V, which had not yet been deciphered, was known in literary and cultural circles in the city of Nancy, but this letter was neither properly identified nor yet digitized. Two years passed between the first mention of this letter by some acquaintances of C. Pierrot, and the moment when she was able to see it and start working on it, thanks to a word-of-mouth game to find the letter that eventually worked. A mixed team of experts, first cryptographic researchers and then historians, was formed.

Now openly available online, the letter consists of two folios: three pages of text and an address on the last page. The first lines and the last two paragraphs on the first page as well as the last lines of the third page (date and signatures) are cleartext. As for many letters produced at that time by the Imperial chancellery or by the Imperial cabinet (Stix, 1934-1936), the letter was written in French. The bulk of the document (two and a half pages) is ciphertext. We counted 1767 symbols taken from a set of 125 different ones[1] of various types: Latin characters, mathematical symbols and so on. Cleartext allowed a quick identification of the letter, written by Emperor Charles V to his ambassador at the French Court, Jean de Saint-Mauris. Sent to Francis I in 1544 as permanent ambassador, Saint-Mauris was related, by his wife, to Antoine Perrenot de Granvelle, Charles' state secretary and main counsellor in the Holy Roman Empire, and both were from Franche-Comté. This provided a substantial leverage for the decipherment process.

2.2 An isolated letter in a French public library

Although the letter meets the usual patterns of encrypted letters from the mid-16th century (diplomatic context, cleartext and ciphertext on the same page, alphabetical and numerical symbols, etc), its preservation history led us to adapt and question the traditional approach to such letters. This letter wasn't hitherto properly identified, although Saint-Mauris is well known to early modern historians, in particular because of his broad and extensive correspondence in which he reported on the complex relationship with Francis I (Cassan, 1878; Potter, 2013), either to Charles and Granvelle, to Maria of Austria (governor of the Lower Countries and Charles' sister), or to Infante Philip (Charles' son) and the state secretary for Spain, Francisco de Los Cobos y Molina. Not only was it misidentified (the date especially was wrong in the library catalog[2]) but it had no reason to be kept in Nancy. Indeed, despite first attempts of channeling the preservation of state papers in Simancas (Spain) from 1540, Saint-Mauris' letters, as many other diplomatic ones, match two common preservation patterns, depending on whether the correspondence is active (the ambassador is the one sending a letter) or passive (the ambassador is the recipient).

The active correspondence of Saint-Mauris[3] has mainly been preserved in the state archives that match the main location of the recipients [4]. It can thus be found in Vienna for the letters to Charles and Granvelle (OeStA-HHStA, Fr 10-16)[5], in Brussels for the letters to Maria of Austria (AGR Audience 420; AGR Audience 1672) and in Simancas for the letters to Infante Philip and Los Cobos[6]. Within this active correspondence, the two previous letters of Saint-Mauris to Charles have been identified: 11th February 1547 (copy) (BM Besançon, Granvelle 40)[7] and 6th February 1547 (Archives nationales, K1487)[8].

On the other hand, the passive correspondence was often kept by the ambassador with his private papers. Saint-Mauris' passive correspondence is no exception. One part is preserved at the public library in Besançon (BM Besançon, Granvelle 40; BM Besançon, Granvelle 70) but it concerns mainly the year 1545. Further letters, especially from Charles, are probably lost or scattered across Europe without global identification[9]. But if the preservation of some letters in Besançon makes sense (Saint-Mauris' correspondence is preserved along with the Granvelle collection), the existence of a single letter in Nancy is much more surpris-

[1] All the symbol counts are approximate since some symbols aren't always well formed and thus look very similar, see Appendix, Fig. 3.

[2] "1546" is written at the end of the letter. However, at that time, several dating systems could be used and the year could begin at Easter and not on January 1st. According to today's dating system, the letter was written in 1547.

[3] Unfortunately it hasn't yet been possible to check every part of this correspondence; it may thus have some deficiencies in the following presentation. Only the letters (or their copies) that are kept in Paris, Besançon and Madrid have been studied for now either directly or thanks to their digitization (Besançon, Madrid).

[4] Recipients do not always match the expected archival collections. Some letters to Maria of Austria are for instance preserved in Vienna.

[5] Seven letters to Charles and Granvelle, and especially a copy of a letter written in February 1547 can be found in Besançon (France) with Granvelle's papers (BM Besançon, Granvelle 40, fol.139, letter to Charles V, 11th February 1547). Another isolated encrypted letter from Saint-Mauris to Granvelle (1548) has been identified in the National library of Spain (Madrid) (BNE, 7913).

[6] Part of Simancas archives are accessible as microfilms at the French National archives (Paris) (AN, K1485-1488). One can found Saint-Mauris' active correspondence to Spain, some minutes from Infante Philip as well as some copies of letters from Saint-Mauris to Charles.

[7] Due to the long transmission delays, Charles hadn't yet received this last letter on 22nd February 1547. He acknowledged the reception of two letters only: 26th January and 6th February 1547.

[8] The letter in Simancas / Paris is however a copy made by the Spanish state secretary. Only plaintext and cleartext are hence transcribed

[9] According to Maxim Hoffman, PhD student in Ghent University, the minute of the 22nd February letter would still exist. David Potter (Potter, 2013) has indeed identified Charles' minutes in Vienna. If a verification couldn't be carried out for this contribution, the authors will conduct some researches in May 2023 both in the Haus-, Hof- und Staatsarchiv (OeStA-HHStA, Fr 10-16) and in the Archives générales du royaume (AGR Misc 95-96) in order to compare their decipherment with the minute and to expand their corpus of encrypted letters from and to Saint-Mauris (AGR Audience 420; AGR Audience 1672).

ing. After some research, the letter would belong to the collections since the 19th century. The library archives unfortunately do not keep tracks of the date or terms of its acquisition. One hypothesis can be formulated: part of the passive correspondence would have been scattered early, one letter bought by an erudite and then given to or bought by the library. That is consistent with the incomplete preservation of the passive correspondence for years 1546-1547[10] but confirming this hypothesis (and the loss of the other letters) will require further enquiries about Saint-Mauris' succession.

2.3 Historical context

When Charles wrote his letter on 22nd February 1547, the main European sovereigns were supposed to be at peace, while Emperor Charles V was dealing with a political and religious conflict within the Empire, the Schmalkaldic War[11]. Nevertheless, between Francis and Charles, and despite the peace treaty of Crépy (1544), war and mistrusts were not really over. The treaty provided the dispositions of former peaces and planned a marriage between Francis' first son (Francis, duke of Orléans) and Charles's daughter, Maria (or Ferdinand's daughter, Anna) while Francis committed to support Charles against the Schmalkaldic League. But one year later, Francis had not yet fulfilled his obligations and his armies were still in Piedmont and Savoy (Babel, 2013). Moreover, after Charles claimed the restitution of Hesdin or at last Thérouanne (North of France) that Francis denied, both sovereigns armed again in Italy (Milanese and Piedmont) (Nawrocki, 2015). Furthermore, in June 1546 Francis concluded the peace of Ardres with Henry VIII (England would keep Boulogne (North of France) until France paid the amount of 2 millions *écus*) (Potter, 2011) and two months later, Francis, duke of Orleans, died. Although the process was obviously more complex and non-linear, Francis' intentions moved back to war, at least against Charles: he secretly reconnected with the Schmalkaldic League (Potter, 1977) and did not push back the offer of a defensive alliance against the Emperor, which Henry VIII was also supposed to join. At the same time, the French King did some military preparations.

At the end of 1546 and beginning of 1547, uncertainties were numerous on both sides and the Imperial presence in Milan and the French one in Piedmont still fed tensions. Charles' situation in the Empire certainly became better: in January 1547, Ülrich von Württemberg came to an agreement with him and the cities of Ulm and Frankfurt submitted themselves (Potter, 2011). However, Francis' intentions were still alarming Charles, who suspected them as either resuming war in Italy or supporting Charles' opponents (League of Schmalkalde, Ottoman Empire). Indeed, undercover but separate negotiations took place between France, England and the Schmalkaldic League in order to conclude an alliance, even though tensions remained between England and France, because of Boulogne but most of all because of Scotland. In January 1547, an agreement between Henry VIII and the League was almost concluded while the negotiations about the conditions of the French financial loans (towards the League) were still ongoing (Pariset, 1981).

However, on 28th January, Henry VIII died: Edward VI, his only legitimate (but very young) son, ascended the throne. The negotiations seemed jeopardized as the new King and his ministers expressed their unwillingness to support the Schmalkaldic League (Nawrocki, 2015). French military preparations on the other hand were continuing, and after being presented in late 1546 as a defensive preparation either against the Emperor (when talking with English ministers) or against the King of England (when talking with Imperial ministers), they were by then a defensive preparation for a new war that Charles V would declare in Italy as soon as he had brought back the Empire under his authority (Potter, 2011). Francis' motives remained ambiguous for foreign informants and ambassadors like Saint-Mauris who suspected either preparations against the Emperor or preparations against the new King of England,

[10] At this stage, apart from the minutes (Vienna) and the copies of letters (Simancas, Brussels), only one other letter from Charles to Saint-Mauris has been identified for the first months of 1547 (until Francis' death in late March): David Potter (Potter, 1977) refers to a letter from 19th January 1547 which has been edited from a Viennese copy (von Druffel, 1878, p.39-45).

[11] Since 1542, several German Lutheran cities and princes whose religion was prohibited had been revolting against Charles and gathering in a League (the Schmalkaldic League) conducted by John Frederick, elector of Saxony, and Philip I, landgrave of Hesse. Thanks to the treaty of Crépy which momentarily interrupted the Italian Wars, Charles launched a military and political campaign against the League. He used the invasion of the duchy of Brunswick-Wolfenbüttel in 1542 by John Frederick and Philip as an excuse; he banished them and convinced Maurice, duke of Saxony and John Frederick's cousin, to join him in exchange of his cousin's lands and electoral dignity.

with whom negotiations about Boulogne were still ongoing in order to obtain a confirmation of the treaty of Ardres and an early return of the city.

3 Decryption methods

3.1 Names and statistics

As no other letter from Charles or Saint-Mauris was preserved in Nancy, it was first decided to work on it as a single letter in order to test the cipher and its strength. The first step was to name each of the 125 different symbols (see Fig. 1). These names were useful to identify several occurrences of a particular symbol, distinguish families of similar ones, and record our observations (statistics, patterns...). Later, it was also necessary for a computer treatment of the ciphertext. Our first observation was that among those 125 symbols, 50 were "simple" ones, and 75 were "complex" ones, i.e., there is at least one occurrence in the ciphertext of this symbol with a dot or a hyphen around it (examples are shown in Fig. 1). Among the 75 complex symbols, there were only 17 "root" symbols (without any dot or hyphen), for example aire. Among the 50 simple symbols, we noticed that 8 symbols appeared only once.

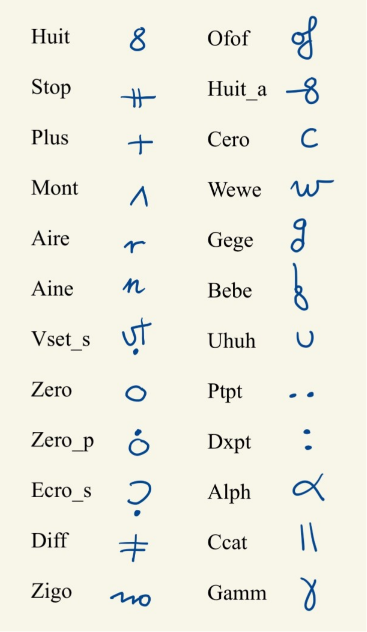

Huit
Stop
Plus
Mont
Aire
Aine
Vset_s
Zero
Zero_p
Ecro_s
Diff
Zigo
Ofof
Huit_a
Cero
Wewe
Gege
Bebe
Uhuh
Ptpt
Dxpt
Alph
Ccat
Gamm

Figure 1: A sample of symbols and their names. The symbols stop, plus and mont are simple symbols, whereas vset_s, zero, and zero_p are complex symbols.

After re-encoding the ciphertext as a list of strings in the Python computer language, we ran small programs to get quick and reliable confirmations of our observations and intuitions. First we analyzed the frequency of each symbol, and sequences of two or three symbols. The most frequent symbols are[12] huit (8.3%), plus (7.8%), and stop (6.2%). The most frequent bigrams are huit stop (2.2%), mont huit (1.7%), and huit plus (1.3%). The most frequent trigrams are plus stop aire (0.39%), huit stop dxpt (0.39%), and gege mont huit (0.39%). Since we have 125 symbols and only 24 letters in the French alphabet[13] it is clear that a plaintext letter can be encrypted by different symbols. This is a classical trick in Renaissance cryptography to avoid easy frequency analysis. We thus tried to find sets of symbols whose cumulative frequency would match the frequency of a given letter in *Moyen Français* (Fig. 2). For instance, given that dxpt has a frequency of 3.1%, we could have the set plus, stop, dxpt representing the letter 'e', with a cumulative frequency of $7.8 + 6.2 + 3.1 = 17.1$, which is near the frequency 17.2 of 'e'. Alas this led to a dead end, and likewise for bigrams and trigrams.

e	s	u/v	n	a	t	i	r
17.2	8.1	8.1	7.4	7.2	7.2	6.6	6.2
o	l	c	d	m	p	q	g
5.7	5.5	3.3	3.3	2.9	2.6	1.6	1.4

Figure 2: Frequency of letters in *Moyen Français* (in percent). These statistics come from an analysis of Rabelais' novel, *Pantagruel*, published in 1532.

3.2 Looking for words and patterns

We then searched for repetitions: the same sequence of symbols appearing at least twice in different parts of the ciphertext. We found such a repetition of 11 consecutive symbols (vset_s huit stop uhuh bebe zero_p mont aire aine huit stop), another one of 10 symbols, one of 8 symbols, one of 7 symbols and other shorter ones. With the hope that these repetitions of ciphertext symbols correspond to full words in the plaintext, we tried to make them match with words in *Moyen Français*. We efficiently restricted the search with the following remark. For the above repetition of 11 symbols, since the

[12] We used our names here.

[13] The characters 'i' and 'j' are the same, as for 'u' and 'v'.

frequency of say huit is 8.3% in the ciphertext, it may be an 'e' , 's', 'u/v', or 'n' in the plaintext according to Fig. 2 (a small margin of error is allowed, but, for instance if it corresponds to 'c', then the frequency of 'c' exceeds the expected 3.3%). At one point we thought that the 8-symbol repetition could correspond to *royaumes* (Kingdoms in English) and the 7-symbol one to *écuries* (stables in English), but this promising idea also led to a dead end. Yet, looking at our repetitions and thinking they were likely to be words, we noted that plus was very often at the end of words that existed also without it. We concluded that plus was likely to be a symbol for the letter 's'.

Not only did we search for exact repetitions but we kept in mind that one letter in the plaintext probably had several symbols to encipher it. For this reason we looked at near repetitions, that are repetitions of sequences of symbols that are exactly equal except for one inner position where they are allowed to have different symbols. For instance, we found the sequence of 10 symbols (ecro_s, ofof, huit_a, uhuh, plus, aire, cero, wewe, mont, plus) and later (ecro_s, ofof, huit_a, uhuh, plus, aire, cero, wewe, ptpt, plus). We conclude that mont and ptpt were likely to represent the same letter.

Moreover we noted that the symbol zede_p was always followed by the same symbol, namely gamm. We thought that zede_p could encode 'q' (their frequencies are 0.6% and 1.6%) and gamm could encode 'u' (frequencies 2.3% and 6.5%), since in French the letter 'q' is almost always followed by 'u'. A similar search for the letter 'x' was indecisive.

Another interesting idea was to try to split the 120 symbols between vowels and consonants. With 125 symbols, there can theoretically be 2^{125} possible partitions between vowels and consonants. However, assuming a word has at most 3 consecutive vowels (as in *oiseau*) and 3 consecutive consonants (as in *prendre*), it is possible to restrict the number of possible partitions. In the 11-symbol repetition above, assuming huit, stop, and uhuh are vowels, the next symbol bebe is necessarily a consonant. If we only consider the 14 most frequent symbols, yielding $2^{14} = 16384$ subpartitions, we find only two possible partitions of the full 125 symbols. Unfortunately, this also led to a dead end.

At the end we had several hypotheses that appeared to be right. Basic statistics led us very quickly to decide that no symbol (or even pair or triplet of symbols) was there to represent a space, which was correct. Basic statistics again gave possible values for the most common symbols, for instance we thought that huit was either 'e', 'u', 's' or 'n' (which was correct, huit is an 'n'). Looking at words told us that plus encoded 's'; and nearly repetitions combined to statistics led us to conclude that diff, zigv, zigo were the same letter and encoded one of 'e', 's', 'u', 'n', 'a', 't', 'i', 'r', 'o', or 'l' (which was correct, they are 'u'). Similarly we thought that alph and ccat encoded the same letter (which was correct, they both encode 'i'), and that ptpt and mont encoded the same letter (which was correct, they are both 'e').

Other hypotheses were wrong and led to a dead end. As we will see later, the main trick of Saint-Mauris' cipher consists in hiding vowels, and for this reason our guesses concerning vowels were hazardous, while it would have worked for other ciphers from that time which equally encrypted vowels and consonants. For instance, we thought that stop, bebe, dxpt and ptpt encoded a vowel, which was partially wrong, since the first two are respectively 't' and 'r' but the last two are respectively 'a' and 'e'. Similarly, hidden vowels and almost systematic bigrams were the reason why we were misled about zede_p, thinking is was a 'q' instead of 'qu'. Finally we looked for repetitions of several symbols with an extra symbol interspersed in order to identify nulls, but this was unsuccessful because nulls were not frequent enough in the ciphertext, and we were not aware of it with a single document. For instance we thought that aire might be a null but this was wrong.

3.3 Increasing the amount of data

We were puzzled with several unexplained observations: why did families of symbols that were graphically rotations to each other exhibit similar behaviour? How could we see repetitions of 11 symbols when the writer surely had two or more choices for each letter to be encrypted? At this point, the study of other encrypted letters from and/or to Saint-Mauris was needed to corroborate hypotheses. For practical convenience (the letters were digitized), the choice was made to work on the letters which were preserved in Besançon (BM

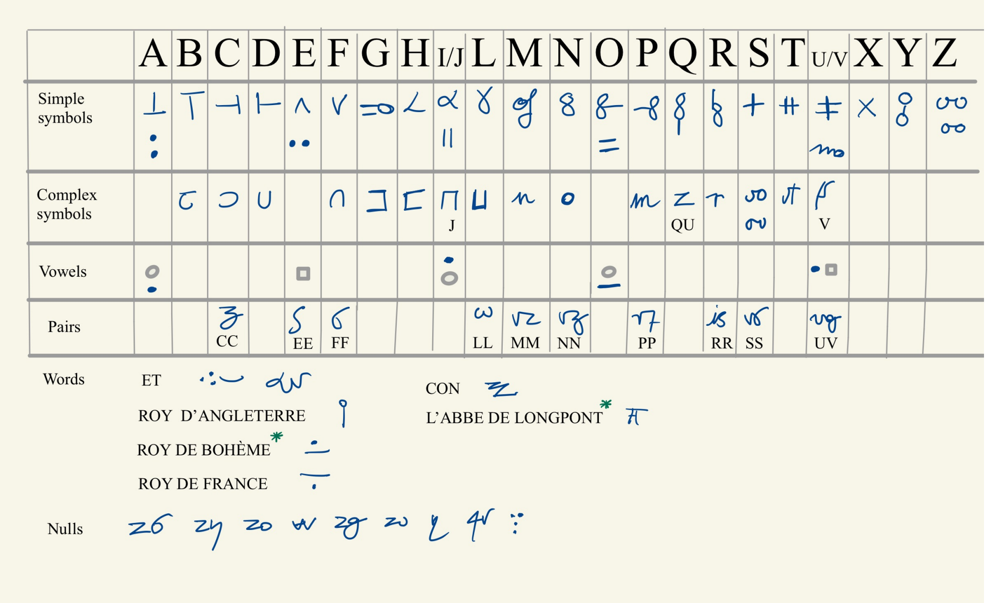

A B C D E F G H I/J L M N O P Q R S T U/V X Y Z
Simple symbols
Complex symbols
J
QU
V
Vowels
Pairs
CC
EE
FF
LL
MM
NN
PP
RR
SS
UV
Words
ET
CON
ROY D'ANGLETERRE
L'ABBE DE LONGPONT*
ROY DE BOHÈME*
ROY DE FRANCE
Nulls

Figure 3: The reconstructed cipher key. Some symbols in the nomenclator can only be guessed from the historical context: they are indicated by a green asterisk in the table.

Besançon, Granvelle 70). Some of them, especially the ones written by Charles and Granvelle (even though two years earlier) were encrypted with almost the same cipher and deciphered in the margin. This was sufficient to start the reconstruction of the cipher key (Fig. 3) and decipher the main part of the letter.

Saint-Mauris' cipher perfectly matches the Renaissance cryptographic practices, especially for European diplomacies. It relies on homophonic substitution and a nomenclator. As for every homophonic substitution, each plaintext letter can be represented by one (consonant) or two (vowel) ciphertext symbols. However, it goes further. Each consonant, if followed by a vowel, can also be encrypted by an extra complex symbol. In this case, the complex symbol is associated with a diacritical mark: dot at the bottom for 'a'; dot on top for 'i'; hyphen at the bottom for 'o'; dot on the left for 'u'. If there is no diacritical mark, it means that the symbol should be deciphered as consonant followed by 'e'. In addition, a ciphertext symbol exists for each repeated consonant (for instance a '3' for 'cc').

The letters in Besançon helped a lot for the value of the usual ciphertext symbols, much less for the nomenclator. Some ciphertext symbols remained a mystery. Four symbols did not appear in the letters in Besançon, but were crucial for understanding the letter in Nancy[14]. Surprisingly, the reconstruction of the nomenclator was quite easy and questions the complementary security that it is supposed to grant. The context of the letter, as well as the similarity between two symbols which encrypted kings, enabled us to identify two kings (in addition to the French King who was several times mentioned): the English King was associated to a recent death, and the Bohemian King to the Empire and to Charles' family. The last symbol, which encrypted Gabriel de Guzman, abbot of Longpont, was harder to uncover. Its decryption was made possible by Saint-Mauris' letter on 6th February, in which he mentioned his negotiation[15]. For this case, it would save little in case of an interception but reminds the main purpose of encryption: delaying the reading of the letter (if it was intercepted) and not fully preventing it.

[14] Three historians or cryptographers had previously reconstructed the key but not the nomenclator (Stix, 1934-1936; Tomokiyo, 2022) or they have not made it accessible (Potter, 2013). As we primarily worked on the 22nd February letter and used other letters only to pursue the global understanding of the core key, the nomenclator in this paper is incomplete and presents only the part of it which is used in the letter in Nancy.

[15] We have been able to consult the copy of this letter (AN, K1485-1488) only. The original encrypted letter, which we have not yet identified, and/or the minute of the 22nd February letter would confirm the attribution of this symbol to Longpont.

3.4 A very structured key.

Deciphering[16] the letter revealed both some patterns in the creation of the key itself and specific rules to use it. This structure is double-edged for the cryptanalyst. On the one hand, hidden vowels make usual statistics and methods fail, but on the other hand, any attack becomes easier as soon as the adversary is familiar with the Imperial cryptographic patterns. Although Charles' ciphers are nowadays little known, contemporary enemy cryptographers knew much better their common patterns and, several years later, Philip II himself acknowledged their low security and recurring decryptions.

Hidden vowels. Saint-Mauris' key uses a clever trick that explains both the failure of our first hypotheses and the unexplained observations we made: when they are following a consonant and form thus a bigram, the vowels are somehow hidden as diacritics. To encrypt a message the rules are the following. If you have to write a consonant followed by a vowel then use the complex symbol for the consonant and add around it the corresponding hyphen or dot for the following vowel. If you write a consonant not followed by a vowel or a vowel not following a consonant, just use one of its simple symbols. Always use the corresponding symbol for a pair of identical letters, and often the nomenclator if it exists. Nulls are not very frequent, except to hide important words and names, at least in this letter. Because of these rules, bigrams always consisted of a pair consonant-vowel but no symbol existed for (even frequent) bigrams of the form vowel-consonant (as 'un' or 'en' in French).

Rotation of symbols. With the reconstructed key in hand, we see a startling structure that betrays how the table was created. Symbols have been assigned in alphabetical order, and rotations were done to create new symbols, without mixing up these symbols. For example, the simple symbols for 'a', 'b', 'c', and 'd' are identical up to rotation, as are those for 'e' and 'f'; 'i' and 'l'; 't' and 'u' ; 'y' and 'z'. The symbols for 'n', 'o', 'p', 'q' and 'r' form another family. The observations we made about similar behaviours (':' and '..'), ('=' and '||') or the complex symbol family representing 'b', 'c', 'd', and 'f' are well explained by this structure. That tempers the cryptographic abilities of those who conceived the ciphers. Certainly the global patterns (homophonic substitution, vowel indicators, etc) were suggested and designed by cryptographers. The daily conception of ciphers however was the work of a secretary who was less concerned by the strength of the cipher (finding various symbols without any consistency between them) than by the need to quickly conceive multiple ciphers. The pattern relies here on rotation (as in some Hungarian ciphers (Lang, 2018)) as it relies, in some other ciphers, on alphabetical or numerical order. Indeed, Saint-Mauris' cipher presents also a numerical pattern for pairs ('5' for 'ee', '6' for 'ff' and so on). That truly questions its strength.

3.5 Merging keys

Finally, we compared the key to previous reconstructions we were able to access. The first one (Stix, 1934-1936) was conducted by Franz Stix within a general study of Charles' cryptographic practices from the Vienna archives (OeStA-HHStA, Fr 10-16). The second one (Tomokiyo, 2022) relied on a single letter which Satoshi Tomokiyo found in Madrid (BNE, 7913). Results were convergent but the comparison allowed us to understand better some aspects of the ciphering process and question once again the security that the cipher granted to the letter. Stix reproduced the main part of the key: symbols for letters, bigrams (all the bigrams are developed and not presented, as in our reconstructed key, as complex symbols and vowel indicators), and repeated consonants[17] but neither the null symbols, the ciphertext symbol for 'com/con', nor the nomenclator. On the other hand, the first key reconstructed by S. Tomokiyo presented only a subset of simple and complex symbols. The null symbols were also reconstructed but not the ciphertext symbols for 'et' and 'com/con' nor the nomenclator. All the identified symbols in the three tables were very similar, even though the writing frequently differed. In fact, it was only when comparing with the other reconstructions that we were able to con-

[16] In order to facilitate the understanding of the encryption processes, we have separated simple symbols, complex symbols and vowel indicators for the presentation of the reconstructed cipher key. Nevertheless, according to the usual presentation of Renaissance ciphers, one can assume that the plaintext bigrams were developed (ba, be, bi, bo, bu, ca, ce, ci, ...). The key might thus be structured in 3 parts: the simple symbols (with the nulls), the bigrams and the nomenclator.

[17] For the repeated consonants, the key reconstructed by Stix revealed the symbol for 'pp' and 'rr' for instance, but not for 'ee'.

firm that the repeated consonants symbols were numbers, even in increasing order. For instance ’ll’, ’mm’ and ’nn’ are encrypted with 10, 12 and 13 while ’rr’ and ’ss’ correspond to 15 and 16. Finally, at first sight the nulls that Tomokiyo identified are quite different from ours, but most of them are digit numbers too. There could even be a rule that any number larger than or equal to 20 is a null. For instance the first three symbols of Fig. 3 might be particular spelling for 26,24 and 20. In our case, another null symbol is formed from 4 dots. This is consistent with what we found in Besançon, where more than one dot around a symbol automatically cancels it out.

The comparison with the works of Stix and Tomokiyo highlighted differences and developments in some symbols, such as the complex symbol for ’s’. In the reconstruction of F. Stix it looked like a letter ’s’ with a small circle attached on the top right part. In S. Tomokiyo’s table, this structure was still visible, but one might not interpret it this way if not aware of the other table. In this letter, however, the complex symbol for ’s’ sometimes became almost flat and was hard to distinguish with the symbol for ’z’. These variants of ’s’ are shown in Fig. 4, and we chose to let the ambiguity between ’s’ and ’z’ be visible in Fig. 3. This example however highlights one of the issues of deciphering early modern letters: characters can be written in different ways even when they are the same (bad writing, different secretaries, cipher evolution and so on).

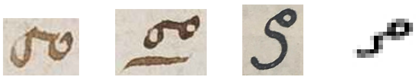


Figure 4: Variants of the complex symbol for ’s’. From left to right: Symbols for ’se’ and ’so’ in the 22nd February letter; symbol for ’se’ in Stix’ key; symbol for ’se’ in Tomokiyo’s key.

Finally, there are several ways to interpret the various writing styles for the ciphertext symbols that occurred in the letters, and which have consequently been passed on to the three reconstructed keys. It could be that the writer was requested to cipher quickly or because he was not mastering the process well. In both cases, that underlines the difficulties of manual ciphering. In fact, in addition to the bad writing of some symbols, many ciphering errors can be pointed out in the letter. They never prevent the complete understanding but are comparatively more frequent than in the other correspondences we have worked on. Further research could help to determine whether these ciphering errors are specific to that letter or if they were common in Charles’ encrypted correspondence, but also to define the type of errors (writings, cross-contamination from other keys and so on). This investigation as well as the reconstruction of the whole nomenclator should help to question the quality of the Imperial ciphers as well as the ciphering mastering of its secretaries.

4 Results

The deciphering enabled us to uncover the content of the letter (mainly about Charles’ concerns towards Francis) and also Charles’ encryption methods in the mid-16th century.

4.1 Content of the letter

In a first part, Charles reaffirmed his concern about Francis’ intentions while he was gathering his military forces in the Empire against the Schmalkaldic League. These concerns were earlier made public by Saint-Mauris at the French Court while Charles had several times expressed, during audiences with Jacques Mesnage, the French ambassador at the Imperial Court, his good will towards Francis and had exhorted him to peace but without clearing away the French doubts[18]. Saint-Mauris was thus encouraged not to openly express Charles’ mistrust or relaunch the negotiations (probably about Hesdin and/or ”demilitarisation” of Northern Italy). On the contrary, Charles ordered him to discover the French intentions towards England following Henry VIII’s death. Saint-Mauris however seemed to remain in the dark about the French intentions towards both Boulogne and a general alliance with England, the Schmalkaldic League, and even Venice (Potter, 2013). As this letter shows, Saint-Mauris was still fishing. This concern about maintaining peace is finally expressed one more time at the end of the first page: Charles immediately accepted the proposal of Claude d’Annebault, Francis’ main counsellor and previous governor of Piedmont, to keep running the cooperation and mutual surveillance of the Milano-Piedmontese border. By countering the French diplomatic and military maneuvers, the

[18] Various letters from Mesnage to Francis, written on 16th January and 20th January 1547 (Ribier, 1666, p. 591-593 and p. 595-600), and on 8th February 1547 (BnF, fr. 17889, fol. 241-242) testified those speeches.

letter reveals part of Charles' foreign policy. It is hardly surprising that his concerns were encrypted while his public demonstration of goodwill by accepting Annebault's proposal was written in cleartext.

In a second part, Charles reported a disturbing rumor: Piero Strozzi, who belonged to an Italian banker family and served Francis both with his financial and military abilities, was planning to assassinate him. Strozzi was indeed sent to the Schmalkaldic League to bring them the French financial subsidies and was suspected of taking advantage of his journey for much more dangerous matters. However, Charles acknowledged that the French King would have refused to support such an assassination project. The fear may originate from the dubious status of Strozzi's missions in the Empire. They were mostly managed directly in the Empire by Jean Sturm, and Strozzi made frequent journeys back and forth, including in Italy, in order to elaborate the French loans (Potter, 1977; Pariset, 1981). When replying on 6th March 1547 (OeStA-HHStA, Fr 10-16; Potter, 2013), Saint-Mauris confirmed that it was only a rumor.

The last part of the letter outlines the state of the Schmalkaldic war. Charles mentioned his upcoming journey to Frankfurt in order to confer with his brother Ferdinand, King of Bohemia and King of Romans, about the operation led by Maurice, duke of Saxony, against John Frederick, elector of Saxony and one of the leaders of the Schmalkaldic League. Nevertheless, if the military situation was improving for Charles, there arose in Prague a revolt which was immediately reported by Jacques Mesnage[19]. In response, Charles encouraged Saint-Mauris to minimize the scale of the revolt as well as the night flight of Ferdinand of Tyrol, Charles' nephew, by transforming it into a simple hasty departure to join his father, Ferdinand, King of Bohemia, and the fight against the Elector of Saxony. In fact, on 22nd February 1547, the revolt was not yet over: Ferdinand, King of Bohemia, was still negotiating with the States of Bohemia, and also with those of Moravia and Silesia, who had joined the first ones.

4.2 Cryptography under Charles V's reign

Imperial cryptographic practices have suffered from a bad reputation, because of both insufficient historical knowledge [20] and the comparison with Philip II's ciphers. Nevertheless, Saint-Mauris' cipher isn't less complex than other European ciphers at the same time. French diplomacy for example already used in the 1530's two or three ciphertext symbols for each plaintext letter (Desenclos, 2021). In the 1540's, Charles' brother Ferdinand, as King of Hungary, used ciphers with two or three ciphertext symbols with his ambassadors (Lang, 2018). In Saint-Mauris' key, the complex symbols act as a second set for plaintext letters as do the vowel indicators. It thus presents two ciphertext symbols for consonants and three for vowels. Moreover, as for many other European ciphers, Saint-Mauris' cipher offers complementary encryption processes: null ciphertext symbols, ciphertext characters for each repeated plaintext consonant, nomenclator.

The diacritical marks for vowels seem to be specific to Imperial ciphers, to the authors' current knowledge. Those vowel indicators could offer the encrypted letter extra strength. As our deciphering attempts show, vowel indicators prevent (or at least slow) any cryptanalysis by frequency analysis. But Charles' encrypted letters were regularly deciphered by enemies who discovered this main pattern of his ciphers. On this basis, Charles' ciphers lost their strength: unlike bigrams (using a different ciphertext symbol each time), using the same diacritical mark for each vowel again made possible frequency analysis. On that perspective, Saint-Mauris' cipher could be considered as less strong than other European ciphers, but it reminds us also of the value of the Imperial ciphers in the history of cryptography, especially for the understanding of Spanish cryptography under the reign of Philip II.

This use both of bigrams and diacritical marks indeed was not new. Since 1527, it can be observed within several ciphers such as the one used between Iñigo Mendoza, ambassador at the French Court, and Charles V. Since then, those diacrit-

[19] His diplomatic papers which include drafts of letters to the French King and letters sent to him by the latter, can be found at the French national library (Paris) (BnF, 17889-17890). On the contrary, the original letters, written by Mesnage to the French King haven't yet been identified; some of them have been edited but not the one to which Saint-Mauris referred (Ribier, 1666).

[20] The correspondences both from Charles V and from Granvelle have been broadly studied and, sometimes, edited. But the cryptographic practices are only quickly mentioned: their existence are acknowledged, sometimes the kind of cryptographic symbols described (see for instance (Berthomeu Masia, 2006)) but the keys are rarely transcribed or even studied.

ical marks were regularly used by Imperial ciphers (Tomokiyo, 2019). They can be observed until 1555 (Tomokiyo, 2022). This vowel encryption process can be considered at the ancestor of encrypted bigrams and trigrams under Philip II's reign. Indeed, Spanish ciphers after 1556 still used diacritical marks for vowels in the exact same way (the same diacritical mark for each vowel whatever the consonant is) (Devos, 1950), but they progressively moved from vowel indicators to proper bigrams and trigrams (two consonants and one vowel such as 'cha', 'che', etc): each bigram and trigram was now encrypted by a different ciphertext character. Certainly, they often matched to increasing numbers (e.g. 10 for "ba", 11 for "be", ...) but vowels could no longer be spotted easily.

5 Conclusion and perspectives

Deciphering this letter may have taught little about the relationship between Charles and Francis. As a large part of Saint-Mauris' correspondence had already been studied, the uncovered content only confirms current historical knowledge. The main value of this work lies in understanding the cryptographical approach of the letter. When deciphering, how to deal with an isolated letter, encryption patterns which aren't well known or documented, and with inconsistent writings? This work led us to question both the ciphering and deciphering process. By working only on one specimen, then by reinserting it in a larger sample, and finally by merging with other similar keys, cryptographic patterns have been highlighted. The decryption of this letter nevertheless is the beginning and not the end of a general study of Charles' cryptographical practices. In the future, thanks to the corpus enlargement (Vienna and Bruxelles mainly), the authors aim to investigate both the cryptographic adaptations to Charles's diplomacy network (Saint-Mauris wrote with the same cipher to Charles, Granvelle, Maria of Austria, Infante Philip and Los Cobos, but they may have adaptations, especially in the nulls and nomenclator) and the exact process of manual ciphering (misuse of complex symbols, ciphering errors, bad writings, nomenclator evolutions and so on). Thereby, the authors hope to consolidate the cipher key from Fig. 3 and contribute to a better knowledge of Renaissance cryptographic practices.

Acknowledgements. The authors thank Sophie Toulouze, Anne Canteaut, Olivier Canteaut, and Lana Martysheva who were involved in making this collaboration possible, the three anonymous referees, and Richard Brent who helped to improve our written English.

References

Archives générales du royaume. Audience 420. Correspondence between the Imperial agents at the French Court and the Holy Roman Empire, 1535-1563.

Archives générales du royaume. Audience 1672/2/E. Correspondence between the Imperial agents in France and the Low Countries, 16th century.

Archives générales du royaume. Manuscrits divers 95-96. Correspondence from Charles V and Maria of Hongria to Jean de Saint-Mauris, copies.

Archives nationales, "Fonds de Simancas". K1485 to K1488. Correspondence between the Imperial agents in France and Spain, 1544-1548.

Rainer Babel. 2013. *La France et l'Allemagne à l'époque de la monarchie universelle des Habsbourg, 1500-1648*. Presses Universitaires du Septentrion, Villeneuve d'Ascq.

Maria José Bertomeu Masia. 2006. *Cartas de un espia de Carlos V. La correspondancia de Jeronima Bucchia con Antonio Perrenot de Granvela*. M. Rieger, Munich.

Bibliothèque municipale de Besançon. Granvelle 40. fol. 139, Letter from Jean de Saint-Mauris to Charles V, 6th February 1547. https://memoirevive.besancon.fr/ark:/48565/th72lvsb095f/6dc4e84d-9393-4e5c-9172-e1360a268aaf.

Bibliothèque municipale de Besançon. Granvelle 70. Lettres et papiers de l'ambassade de Jean de Saint-Mauris, 1544-1576. https://memoirevive.besancon.fr/ark:/48565/t43hg0sk92jl/72919b8a-dacd-44a5-a220-a6987a0378d9.

Bibliothèque Stanislas de Nancy. Letter from Charles Quint to Jean de Saint-Mauris, 22nd February 1547. https://galeries.limedia.fr/ark:/31124/dct0sbwx8vmhspk0.

Biblioteca nacional de España. MSS/7913/127. Letter from Jean de Saint Mauris to Granvelle, 5th July 1548.

Bibliothèque nationale de France. Manuscrits français 17889 to 17890. Ambassade de Jacques Mesnage auprès de Charles Quint, 1544-1546.

Auguste Cassan. 1878. La mort de François I^{er} et l'avènement de Henri II d'après les dépêches secrètes de l'ambassadeur impérial Jean de Saint-Mauris. *Mémoires de la société d'émulation du Doubs*, pages 422–454.

Camille Desenclos. 2021. Écrire le secret quotidien. Pratiques de la cryptographie au sein de la diplomatie française (XVI^{e}-premier XVII^{e} siècle). In G. Braun and S. Lachenicht, editors, *Spies, espionnage and secret diplomacy in the early modern period*, pages 85–103. Kohlhammer.

Jean-Pierre Devos. 1950. *Les chiffres de Philippe II (1555-1598) et du despacho universal durant le XVII^{e} siècle*. Palais des Académies, Bruxelles.

August von Druffel. 1878. *Briefe und Acten zur Geschichte des sechszehnten Jahrhunderts*. M. Rieger, Munich.

Nils Kopal and Michelle Waldispühl. 2022. Deciphering three diplomatic letters sent by Maximilian II in 1575. *Cryptologia*, 46/2:103/127.

Benedek Lang. 2018. *Real Life Cryptology. Ciphers and Secrets in Early Modern Hungary*. Amsterdam University Press B.V., Amsterdam.

François Nawrocki. 2015. *L'amiral Claude d'Annebault, conseiller favori de François I^{er}*. Classiques Garnier, Paris.

Österreichisches Staatsarchiv / Haus Hof und Staatsarchiv. Frankreich, Berichte 10 to 16. Imperial diplomatic correspondence to France, 1542-1548.

Jean-Daniel Pariset. 1981. *Les relations entre la France et l'Allemagne au milieu du XVI^{e} siècle*. Librairie Istra, Strasbourg.

Jean-Daniel Pariset. 1982. La France et les princes allemands. Documents et commentaires (1545-1557). *Francia*, 10:229–301.

David Potter. 1977. Foreign Policy in the Age of the Reformation: French Involvement in the Schmalkaldic War, 1544-1547. *The Historical Journal*, 20/3:525/544.

David Potter. 2011. *Henry VIII and Francis I. The Final Conflict, 1540-1547*. Brill, Leiden.

David Potter. 2013. La fin du règne de François I^{er} et l'avènement d'Henri II d'après les dépêches de Jean de Saint-Mauris. `https://cour-de-france.fr/article2749.html`.

Guillaume Ribier. 1666. *Lettres et mémoires d'Estat des roys, princes, ambassadeurs et autres ministres sous les regnes de François premier, Henry III et François II, tome premier*. François Clouzier, Paris.

Franz Stix. 1934/1936. Die Geheimschriftenschlüssel der Kabinettskanzlei des Kaisers. *Nachrichten aus der Mittleren und Neueren Geschichte*, 1/2:207–226/61–70.

Satoshi Tomokiyo. 2019. Tracing the origin of vowel indicators in Spanish ciphers. `http://cryptiana.web.fc2.com/code/vowel.htm`, retrieved 2023-04-19.

Satoshi Tomokiyo. 2022. Ciphers during the reign of emperor Charles V. `http://cryptiana.web.fc2.com/code/spanish2.htm#SEC14B`, retrieved 2023-04-19.

Appendix: The Decrypted Letter

(We put the decrypted part in italics.)

L'empereur et roy

Chier et feal

Nous avons receu voz deux lettres des XXVIe du passé et VIe du present et par icelles entendu bien amplement tous occourans en ce coustel là et mesmes la responce que vous a fait le roy sur ce que luy avyons fait remonstrer par vous *par vous [sic] et puisque luy ny ses ministres ne se sont extendus davantaige quant à la plus estroicte amyté et moyens d'icelle, sinon qu'il seroit bon remectre la negociation à l'abbé de Longpont, il sera bien laisser la chose ainsi sans en faire plus de mention jusques l'on voye s'ilz retourneront à en parler et en feront plus d'instance, et proposeront aucuns moyens où l'on puisse prendre quelque fondement dont nous advertirez. Et nostre dicte seur vous tenant tousjours cependant ès mesmes termes qu'avez jusques à maintenant sans en riens vous eslargir davantaige en sunvant [sic] ce que vous avons tousjours escript. Et sera bien que nous advertissez de ce qu'aurez pu assentir de leu[r] intention depuii qu'ils auront sceu le trespas [du] roy d'Angleterre et z'ilz n'etendent rien s[e] mouvoir en ce coustel là et si soubz ceste couleur ils se font plus grant amas de gens ensemble de toutes aultres particularitez.* Et quant à ce que l'admiral vous a dit que pour entretenir bonne voisinance et eviter tous scrupules, il seroit bon que l'on observa du coustel de Piedmont ce que faisoient le feu marquis del Gasto et luy d'advertir l'ung l'aultre quant aucunes garnisons se augmentoient ou changeoient d'ung lieu à aultre, vous luy pourrez dire que le trouvons bon et ferons escrire au sieur Don Fernande que à l'advenir il en use ainsi et que de leur coustel ilz facent de semblable à leur gouverneur audit Piedmont. Ledit Don Fernande nous a envoyé le memoire cy joinct dont pourrez parler comme aurrez l'opportunité.

Au surplus l'on nous a adverty que estant dernierement le roy ou coustel de Bresse, aulcuns gentilz hommes ytaliens suyvans le sieur Oracio, eulx monstrans affectionnez à nous, auroyent dit qu'ilz se covoyent certainement que Pierre Strossy en partant dernierement de France et lors qu'il vint au camp des rebelles dit entre aultres choses au roy que s'il vouloit qu'il entreprendroit de nous tuer et qu'il n'en demandoit aulcune recompense ny se soucioyt d'estre apres prins, car il estoit content de mourir moyennant que aussi [nous] mourissions, et que le dit sieur roy luy auroit respondu qu'il ne s'estoit jamais meslé de telles praticques, et qu'encores ne le vouldroit y faire et que ledit Strossy fit ce qu'il vouldroit. Lequel auroit depuis encores dit aillieurs qu'il s'en iroit audit camp des rebelles et trouveroit moyen d'entrer au nostre soubz quelque couleur que ce fut et mectroit sa volonté à execution quoiqu'il en deust advenir. Et pour ce que vouldrions bien que cecy se puist en aucune manière veriffier pour avoir occasion de faire apprehender ledit Strossy et nous en pouvoir justifier en ce coustel là, sera bien que regardez tous moyen possibles pour si faire se peult scavoir si ledit Strossy auroit tenu audit roy les susdictz propoz ou encores aillieurs. Et cecy vous recommandons nous affectueuzement.

En oultre nous sumes deliberé partir d'icy dans cincq ou six jours et tirer contre Francfort pour estre là à propos de tous affaires et pouvoir tinir meillieur correspondence avec le roy de Boheme en l'emprinse de Saxe de laquelle somes actendant nouvelles du succes. Et pour ce que l'ambassadeur Mesnaige auroit par adventure escript par delà et fait grant cas de l'emotion de Praghe. Et aussi que Monsieur l'Archiduc nostre nepveu s'estoit une nuict party secretement vous advisons que quant à ladicte emotion elle est cessee et a esté seulement une assemblee de peuple sans qu'il en soit ensuy aultre chose. Et quant à nostre dit nepveu ayant entendu que son pere delaberoit soy trouver en la dicte emprise contre le jadis electeur, et doubtant que ne luy eussions voulsu permectre d'y aller s'estoit desrobé pour soy y trouver mais il rev[i]ent le mesme jour et ainsi en pourrez respondre si vous en es[t] parlé. À tous chiers et feal Dieu vous ait en sa saincte garde. De Ulme le XXIIe de fevrier 1546.

Charles
Bave

On the Combination of Cryptography and Steganography in 17th Century Germany

Eveline Szarka
University of Heidelberg
Grabengasse 1
69117 Heidelberg, Germany
eveline.szarka@uni-heidelberg.de

Abstract

Assessing and averting possible interception lies at the heart of cryptology. In handbooks printed in 17th century Germany, the authors suggested combining steganography and cryptography to increase information security. This article discusses several techniques for concealing ciphers and demonstrates that the authors of instructional literature had to consider complex interplaying factors when it came to combining cryptography and steganography. Overall, these examples show that the cryptological literature of the 17th century mirrors an increased discussion about the visibility of ciphers.

1 Introduction

The history of early modern cryptology is first and foremost a history of alphabetical and numerical encryption techniques. Historians have been primarily interested in the invention and evolution of polyalphabeticity, which, according to scholarship, peaked in the 15th and 16th centuries (Strasser, 2007, p. 297, 321, Kahn, 1967, p. 154). For this reason, the cryptological literature of the 17th century, especially those printed in Germany, received relatively little scholarly attention.[1] However, the authors of handbooks on secret communication showcase the skillful use of diverse resources to develop ever more sophisticated techniques. Applying polyalphabeticity was not the only way to increase information security.

Steganography is another blind spot in the history of secret communication. In contrast to cryptography (text encryption), steganographic techniques hide messages entirely so that they escape the attention of unauthorized persons. In the early modern era, this form of secret communication was more diverse and important than is assumed today.

To achieve an even higher level of information security, early modern scholars also suggested combining cryptography and steganography. The aim of this article is to offer an insight into multiple historical techniques concerned with the concealment of a ciphertext. To create an adequate basis of understanding, I will briefly discuss early modern steganography and its relevance for the history of cryptology. It will be argued that, when it came to the combination of cryptography and steganography, scholars had to consider multiple interplaying factors such as the knowledge of interceptors, the naturalness of the steganotext, the choice of the steganotext, the complexity of the cipher, as well as the encryption process on the receiver side.

On a broader level it will be shown that these techniques reflect a broader historical change in cryptology that began around 1600. Although alphabetical and numerical substitutions remained important, 17th century sources mirror an increased discussion about the visibility of ciphers.

2 Early Modern Steganography

The field of steganography (from the Greek word στεγανός, 'hidden' and γραφία, 'writing') dates back to antiquity. Steganography is a form of secret communication that seeks to hide messages in innocuous texts, objects, art works etc. There are two forms of steganography:

[1] Gerhard F. Strasser examines 17th century handbooks published in Germany, although with a focus on universal languages schemes, see Strasser (1988). For England see Ellison (2017).

Linguistic steganography is the term used to describe methods by which a text is hidden in a set of data. Writing with invisible inks or hiding slips of information in objects, on the other hand, are examples of technical steganography. This article discusses linguistic steganography.

Semagrams were a popular means of hiding information in early modern Europe. Here, the secret message is embedded into an inconspicuous message (*steganotext*). The semagram relies on visual cues such as tiny ink dots placed above certain letters or a slightly different appearance of a character to denote the relevant letter of a secret message. The advantage of this technique is that the sender does not have to compose a text, but can also send, for example, journals or books to the receiver to convey clandestine intelligence. Due to the use of visual cues, however, detection is still possible.

Null ciphers work similarly, but the plaintext is inserted into a larger text body based on a rule. Despite their name, null ciphers are not concerned with encryption, but the significant text is surrounded by insignificant data (*nulls*). An example would be to only read the first (*acrostic*) or last (*telestich*) letters of each word.[2] Depending on the complexity of the rule, null ciphers are a safe way to convey clandestine messages. In contrast to semagrams, the sender would have to compose a text that might sound constructed.

Since steganographic techniques ideally do not arouse suspicion, they are also prone to escape the historians' gaze. This is probably the reason why there is hardly any historical research on this topic.[3] Invisible inks have been found in early modern letters (Britland, 2018, p. 208, Rous, 2011, p. 250), but other forms are much harder to spot. For this reason, cryptological handbooks are – besides metatexts that attest for its application[4] – the only valuable source to gain insight into early modern steganography.

Many cryptological works of the 16th century mention forms of linguistic steganography. However, in the 17th century, steganographic techniques take up more space in cryptological handbooks; a trend that correlates with changing attitudes towards the visibility of ciphers. The Nuremberg professor Daniel Schwenter, military expert[5] and author of the first cryptological handbook in the German language, the *Steganologia & Steganographia nova* (1617)[6], writes (p. 250):

> "Hingegen aber ist es nicht rahtsam inn Brieffen mit verborgenen Charactern zu schreiben / in dem die sach voller argwohn / und man solche Brieff nit allein auffangen / sondern auch verstehen möchte […]."[7]

The polymath Georg Philipp Harsdörffer states in the *Delitiae Mathematicae et Physicae* (1653) that during sieges, letters containing indecipherable characters, numbers, or images will not pass the guards.[8] He also advises to hide messages altogether (p. 56). The risk of a death penalty for messengers who carried encrypted information (Harsdörffer, 1653, p. 56, Francisci, 1673, p. 176) was not limited to the 17th century, but also a relevant factor. Moreover, it is likely that rising literacy rates and the dissemination of cryptological knowledge increased the amount of people who were able to break ciphers.[9] It seems then, that the mastering of steganographic techniques became increasingly important in the 17th century.

Still, little is known about the value of steganography in cryptological literature, the authors' receptions, and improvements of existing methods as well as inventions after 1600. What factors did scholars like Schwenter assess to guarantee information security? What additional security measures did they apply? And what about the advantages and disadvantages

[2] For example, the *Steganographia* (c. 1499) penned by Johannes Trithemius' applies the null cipher to hide the secret behind the names of spirits.

[3] Despite having surveys on steganography covering the time from antiquity to modern times (Schmeh, 2009 or Macrakis, 2014), there is still a lack of in-depth studies on the early modern era.

[4] Strasser writes that the so-called *Cardano Grille*, a physical steganographic key, "was employed from the 16th well into the 18th centuries in the diplomatic correspondence of a number of countries." (Strasser, 2007, p. 291).

[5] Several parties asked Schwenter for advice during the Thirty Years' War. Mährle (2000), p. 375. Schwenter also published a work on war fortification, see Gärtner (1999), p. 244.

[6] Much ink has been spilled about the exact publication date of Schwenter's handbooks. However, it seems likely that the first edition with the *nova*-title was published in 1617, as on p. 50 Schwenter refers to a book that had been published two years earlier, namely Franz Kessler's *Secreta* (1615). The publication dates of the second and third editions remain unresolved.

[7] "However, it is not advisable to write with secret characters in letters / as the matter is full of suspicion / and one can not only intercept these letters / but also understand them."

[8] Post espionage in general and the extent of interception remains to be examined. In a survey on early modern Saxony, Rous counted eleven boxes with intercepted letters from the Thirty Years' War. Rous (2022), p. 380.

[9] According to the writer Erasmus Francisci everyone had become so clever, that it was almost impossible to come up with secure ciphers. Francisci (1673), p. 173.

when it came to decide on the one or the other method? Regardless of whether or not these techniques were applied, handbooks are a crucial source for the history of science and technology, as they demonstrate how changes in communication culture spurred the invention of new ways to protect intelligence.

Information security could, for example, be increased by the combination of cryptography and steganography. In simple terms, a plaintext is transformed into a ciphertext and hidden as / in a steganotext.

plaintext	ciphertext	steganotext
hello →	ALFFC →	**A**nd **L**inda **f**ed **f**eral **c**ats.

Table 1: Example of a concealed cipher, based on a simple monoalphabetic substitution and acrostic null cipher.

Concealed ciphers were known before 1600,[10] however, 17th century handbooks contain more variations and innovations thereof. The following examples demonstrate that hiding ciphertexts required the assessment of various interplaying factors.

3 Secret Alphabets

In addition to common alphabetical or numerical ciphers, handbooks often include secret alphabets consisting of alchemical, zodiac, or geometrical symbols. We know that in 17th century Germany symbols were used for encryption, although they were not supposed to be hidden.[11] Overall, the use of secret alphabets was not a secure method to conceal a message, as the cipher could be cracked easily just like other encryptions based on a monoalphabetic substitution.

Schwenter assesses several commonly known alphabets in the fifth book of the *Steganologia* (1617), such as Cabbalistic alphabets discussed by Cornelius Agrippa in the third volume of *De occulta philosophia* (1533), all of which Schwenter deems to be impractical as they are difficult and slow to write (p. 146). He also examines several geometric alphabets consisting of rectangles, triangles, or circles.

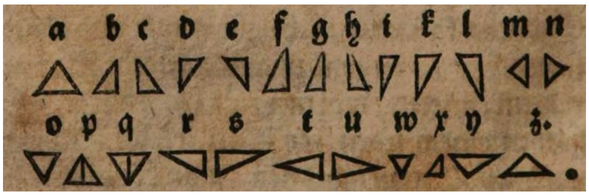

a b c d e f g h i k l m n
o p q r s t u w x y z.

Figure 1: Triangle Alphabet (Schwenter, 1617, p. 152). Bayerische Staatsbibliothek München, Res/Path. 801#Beibd.4, urn:nbn:de: bvb:12-bsb10926734-6.

Unfazed by the existing techniques, Schwenter introduces his own cunning alphabet (p. 154). As a professor of Oriental languages at the Altdorf Academy near Nuremberg, he took inspiration from the Hebrew, Syriac, and Arabic writing systems. According to the key, the symbols substitute consonants and dots signify vowels. As figure 2 shows, the dots are placed below or above the preceding consonants. The first symbol from the right is to be used as a proxy consonant in case a word starts with a vowel. Additionally, Schwenter proposes to change the writing direction.

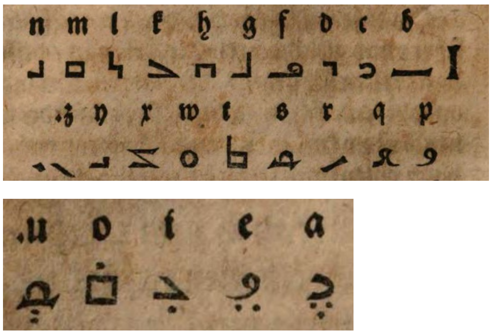

n m l k h g f d c b
z y x w t s r q p
u o i e a

Figure 2: Schwenter's Secret Alphabet (Schwenter, 1617, p. 154). Bayerische Staatsbibliothek München, Res/Path. 801#Beibd.4, urn:nbn:de:bvb:12-bsb10926734-6.

Schwenter then encrypts the sample text "Saliter / Schwefel und Weinstein angezündet / zerschmeltzen etliche Metall fast im augenblick."[12]

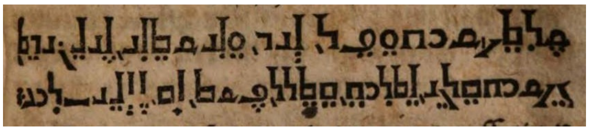


Figure 3: Example of Schwenter's Secret Alphabet (Schwenter, 1617, p. 155). Bayerische Staatsbibliothek München, Res/Path. 801#Beibd.4, urn:nbn:de:bvb:12-bsb10926734-6.

This alphabet is a simple monoalphabetic substitution, but the secret message disguises

[10] Trithemius' *Ava Maria Cipher* presented in his *Polygraphia* (1518) is a prime example on how to hide an enciphered text. The letters of the alphabet are encrypted by Latin nouns, verbs, and adjectives that should read like innocuous religious texts or prayers. See also a recent study by Paolo Bonavoglia (2020) on a technique used in late 16th century Italy, as well as Strasser (2007), p. 317.

[11] See for example Stützel (1963).

[12] "Salpeter / sulfur and wine scale ignited / immediately melt all kinds of metals."

itself as an innocuous text; using this method, interceptors would ideally assume a text written in a foreign language and therefore refrain from attempting to break the cipher. In this case, the ciphertext and the steganotext are identical.

It can be assumed that Schwenter proposed to encrypt only certain passages and not an entire letter. In that case, the relationship between the script and the surrounding text would have been of utmost importance. The letter would have had to refer to languages or scripts. The specific communication context should also have legitimized the exchange of information related to this topic.

The key to using secret alphabets was to find a middle ground between something that looked familiar, but not too familiar. In contrast to the triangle letters or other secret alphabets, this way of writing could be mistaken for a foreign script in use. It was still better to encrypt a text with Greek letters than with an invented alphabet, but in 17th century Germany, there was still the possibility that the correspondence could fall into the hands of scholars who were able to read Greek. The chance of someone being fluent in Oriental languages was smaller. Thus, the success of this technique depended on the interceptors' language skills.

To protect the message even in the event of detection, Schwenter adds another layer of security by suggesting a right-to-left-script as is usual in the Arabic, Syriac, or Hebrew writing systems. While *inversion* – the technical term for writing backwards – is usually not a safe way to disguise a message, it somewhat improves the safety of an enciphered text as cryptanalysis heavily depends on the recognition of linguistic patterns such as frequent diphthongs, pre- or suffixes, or parts of words. This process is already complicated by the fact that the vowels are substituted by diacritics. Schwenter proudly declares his invention to be the best, swiftest, and safest of all secret alphabets (p. 155).

4 Musical notes

Musical note ciphers were already used in the Middle Ages (see for example de Luca / Haines, 2018, or Code, 2022, p. 13). However, it is unclear if the ciphers were supposed to be visible. Musical note ciphers were particularly popular in 17th century handbooks.[13] Gustavus Selenus (Duke August II of Brunswick-Lüneburg) discusses several musical ciphers in his *Cryptomenytices* (1624), which he attributes to Count Friedrich of Oettingen-Wallerstein.[14]

These techniques are based on the *Polybius Square*. In its original form, the square, first introduced by the Greek historian Polybius in the second century BC, consists of five rows and five columns with spaces for up to 26 letters of the alphabet. Thus, all letters are represented by a two-digit number, depending on which cell they are located in. The first digit signifies the row, the second the column. Although by using the square the ciphertext is twice as long as the plaintext, from a combinatorial perspective it comes handy in situations when cryptographic knowledge is transferred from basic alphabetical substitutions to other symbolic representations or media. For example, Polybius applied the square to encrypt a message by operating five torches at two places (5^2 = 25 letters of the alphabet) (Polybius, 1925, p. 215ff.).

To simplify and speed up the encryption process, Selenus reduces the alphabet to 16 letters as is shown in table 2.

	1	2	3	4
1	a	b	c	d
2	e	f	g	h
3	i	l	m	n
4	o	r	s	u

Table 2: Selenus' version of the *Polybius Square* (Selenus, 1624, p. 321).

Now the sender draws five lines that serve as note lines as is common in musical sheets, although only four are to be used. These four lines correspond to the numbers of columns and rows (4^2 positions to signify 16 letters). The letter "g" as is shown in Selenus sample text "Gustavus" is therefore encrypted by the number 23 (second row and third column). The sender must therefore place the first note on the second line and the second note one on the third.

[13] For an overview of various musical ciphers of the 17th century see Code (2022).

[14] The Count penned a cryptological manuscript in 1601 that was possibly offered to emperor Rudolf II. Strasser (1997), col. 784. See also Strasser (1982), p. 86f.

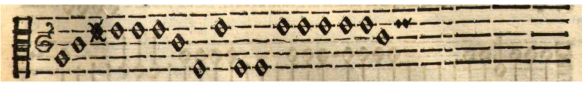


Figure 4: Selenus' example of a musical cipher (Selenus, 1624, p. 321). Staats- und Stadtbibliothek Augsburg, 2 H 336, urn:nbn:de:bvb:12-bsb11197933-2.

However, the application of this technique produces odd melodies; depending on the knowledge of interceptors, this could have posed a problem. Selenus therefore presents another technique with polyphonic vocals and in which only the tenor is encrypted, resulting in a somewhat harmonious musical piece.[15] Although this method is quite secure, also because the ciphertext is surrounded by nulls, it would have been very tedious to compose this kind of music, not to mention the skill that it required.

In the *Cryptographia* (1684), the second cryptological handbook published in the German language, Johann Balthasar Friderici[16] suggests an easier cipher. According to the key shown in figure 5, the letters are substituted by three notes each, however, the tone is irrelevant. Instead, the letters are encrypted by a combination of whole, half, and quarter notes. To even out an odd rhythm, Friderici proposes to add nulls in the form of inverted quarter notes.

While at first, this technique might look like the solution to avert detection, there are two issues. First, just as with the last example discussed by Selenus, the sender of the secret message would have to know how to compose a coherent melody. Second, in music composition, the rhythm is just as important as the melody, and this case, it could look funny to the trained eye. Nonetheless, from a cryptographic perspective, this was still safer than methods that relied on the note height, which seem to have been more commonly known. The more natural the music "looked", the better.

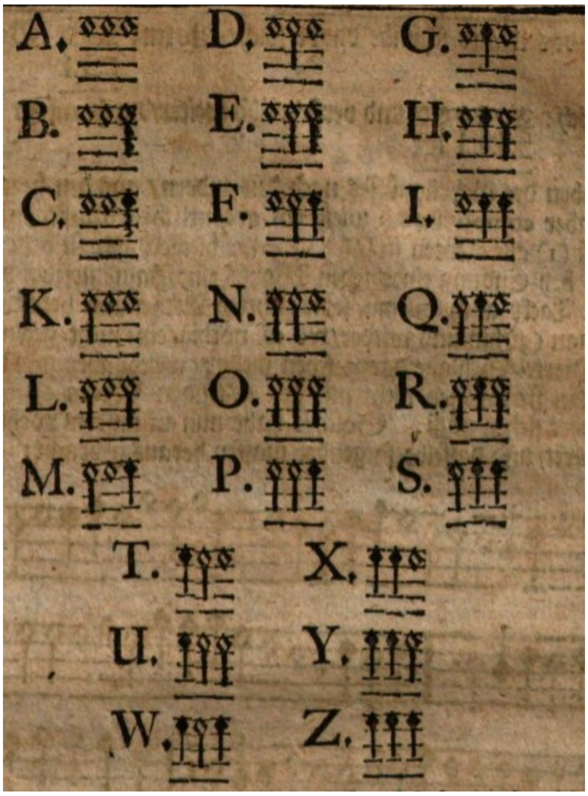


Figure 5: Key for a musical cipher (Friderici, 1684, p. 181). Bayerische Staatsbibliothek München, Res/4 Graph. 19, urn:nbn:de:bvb:12-bsb10897281-3.

Overall, di- or trigraphic ciphers were a bit safer than musical notes that relied on monographic substitutions.[17] Similar to the use of secret alphabets, the success of the technique furthermore relied on the surrounding text. The cipher could remain obscured if the sender wrote about music. Of course, it was also possible to only transmit a musical sheet without additional information. Lastly, the context would have to justify why two people exchanged musical sheets or information about music.

5 Invoices

In the second volume of the *Delitiae Mathematicae et Physicae* (1651), Georg Philipp Harsdörffer presents a cryptological technique for people in besieged cities who want to convey intelligence to persons outside the city walls (p. 6ff.). The numerical ciphers are inserted to a merchant's invoice. As is shown in table 3, the consonants b-m are signified by the numbers 1-

[15] According to Selenus, it was composed by a musician from Luneburg named Friedrich Hollandt (p. 325). In 2017, the piece was played by Pennsylvania State University's orchestra on behalf of Gerhard F. Strasser. See Klaus Schmeh's upload: https://youtu.be/XjHtiXE8Iys.

[16] To this day, it is unknown if this was the author's real name or a *nom de plume*. Friderici plagiarized a lot of his material, although the extent of his plagiarism remains to be examined.

[17] See Porta, 1602, p. 156 or Schwenter, s.d., p. 303. We know, for example, that by the end of the 16th century, complex musical ciphers using di- and trigraphs were used by the papal cryptographic service (Strasser, 1997, col. 784), although it is unclear whether the cipher was supposed to be hidden.

10, the letters n-z by 20-100 and the vowels by zeros.

1	2	3	4	5	6	7	8	9	10
b	c	d	f	g	h	j	k	l	m
20	30	40	50	60	70	80	90	100	
n	p	q	r	s	t	v	x	z	

0	00	000	0000	00000
a	e	i	o	u

Table 3: Key for a cipher to be used in an invoice (Harsdörffer, 1651, p. 6ff.).

Harsdörffer's example message "Jch kan den Ort drey Tage halten"[18] is encrypted by numbers that are then listed as the debts of certain people. As the first and fourth word begin with a vowel, he advises his readers to attach the vowels to the preceding names, as demonstrated here with Claus Pfitz and Moritz Curz:

Claus Pfitz**i**	26	Jch
Conrad Groß	8020	kan
Friederich Beerlin	30020	den
Moritz Curz**o**	5070	Ort
Friederich Demm	35027	drey
Dieterich Plock	7005	Tag
Georg Schwetz auf 3 mal	6097020	halten.

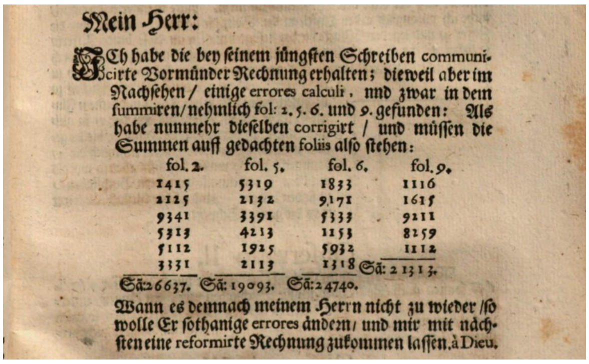
To increase the complexity of the cipher, the encryption operates with one, two, three, four, and five-digit numbers to make cryptanalysis even more difficult. Unfortunately, Harsdörffer failed to take the recipient into account, as the ciphertexts result in multiple plaintexts. For example, the number "30020" encrypts the plaintext "den", but also "daan" "daaca", "deca", "pan", or "paca". Therefore, this is not a practical technique.

Another example of an encrypted invoice is included in Friderici's *Cryptographia* (1684) (p. 166ff.). A variation of the *Polybius Square* serves as the base.

I	a	b	c	d	e	f	g	h
II	i	k	l	m	n	o	p	q
III	r	s	t	u	w	x	y	z
	1	2	3	4	5	6	7	8

Table 4: Friderici's version of the *Polybius Square* (Friderici, 1684, p. 166).

The ciphers are merged into four-digit numbers (quadgraphs), while the number 9 marks a word ending. To hide the message "Dein Verwalter ist dir nicht getrew / schaffe ihn ab"[19], Friderici composes a letter asking the recipient for the correction of an erroneous invoice. The recipient first jots down the numbers listed in the first column from top to bottom and so on, splits up the numbers into two-digits (leaving out the number 9) and consults the square to decrypt the secret message.



Mein Herr:

JCh habe die bey ſeinem jüngſten Schreiben communicirte Vormünder Rechnung erhalten; dieweil aber im Nachſehen / einige errores calculi, nnd zwar in dem ſummiren/nehmlich fol: 2.5.6. und 9. gefunden: Als habe nunmehr dieſelben corrigirt / und müſſen die Summen auff gedachten foliis alſo ſtehen:

fol. 2.	fol. 5.	fol. 6.	fol. 9.
1415	5319	1833	1116
2125	2132	9171	1615
9341	3391	5333	9211
5313	4213	1153	8259
5112	1925	5932	1112
3331	2113	1318	Sã: 21313.
Sã:26637.	Sã: 19093.	Sã:24740.	

Wann es demnach meinem Herrn nicht zu wieder /ſo wolle Er ſothanige errores ändern/ und mir mit nächſten eine reformirte Rechnung zukommen laſſen. à Dieu.

Figure 6: Friderici's encrypted invoice (Friderici, 1684, p. 167). Bayerische Staatsbibliothek München, Res/4 Graph. 19, urn:nbn:de:bvb:12-bsb10897281-3.

Unlike common numerical ciphers, the surrounding text and the appearance of an invoice legitimates the presence of numbers. In contrast to secret alphabets and musical notes, the knowledge of the interceptor would have been irrelevant as long as the amount of money made sense. Moreover, digraphic encryptions presented as quadgraphs as well as the addition of nulls significantly increases information security.

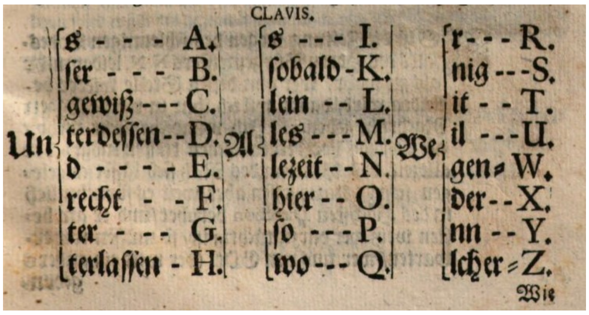
6 Encrypted Null Ciphers

Secret alphabets, musical ciphers, and numerical ciphers in invoices feigned to be something else which ideally resulted in an undetected exchange of information. The downside was that the choice of the steganotext was somewhat limited, as the sender of a message would have to reference languages/scripts, music, or amounts of money. Techniques for concealed ciphers that utilized the Latin alphabet had two advantages: The knowledge of possible interceptors was irrelevant, and they allowed the flexibility to choose any kind of steganotext.

[18] "I can hold this place for three days."

[19] "Your steward is disloyal to you / get rid of him."

According to a technique discussed in Friderici's *Cryptographia* (1684), the letters of the secret message are encrypted by words that are then embedded into a larger body of non-significant text (p. 111). In that sense, this method is similar to Trithemius' *Ava Maria Cipher,* although there are more word types, which are also unrelated to the nature of the information.[20] Friderici uses the key shown in figure 7. The letters of the alphabet are divided into three groups, each of which is assigned a different word beginning ("UN-", "AL-", or "WE-").



CLAVIS.

Un		Al		We	
s	A.	s	I.	r	R.
ſer	B.	ſobald	K.	nig	S.
gewiß	C.	lein	L.	it	T.
terdeſſen	D.	les	M.	il	U.
d	E.	lezeit	N.	gen	W.
recht	F.	hier	O.	der	X.
ter	G.	ſo	P.	nn	Y.
terlaſſen	H.	wo	Q.	lcher	Z.

Wie

Figure 7: Friderici's key for an encrypted null cipher (Friderici 1684, p. 111). Bayerische Staatsbibliothek München, Res/4 Graph. 19, urn:nbn:de:bvb:12-bsb10897281-3.

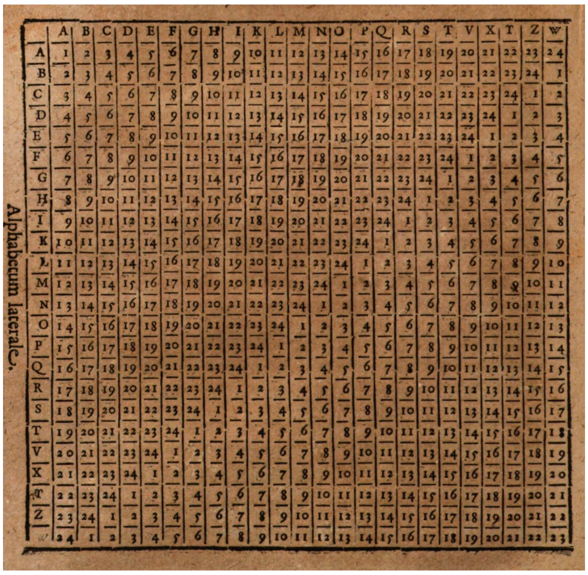
The sample message "wir sind verrathen"[21], is encrypted by the ciphertext:

> "WEGEN ALS WER WENIG ALS ALLEZEIT UNTERDESSEN WEIL UND WER WER UNS WEIT UNTERLASSEN UND ALLEZEIT."[22]

According to Friderici, the sender would now have to integrate these words into a coherent message in the right order. He continues that this technique is especially convenient for the recipient, as they can skim the letter searching for words starting with "UN-", "AL-", and "WE". However, he acknowledges that, in order to avert confusion, the sender would have to refrain from using any of the above listed words in the text surrounding the cipher. In sum, Friderici states that while this technique is rather tedious, it is also "sehr sicher und darzu von allem Verdacht befreyet"[23] (p. 112).

7 Encrypted Semagrams

In the second half of the 17th century, the Jesuit scholar Caspar Schott introduced in the *Schola Steganographica* (1665) a technique that, according to him, could be used for correspondence with princes and governors. He uses a *Tabula Recta* and a keyword as introduced by Blaise de Vigenère in his *Traicté des chiffres* (1586), although Schott refers to Athanasius Kircher's *Abacus Numeralis* using numbers instead of letters.



	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	V	X	T	Z	W
A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
B	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1
C	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1	2
D	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1	2	3
E	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1	2	3	4
F	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1	2	3	4	5
G	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1	2	3	4	5	6
H	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1	2	3	4	5	6	7
I	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1	2	3	4	5	6	7	8
K	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1	2	3	4	5	6	7	8	9
L	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1	2	3	4	5	6	7	8	9	10
M	12	13	14	15	16	17	18	19	20	21	22	23	24	1	2	3	4	5	6	7	8	9	10	11
N	13	14	15	16	17	18	19	20	21	22	23	24	1	2	3	4	5	6	7	8	9	10	11	12
O	14	15	16	17	18	19	20	21	22	23	24	1	2	3	4	5	6	7	8	9	10	11	12	13
P	15	16	17	18	19	20	21	22	23	24	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Q	16	17	18	19	20	21	22	23	24	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
R	17	18	19	20	21	22	23	24	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
S	18	19	20	21	22	23	24	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
T	19	20	21	22	23	24	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
V	20	21	22	23	24	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
X	21	22	23	24	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
T	22	23	24	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Z	23	24	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
W	24	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

Alphabetum laterale.

Figure 8: The Abacus Numeralis (Schott, 1665, p. 77). Staatliche Bibliothek Regensburg, 999/4Art.31, urn:nbn:de:bvb:12-bsb11060024-5.

The key, as shown in figure 8, consists of a table with letters of the alphabet both in a vertical and horizontal order as well as numbers from 1-24 allowing for polyalphabeticity. First, the sender writes the secret message onto a piece of paper for reference. The note shall be:

> "State cauti, cras hora decima noctis venient hostes, ut invadant urbem ex parte orientalis plagae."[24]

Then, they must decide on a keyword; Schott uses the sentence (p. 79):

> "Omnia sunt hominum tenui pendentia filo."[25]

The sender now composes a text in any language with any kind of information. To encrypt the message, they locate the first letter of the keyword, "O" in the top row. Then they move down to where the first letter of the secret

[20] The Ava Maria Cipher also made use of nulls, but the syntactic structure remained somewhat fixed. Additionally, the choice of steganotext (prayers, religious texts) was limited. For this cipher, see Gamer (2022), p. 148.

[21] "We have been betrayed."

[22] "BECAUSE AS WHO LITTLE AS ALWAYS MEANWHILE BECAUSE AND WHO WHO US FAR REFRAINED AND ALWAYS."

[23] "Very secure and, in addition, free from any suspicion."

[24] "Be careful, tomorrow at ten o'clock at night the enemy will invade the city from the eastern quarter."

[25] "All men are hanging by a thin thread."

message “S” is located on the far-right column and finds the number 7. Thus, the first letter is encrypted by the number 7. To encrypt the second letter of the secret, the sender finds the number located between “M” (second letter of the keyword) and “T” (second letter of the secret), which is the number 6, etc.

So far, this is nothing new. But, to “write” the first letter of the secret message (“S”), one should not jot down the number 7, but rather place a small dot or line below the seventh letter of the steganotext. To signify the following letter “T”, which is encrypted by the number “6”, the sender starts from the first mark, counts six letters, and places a dot on the thirteenth letter of the steganotext, etc. This technique was quick and easy to use. Moreover, the sender of a secret message could apply it to any kind of steganotext. The language sounded natural, and the parties could disregard any concerns relating to the communication context or knowledge of possible interceptors.

The writer and poet Erasmus Francisci paid reverence to this method in the third volume of *Die lustige Schau-Bühne* (1673), which contains six fictional conversations between six friends (Helduser 2011),[26] providing valuable insight into contemporary communication culture. The fictional friends state that letters containing ciphers would be confiscated immediately, which is why it would be better to hide a secret message (Francisci 1673, p. 173). This is followed by a story about a French diplomat named De la Haye who was supposedly incarcerated after “the Turks” had intercepted one of his letters that contained some ciphers (p. 177). One of the characters – or rather Francisci – deems Schott’s method to be the safest technique of all and adds that it would be even more secure to make the dots with invisible ink. To this date, it is the only known early modern technique to combine complex cryptography (polyalphabeticity and keyword) with linguistic and technical steganography.

8 Conclusion

As this article has shown, skills for hiding messages became increasingly important during the 17th century. To ensure the highest possible level of information security, scholars combined cryptography and steganography. Although the simultaneous application of these two forms of secret communication fanned out the possibilities of hiding information, they also required a lot of skill and consideration.

The authors of cryptological handbooks had to consider the knowledge of possible interceptors, the naturalness of the language or data, and flexibility in choosing an adequate steganotext. The scholars even made sure that they increased the complexity of the ciphers through the application of inversion, polygraphic substitution, as well as keywords and poly-alphabeticity. The techniques presented in this paper varied in sophistication, and were sometimes not very practical, but they all exemplify the scholars’ ability for computational thinking and creative problem-solving strategies.

As a next step, it would be interesting to see if and to what extent the techniques discussed have been applied in correspondences. However, since steganography obscures the presence of a secret message, this will prove to be a difficult task. In sum, historians still have a lot to discover as far as 17th century cryptology – and hiding information in specific – is concerned. This paper offers only a small glimpse into early modern steganography, which, just like the messages it seeks to conceal, is still hiding in plain sight.

Acknowledgements

This article results from a project kindly funded by the Swiss National Science Foundation (grant no. 210981). Warmest thanks to Theresa Bayer for her helpful suggestions regarding the structure of the text.

References

Cornelius Agrippa. 1533. *De occulta philosophia*, vol. 3. Cologne.

Paolo Benavoglio. 2020. A Partenio’s Stegano-Crypto Cipher. *HistoCrypt 2020. Proceedings of the 3rd International Conference on Historical Cryptology*:36-45.

Karen Britland. 2018. “What I Write I Do Not See.” Reading and Writing with Invisible Ink. Katherine Ellison, Susan Kim (ed.). *A Material History of Medieval and Early Modern Ciphers. Cryptography and the History of Literacy*, Routledge, New York

[26] He also refers to Hardsörffers method hiding ciphers in invoices as a secure way to convey information, p. 175.

(Material Readings in Early Modern Culture):208-222.

David Løberg Code. 2022. Can musical encryption be both? A survey of music-based ciphers. *Cryptologia.* https://doi.org/10.1080/01611194.2021.2021565.

Katherine Ellison. 2017. *A Cultural History of Early Modern English Cryptography Manuals*. Routledge, Abingdon.

Erasmus Francisci. 1673. *Die lustige SchauBühne von allerhand Curiositäten*, vol. 3. Nuremberg.

Johann Balthasar Friderici. 1684. *Cryptographia*. S.l.

Maximilian Gamer. 2022. *Die Polygraphia des Johannes Trithemius nach der handschriftlichen Fassung. Edition. Übersetzung und Kommentar*, vol. 1. Brill, Leiden (Mittellateinische Studien und Texte 56/1).

Barbara Gärtner. 1999. Daniel Schwenter (1585-1636). Ein barocker Mathematiker. Rainer Gebhardt (ed.). Rechenbücher und mathematische Texte der frühen Neuzeit. Adam-Ries-Bund, Annaber-Buchholz:241-247.

Urte Helduser. 2011. Erasmus Francisci: Die lustige Schau-Bühne von allerhand Curiositäten. Nikola Roßbach, Thomas Stäcker (ed.). *Welt und Wissen auf der Bühne. Die Theatrum-Literatur der Frühen Neuzeit. Repertorium*. Herzog August Bibliothek, Wolfenbüttel. http://www.theatra.de/repertorium/ed000043.pdf.

Georg Philipp Harsdörffer. 1651. *Delitiae Mathematicae et Physicae*, vol. 2. Nuremberg.

Georg Philipp Harsdörffer. 1653. *Delitiae Mathematicae et Physicae*, vol. 3. Nuremberg.

David Kahn. 1967. *The Codebreakers. The Story of Secret Writing*. Macmillan Publishing Co. Inc., New York.

Franz Kessler. 1615. *Unterschiedliche bißhero mehrern Theils Secreta*. Oppenheim.

Elsa de Luca, John Haines. 2018. Medieval Musical Notes as Cryptography. Katherine Ellison, Susan Kim (ed.). *A Material History of Medieval and Early Modern Ciphers. Cryptography and the History of Literacy*, Routledge, New York (Material Readings in Early Modern Culture):30-47.

Kristie Macrakis. 2014. *Prisonders, Lovers, and Spies. The Story of Invisible Ink from Herodotus to al-Qaeda*. Yale University Press, New Haven.

Wolfgang Mährle. 2000. *Academia Norica. Wissenschaft und Bildung an der Nürnberger Hohen Schule in Altdorf (1575-1623)*. Franz Steiner Verlag, Stuttgart (Contubernium 54).

Polybius. 1925. *The Histories*, vol. IV. Ed. by the Loeb Classical Library, transl. by Paton, W.R. Harvard University Press, London, New York.

Giambattista della Porta. 1602. *De furtivis literarum notis*. Naples.

Anne-Simone Rous. 2011. Geheimschriften in sächsischen Akten der Neuzeit. *Neues Archiv für sächsische Geschichte*, 82:243-253.

Anne-Simone Rous. 2022. *Geheimdiplomatie in der Frühen Neuzeit- Spione und Chiffren in Sachsen 1500–1763*. Franz Steiner Verlag, Stuttgart.

Klaus Schmeh. 2009. *Versteckte Botschaften. Die faszinierende Geschichte der Steganografie*. Heise, Hannover.

Klaus Schmeh. 2017. Selenus' musical cipher played by the orchestra of the Pennsylvania State university on behalf of Gerhard F. Strasser. https://youtu.be/XjHtiXE8Iys.

Caspar Schott. 1665. *Schola Steganographica*. Nuremberg.

Daniel Schwenter. [1617]. *Steganologia & Steganographia nova*. Nuremberg.

Daniel Schwenter. s.d. *Steganologia & Steganographia aucta*. Nuremberg. (2nd edition).

Gustavus Selenus. 1624. *Cryptomenytices*. Luneburg.

Gerhard F. Strasser. 1982. Die kryptographische Sammlung Herzog Augusts: Vom Quellenmaterial für seine „Cryptomenytices" zu einem Schwerpunkt in seiner Bibliothek. *Wolfenbütteler Beiträge* 5:83-121.

Gerhard F. Strasser. 1988. *Lingua Universalis. Kryptologie und Theorie der Universalsprachen im 16. und 17.*

Jahrhundert. Otto Harrassowitz, Wiesbaden (Wolfenbütteler Forschungen 38).

Gerhard F. Strasser. 1997. Musik und Kryptographie. Ludwig Finscher (ed.). *Die Musik in Geschichte und Gegenwart, Sachteil 6*:783-790.

Gerhard F. Strasser. 2007. The Rise of Cryptology in the European Renaissance. Karl de Leeuw, Jan Bergstra (ed.). *The History of Information Security. A Comprehensive Handbook*. Elsevier, Amsterdam:277-325.

Hermann Stützel: Chiffrierwesen im Dreissigjährigen Krieg. 1963. *Württembergisch Franken 47, Neue Folge 37*: 109-115.

Johannes Trithemius. c. 1499. *Steganographia*. First published in 1606. Frankfurt.

Johannes Trithemius. 1518. *Polygraphia*. [Reichenau].

Blaise de Vigenère. 1586. *Traicté des chiffres*. Paris.

Scherbius and the Enigma
Political, Economic and Military Conditions

From the Order of the Imperial Army 1917 to the Ruin of Chiffriermaschinen AG 1925

Claus Taaks
Munich, Germany
claus_taaks@web.de

> "Luck rarely helped this man in his search for new paths. No matter how often disappointment weighed down his heavy, serious character, the belief in the final success of an idea that was recognized as right never left him. Failures, therefore, could not paralyze him, but only spurred him on to renewed efforts."[1]

Abstract

In 1917, the German War Ministry commissioned Arthur Scherbius to invent a cipher machine.

The early history of the device was determined by political and economic disasters.

The Enigma was taken up by dubious businessmen in 1920. In 1925 that ended in a catastrophe.

Introduction

In 1914, Arthur Scherbius (1878-1929) decided to contribute to the war effort and, in conjunction with a well-known ceramics manufacturer, he invented processes for making radiators out of ceramic instead of metal. These radiators have been haunting the literature since Kahn as "ceramic heating parts".

In October 1915, he was drafted into the telegraph troops, and probably he was assigned to its directorate at the Grand Headquarters (GHQ). Wilhelm Fenner later referred to him as "the talented engineer Dr SCHERBIUS of the Supreme Army Command" (TICOM DF-202. p. 9). The telegraph troops were the only military unit whose command was at GHQ. There they had to ensure a very high volume of secured message connections.

On 1st February 1917, he was seconded to the Weapons and Munitions Procurement Office (WuMBA), the central steering body of the German war economy. There he was deputy head of a department, probably the department for electrical machines and equipment.

1. 1917. The order to develop a cipher machine

In the spring of 1917, most likely at his transfer to the WuMBA, Scherbius received the order from the War Ministry to develop a cipher machine. The order itself and the specifications have not yet been found, but they are mentioned in several documents (BArch, 1919).

2. 1918. First patent, Probemaschine

As is well known, Scherbius applied for the first patent of his rotor machine on 23rd February 1918. The keyboard consisted of a square of 5x5 letter keys, the result was indicated by 5x5 glow-lamps. Due to a lack of materials a "writing" version was announced for the time after the war.

A demonstration machine with two rotors was built, probably in the workshop of the company “Dipl.-Ing. E. Richard Ritter & Co.”, founded in 1911, which sold and installed electrical household appliances. This machine, as Scherbius announced in a letter to the Navy Office on 15th April 1918, was to be demonstrated first at the GHQ in Spa, then at the Reichs-Marineamt (Naval Office) in Berlin.

The principle of the machine was recognised as secure and the naval command demanded 10 rotors that could be exchanged on one axis in any order. Then, also in 1918, two machines with 7 rotors were tested. The great number of electrical contacts made these electromechanical machines so unreliable that military use was out of the question (BArch, RM 5/3566).

[1] Meyer-Delius, Heinrich: Elektrotechnik und Maschinenbau, Vol. 47, 1929, issue 28, 14th July, 1929, p. 610f, translated by Jim Rawlinson

"Security by a great number of rotors" seemed impossible, the first Scherbius-machine had failed.

Additional patent: Pneumatic or hydraulic machines

On 2 June, Scherbius applied for an additional patent for pneumatic or hydraulic machines and proposed how such a machine could be realised. (Pneumatic controls had been known for a long time, since the 18th century.) Such cipher machines were never built in Germany, but the impact of Scherbius' proposal, especially DE425147 of 1920, was considerable abroad: They were developed and patented in Great Britain and in Czechoslovakia (main patents: Hugo Koch 1919, Scherbius 1920, Sidney Hole 1922, Josef Sieber, later called Štolba) and built in Britain in 1925 and 1926, in Czechoslovakia as first version of the Štolba-machine in 1930. (From the second version on, the Štolba was electromechanical.)

In Great Britain, research was carried out until 1934, when Wing Cdr Oswyn Lywood, RAF, ran out of patience and commissioned the "RAF Enigma", from which the Typex was later developed (Ferris, 2005).

4 1919. First version of the Handelsmaschine with line-by-line scrambling

After the armistice, new institutions were set up to close down the army. Scherbius became a consultant in the electrotechnical department of the Reichsverwertungsamt (Office for the realisation of military property) and was in a section of the Armistice Commission and in the Army Peace Treaty Commission. These commissions were allowed to send enciphered messages and possibly he was ciphering there.

He had to reduce the number of rotors of the electromechanical machine to 3 or 4, but this did not meet the military's requirements. Scherbius did not give up. From early 1919 on he searched for processes to improve the security of a "writing" machine, the version, which was demanded by the military and other possible customers.

The four rotors of the first version of the Handelsmaschine were supplemented by a new invention: Scherbius added a line-by-line scrambling ("Umwürfelung", Transposition) to the "Trithemius" of the rotors.

5. 1920. Discussions between Scherbius, the Reichswehr, the Reichspost and the Foreign Office.

The reason for these discussions in 1919 - 1920 was the different interests of the participants. Scherbius knew the representatives of the ministries, they were nearly all former officers of the Second Bureau of the GHQ.

Scherbius demanded payment for his two years of development work and orders for batch production, but his wartime client no longer existed.

The "Provisional Reichswehr" (from 1st January 1921 on "Reichswehr") was not allowed to cipher, but nevertheless had great interest in the device, and wanted to wait for the development of a "writing" machine for regular use. Although the necessary funds had not been granted the Reichswehr insisted that the scrambling ("Umwürfelung") had to remain secret and be reserved for the military alone (BArch, 1919).

The interest of the Reichspost was particularly great, especially because of the competition between Telefunken and Marconi. Ever since the armistice, there had been plans to restore telegraph connections throughout Europe, both by cable and by radio, and to expand them on a massive scale.

In 1920, delegations from almost all European countries, including the German Reich, took part in a conference in Paris and called for a massive development programme. Possibilities of keeping radio telegrams secret were also debated. These were to be secured against "unauthorised eavesdropping", both nationally and internationally. The fact that cipher machines were not mentioned is understandable: on the one hand, there were restrictions imposed by the Versailles treaty and by individual governments, and on the other hand, only two states or companies were leaders in this field, Germany and Sweden. This programme of a worldwide telephone and telegraph traffic was the impetus for a huge development of the communication industry after WWI and thus opened up new prospects for the marketing of cipher machines especially for the inventors Arvid Damm and Arthur Scherbius (Conférence, 1920).

In the negotiations with Scherbius, the Reichspost insisted on both methods, rotors and scrambling ("Umwürfelung") line-by-line. The Reichspost played off two inventors and manufacturers of different machines against each other, Arvid Damm and Arthur Scherbius.

The Foreign Office, which was allowed to cipher, insisted, like the Reichswehr and Reichspost, on a “writing” version, held back and announcing that although there was interest, there was not enough demand for a machine at the time. It followed the development very closely but wanted to wait until the “writing” machine had been tested. (BArch, 1919, Reich).

On 1st April the company “Scherbius & Ritter” was founded to develop and produce cipher machines and Birka controllers (thermostats) for household appliances, especially heating pads. Scherbius was also responsible for their further development, as is evident from the patents he registered. The company was to serve the interests of both founders and to generate income (HR Scherbius & Ritter).

How did Scherbius finance his share in the company? After the war he had a relatively high and regular income from large companies that used his pre-war patents and paid in foreign currency, General Electric and others.

Scherbius was aware of the events of the time in Germany (Kapp Putsch, uprisings in large parts of Germany). However, he was keen to push through his plan: The introduction of his cipher machine in a military version for the Reichswehr and in a civilian version for the Reichspost and others.

On 5th April Scherbius made an offer of the assignment of all rights and, in return, the purchase of a fixed number of machines

- “Writing“ machines: Postmaschinen, or Handelsmaschinen.
- Glow lamp machines: Artilleriemaschinen or Zahlen-Code-Maschinen (number code machines). The only use of the glow lamp machine considered at the time was artillery observation.

On 25th June Scherbius made another offer to the Reichspost, but the Reichswehr insisted on the absolute secrecy of the scrambling (“Umwürfelung”), while the Reichspost insisted on its international distribution. Scherbius' claim to the development costs was accepted, but none of the ministries wanted to cover the costs for two years of development.

On 1st July the Admiralty once again insisted that the “Scherbius'sche Umwürfelungssytem” had to remain secret. But the navy had no money to pay the inventor.

On 13th August, the Reichspost informed Scherbius: “Unfortunately, the government is not in a position to spend funds on the purchase of your patents for a cipher machine. Consequently, the use of your invention is herewith released for foreign countries as well.” The Reich Telegraph Administration was still interested in a machine but wrote: “In view of the efforts of foreign inventors known to you for the introduction of cipher machines for telegraph operation, it can only be recommended that you speed up your work.”

Both methods, polyalphabetic cipher (“Trithemius”) and scrambling (“Umwürfe lung”), were finally expressly released. Scherbius had already received patent application numbers for each of them but withdrew these applications a short time later.

The ministries, including the Reichswehr, continued to be interested exclusively in a “writing” version of the Scherbius machine and only slightly in glow-lamp machines.

The project of a cipher machine for the Reich Telegraph Office was not abandoned. On 17th September Scherbius offered a cipher machine with a delivery time of nine months, which was to show the result by a Wheatstone puncher from Siemens, i.e. an new version of the Postmaschine.

The competitor, AB Cryptograph (Damm), was represented by Telefunken and continued to apply for the delivery of their version of the postal machine, of which the Swedish Telegraph Office was already testing two machines.

On 2nd December, Scherbius demonstrated to the Reichswehr and Reichspost the prototype of the “big machine” he had developed. State Secretary Bredow (Reichspost) presented a draft contract on 6th December, and a contract was concluded on 19th. Later this contract was apparently never mentioned again. However, the fact that a version of this machine was tested by the Reichspost is mentioned several times (BArch, 1919).

On 26th September 1920, Scherbius applied for the patent of a writing machine, DE425147, which contains the line-by-line scrambling for pneumatic machines, and on 10th May 1922, DE385682 for electromechanical machines, the process that the Reichswehr wanted to keep secret and the Reichspost had demanded for international use, was made public.

This made the machine unusable for the military. In this respect, it had failed for the second time.

6. 1921/22. Gewerkschaft Securitas, Securitas, N.V. Ingenieursbureau Securitas and Securitas GmbH

On 21st November 1921, Scherbius transferred 10 patents (1 granted, 9 pending) to the “Gewerkschaft Securitas”, a small Berlin “workshop for precision mechanics” that manufactured “apparatuses for wireless telegraphy and telephony” called “Audioma” and had close relations to a number of similar companies.

In return for the patents and the assurance of further cooperation, Scherbius and Scherbius & Ritter received a minority share in the Gewerkschaft Securitas of maximum 40 %.

On 4th May 1922 the N.V. Ingenieursbureau Securitas was registered in the Netherlands; it was to represent the Gewerkschaft Securitas internationally, register patents and grant licences. The Gewerkschaft Securitas held 60 % of the shares, 40 % were held by Dutch investors. It was already representing Scherbius patents when on 5th May Hugo Koch filed his first patent, corresponding to the earliest state of the Handelsmaschine at the beginning of 1919 and was intended for the operation by media other than electricity too. The Berlin representative Carl Duhm applied for a patent of the Dutch company in Berlin. (HR, ChiMaAG)

There had also been setbacks in the development of Damm’s cryptograph, of which an internationally usable postal machine was planned. At the beginning of December 1922, the three major telegraph companies, Marconi, TSF and Western Union agreed on the machine developed by Damm, and Telefunken joined them on the grounds that the Scherbius machine was still in the development stage. Four copies of Damm’s machine were to be manufactured in Paris by TSF and tested by the four companies. The Reichspost and the Scherbius machine were thus out of the running.

But even this project, which was pushed forward by Damm, was not successful. However, AB Cryptograph was able to continue with Hagelin taking over the supervision of Damm's company in Paris on behalf of Emanuel Nobel (Hagelin, 1994 and BArch, 1919).

On 19th December the lawyer and businessman Adolf Schläfke, and Wilhelm Fritsch, director of the Berlin branch of the Volksbank, founded Securitas GmbH, which was to build the Scherbius machines in small batches in its workshop at Bahnstraße 21 (today Crellestraße) (BArch, 1919 and HR). The managing director was Carl Duhm. Some machines were built there including an early version of the Handelsmaschine, as described by Scherbius in the ETZ (Elektrotechnische Zeitschrift) in 1923, the Diplomatenmaschine which was similar to it, and, with a glow-lamp display, the Artilleriemaschine, which presumably corresponded to the “small Militärmaschine” mentioned by Wik. (Wik 2018)

7. What's a jemmy compared with a share certificate? What's breaking into a bank compared with founding a bank? What's murdering a man compared with employing a man?[2]

In retrospect, it becomes clear who was involved in the “Gewerkschaft Securitas” - apart from Scherbius and Ritter. David Kahn suspected that the company was founded with the purpose “to funnel risk capital into cipher machines” (Kahn 2012, p. 41). The reality in the early 1920s was different:

Businessmen with connections abroad, especially in the Netherlands, took over German banks and manufacturing companies with the help of foreign finance companies. They took advantage of the lack of capital in Germany, the inflation which in 1921 was already very high, the lack of food and the desperate situation of the German government, which tried to stop the disintegration of the country. For their dubious deals, they bought high-ranking former officers, aristocrats and influential politicians, who were given supervisory board positions, shareholdings or simply "love gifts" and loans.

A small group of such businessmen, some of whom had been noticeably active before the war, had already taken over Wollheim Industriegesellschaft. In this group's prestigious office at Voßstraße 18 in Berlin, they planned the purchase or foundation of companies that promised fabulous profits. In the case of the Gewerkschaft Securitas, they published the, supposedly imminent, international introduction of the Scherbius machine, which would make them and the shareholders rich.

The Gewerkschaft Securitas brought together people who were to play a decisive role in the joint stock company to be established: Gottfried Eberbach and his brother Adolf who

[2] The Threepenny Opera (1928) act 3, sc. 3

were known before the war for financial manipulations, the Volksbank which was close to the Zentrum party and the Christian Trade Unions. Some of their officials were involved in such deals during the period of inflation.

On 11th January 1923 the “Ruhrkampf” (the occupation of the Ruhr) began. The French army invaded the Ruhr district, the German government encouraged passive resistance by the workers and the government paid their unemployment allowance. Thus, after a very short time, the government was overburdened in every respect. Inflation, which was already very high by then, rose rapidly, industrial production fell. It was the beginning of an economic and political catastrophe. The very existence of the Reich was at stake.

8. 1923. Chiffriermaschinen AG: “Capitalists‘ Group” vs. “Inventors‘ Group”

The partners of the Gewerkschaft Securitas had 10 patents, an international representation in Amsterdam, and a workshop in Berlin, but not their own development department, and they had only sold a few prototypes of the machine. The business was intended to be a worldwide enterprise, and this was to be made possible by a joint-stock company, the foundation of which was being prepared. This process required:

- Restructuring of the Gewerkschaft Securitas,
- Demonstrations of the machine by former Post Minister Giesberts and articles in newspapers and trade journals,
- A positive report on the machine by a recognized expert.

On 10th February 1923, the owners of Securitas GmbH (workshop in Berlin) transferred their shares to N.V. Internationaal Financierungsbureau, Amsterdam, which was represented by Gottfried Eberbach. Carl Duhm became the managing director.

On 13th June, former Post Minister Giesberts demonstrated the commercial machine before the Berlin Chamber of Commerce, and on 14th June a detailed article appeared in the Vossische Zeitung entitled “Depeschengeheimnis der Funkentelegraphie” (Dispatch secrecy of radiotelegraphy), with the comment that two machines were already in trial operation at the Telegraphisches Versuchsamt and the Foreign Office of the Reich. In July, Scherbius published an article “Radiotelegraphie und Geheimschrift” (Radiotelegraphy and secret writing) in a trade journal. This also refers to the testing by the Reichspost (Kahn, DK 110_04).

From 27th to 29th June, Rudolf Schauffler from the cipher department of the Foreign Office, wrote “Berichte über vorläufige Untersuchungen betr. die Scherbius'sche Maschine (Frage der Sicherheit)” (Report on the initial testing of the Scherbius machine (the question of security). The conclusion of this report was that the Handelsmaschine was breakable, “on the basis of theoretical considerations and possibly with the help of a 'counter-machine'”, but that the effort for this was so very high that this machine could still be used (Politisches Archiv of the Foreign Office).

On 9th July, Chiffriermaschinen AG (ChiMaAG) was founded in Berlin. The Gewerkschaft Securitas made the contribution in kind (the inventory of Securitas GmbH: patents and construction drawings, a few Handelsmaschinen and parts of the Diplomatenmaschine (writing), artillery machine (glow-lamps), parts and tools worth 300,000 marks.

Four investors, represented by N.V. Algemeene Handelsmaatschappij in Amsterdam, subscribed for shares with a nominal value of 200,000 marks, of which one quarter was paid immediately and three quarters were due after the ChiMaAG was founded.

The founding board consisted of Franz Henke, who then did not take up his post, Carl Duhm, Bruno Weigandt. Nearly the entire supervisory board consisted of persons belonging to the Zentrum party and the Christian Trade Unions and several of them also to the “Wollheim Industriegesellschaft”. Besides politicians and lawyers, there were also some very dubious figures of the inflation era.

On 12th August, the full shares of the four investors had not yet been paid, the Dutch company offered an “early share issue” with a high premium, at 25,000 % of the nominal value. This was followed by intensive advertising, in which the bright prospects of the “postal machine” were painted. The issue of shares was supposed to bring the four investors a high profit, even before they had paid their shares. But this did not materialise. From then on, the management of the ChiMaAG consisted of the capitalists’ group, and a minority, the inventors’ group around Scherbius & Ritter, which was not represented in the board.

Immediately afterwards, an event occurred that the capitalists’ group had been waiting for:

Anton Höfle of the Zentrum party became post minister.

Since the permanent collaboration with the inventors' group no longer existed, the capitalists' group needed a technical director. On 2nd September Paul Bernstein (1891-1976), a specialist in precision mechanics, was hired. At the same time, several prominent engineers from the radio and telegraph industry were offered positions on the board.

As the capitalists' group had not yet paid the nominal value of their shares, the Gewerkschaft Securitas withheld the patents and did not transfer them to the ChiMaAG.

At the preliminary discussion of the "capitalists' group" for the General Assembly on 24th September it was proposed that Erich F. Huth, inventor and owner of the renowned company Radio-Huth, was to be appointed General Director and Technical Director of ChiMaAG, and a contract with N.V. Internationaal Financierungsbureau which had been already concluded by the supervisory board was to be put on the agenda. Adolf Hermkes, chairman of the import and export company "Damaraland" in Amsterdam and member of the supervisory board of ChiMaAG, described the expected brilliant business prospects.

Resolutions of the General Assembly at Voßstraße 18:

- A contract was concluded with Erich F. Huth. Additional members were appointed to the supervisory board, some of them prominent and influential. These included politicians of the Zentrum party, representatives of the banks belonging to the capitalists' group, of the group Schiele & Bruchsaler manufacturer of precision mechanics, and Franz Ullstein, publisher of the largest German newspapers.
- The contract with the Dutch Financierungsbureau was accepted by the capitalists' group, against the protest of the inventors' group. The Dutch Financierungsbureau got from the ChiMaAG its rights to receive the patents, which had been withdrawn by the Gewerkschaft Securitas. ChiMaAG received in return 60 Kuxe (shares) out of 100 of the Gewerkschaft Securitas and thus became the majority owner of the Gewerkschaft and had a share in all patent rights, also abroad. The Gewerkschaft Securitas undertook to transfer its shares in ChiMaAG (300 million marks nominal) to the N.V. Internationaal Financierungsbureau.

Satisfied, the lawyer of the capitalists' group stated: "The interests of the company are also served by the conclusion of the contract in so far as this makes any litigation with the inventors unnecessary for the company."

In a further move, the N.V. Internationaal Financierungsbureau, behind which stood the capitalists' group, received the majority of shares of the ChiMaAG. Triumphantly, the Vossische newspaper reported that ChiMaAG was from then on "involved in the cipher machine business all over the world" (HR ChiMaAG).

The calculation of the capitalists' group seemed to work, but then there were several setbacks: On 26th September the Stresemann government broke off the Ruhr campaign. Inflation had reached its peak (US$ 1 equalled 4.2 trillion marks), the economy, including the industry, the banks and the supply of food, especially in the Ruhr district, threatened to collapse, separatist movements threatened the government, there were uprisings, including the Hitler-Ludendorff putsch (Beer Hall Putsch) with the march on the Feldherrnhalle on 9th November.

On 15th November the Rentenmark, a currency backed by the land used for agriculture and business, was introduced. Credit was tightened. The recipe of the inflation profiteers - assets and financing in guilders, debts in marks - became a loss.

On 29th November Scherbius' article about the cipher machine, then called "Enigma", appeared in the ETZ (Elektrotechnische Zeitschrift), specifically discussing the first version of the Handelsmaschine constructed by Scherbius & Ritter. The ChiMaAG was not mentioned in it. Probably he had written the article much earlier.

The administration of the First Marx cabinet began on 30th November. Post Minister Anton Höfle was a "friend" of Adolf Hermkes and Reich Chancellor Wilhelm Marx a fellow party member, so the big order for the postal machine should still be possible.

On 11th December, ChiMaAG demonstrated the Enigma at the Congress of the Universal Postal Union in Berne, exchanging enciphered messages with the Reichspost Ministry in Berlin. Presumably Adolf Eberbach, brother of Gottfried Eberbach, and his friend and business partner Carl Winkler, co-founder of Wifag AG in Berne, with whom the ChiMaAG was also conducting licensing negotiations, were involved. An article in the Kölnische Zeitung in February 1924, obviously launched by the capitalists' group, claimed that a revenue of 1 million Swiss francs was in prospect, and that orders from the Reichspost could also be expected. The share price doubled as a result of this news (BArch R8127).

Huth did not take up his position on the ChiMaAG board. Technology in ChiMaAG was now only represented by the authorised signatory Paul Bernstein (HR ChiMaAG).

9. 1924. The Crossing of Typewriter and Cipher Mechanism

ChiMaAG was in a tight spot. It had to exploit as quickly as possible the relationship with Minister Anton Höfle, who was persuaded to buy and test 20 machines for the Reichspost and to buy shares in ChiMaAG, and a supply contract with the Reichspost was also being discussed. At the same time, "writing" machines that had already been ordered and paid during the days of the inventors' group had to be delivered to the Reichswehr, to Voith AG and to the Šentel company in Prague (chairman Josef Sieber, later named Štolba). The large-scale production of the postal machine had to come as quickly as possible and no matter what the cost, without the inventors' group.

The cooperation with the Schiele & Bruchsaler group which had already been initiated seemed to be the salvation. Alfred Wallenstein, a typewriter engineer of one of their companies, the Uhrenfabrik vorm. L. Furtwängler Söhne AG, had developed the technically very complex and expensive typewriter "Cardinal" which could only be sold in small numbers. On 19th February 1924, ChiMaAG concluded a manufacturing contract for 1,000 machines to be developed from parts of this typewriter and the cipher mechanism of the Handelsmaschine. The engineers Paul Bernstein and Alfred Wallenstein immediately began development in Furtwangen in the Black Forest.

The machine created there, today called the "writing Enigma" (cryptomuseum.com), was unusable, mechanically unreliable and electrically not VDE-compliant. A few machines were sold, but never used. Its cipher mechanism, which was further developed by Bernstein, was supposed to achieve a very high level of security by means of 4 rotors and 4 drive gears which coprime numbers of notches (Patent DE429122, Bernstein, p. 2, lines 72-78).

This particular strength of the "writing" Enigmas was not given attention abroad because such machines had never been built in large series, were hardly used systematically and not in cases of interest to Great Britain.

As soon as possible after this failure, a better postal machine had to be developed and produced on the basis of the Handelsmaschine. This work was taken over by three engineers of another company of the Schiele & Bruchsaler group. A few of the resulting machines were delivered to the Reich Telegraphy Office, they were described in a ChiMaAG brochure and in some magazine articles and can be recognised by the trapezoidal keyboard that narrows towards the front. Devices from other manufacturers could be connected to it on both sides (puncher and reader, printer or typewriter). (Reiner1988)

From March 1924 on German companies were converted to gold marks, i.e. the mark at its pre-war value. During this "stabilisation crisis" it even had to be accepted that many Germans, especially in the Ruhr area, would starve. "This is a damage to the people's nutrition that can only be accepted, but must also be taken if the aforementioned high goal [stability of the mark] is to be achieved." (Hans Luther, then Minister for food and agriculture, Vossische 1st March 1924). Many companies did not survive this crisis. As with other companies, the ChiMaAG, tried to delay the conversion of the balance sheet to gold marks as long as possible.

Hermkes agreed with the management of the Reichspost and with Höfle personally on the purchase of 20 machines for a total of 100,000 marks and the purchase of shares in ChiMaAG for 200,000 marks. Höfle ordered the payment without a valid contract. Later it turned out that Hermkes had not bought shares for the Reichspost as agreed but had spent a large part for other purposes, included the paying of the inventors' group. The whole transaction did not appear in ChiMaAG's books.

Kurt Danziger, the lawyer of the inventors" group, prepared claims for compensation against Hermkes and Eberbach and legal proceedings for evasion of funds and financial irregularities. The capitalists' group, on the other hand, adjourned a general meeting that was due three times and refused to take up the agenda items of the inventors' group. Thus the capitalists' group managed to hide their actions for a time, until finally Danziger requested officially that the meeting be convened.

From 3rd June on, there was a new government, the Second Marx cabinet, in which the Zentrum party was strongly represented.

The capitalists' group saw new opportunities. The annual report of 7th June written by Hermkes and Eberbach again painted a bright future, only "inhibited by the general economic depression". "The machines in use so far have proven themselves very well from a technical point of view."

At the General Assembly of 28th June, all objections by the inventors'group were rejected as was their motion for adjournment. Since this decision was not legally permissible, the meeting was nevertheless adjourned. The press reports on this fiasco of the capitalists' group were devastating for ChiMaAG, but the Board of Directors had still managed to force their position through.

Duhm and Hermkes (and a secretary) exhibited the Handelsmaschine and the Artilleriemaschine at the World Postal Congress in Stockholm. Foreign institutions (military, secret services), including some from Sweden, examined the Enigma and tested some machines. The ChiMaAG lent them the machines for trial.

On 29th July at the meeting of the supervisory board, only the group around Hermkes was present, who was appointed chairman of the board. The conditions of his employment were decided, and very probably also the commission he was to receive. The General Assembly meeting on 13 January 1925 was adjourned because of the "differences of opinion".

In the meantime, the Naval command had decided to test the glow lamp Enigma – the Navy urgently needed cipher machines and the "writing" version was still not usable.

During all these disputes, machines were ordered, built in small series and partly already delivered: 10 machines "Funkschlüssel C" to the naval command, 12 Handelsmaschinen to the company Šentel in Prague, machines to the company Voith, 2 glow lamp machines (Enigma B) to the Swedish General Staff. Several Enigmas of different types went to military and secret services abroad, also to GC & CS.

10. 1925. The ChiMaAG Disaster

On 15th January 1925, the First Luther cabinet came together, in which the DNVP (far right nationalists) had more seats than the Zentrum party.

On 10th February, Post Minister Anton Höfle was arrested for embezzlement and passive bribery. There was no mention of ChiMaAG, it was mainly Julius Barmat, Julius Barmat, a businessman originating from Ukraine and resident in the Netherlands, who was alleged to have bribed him. But it later came out that he had also received a loan from the Depositen- und Handelsbank, a "Schieberbank" (gangster bank) of Hermkes.

Hermkes resigned from office on 21st February.

The general meeting was finally held on 23rd March. In the annual report for 1924, all the actions of the capitalist group as well as the resulting losses were not mentioned and not included in the balance sheet and profit and loss account.

"The year 1924, the first full business year of our company, showed that Chiffriermaschinen A.G. was able to develop forwards and upwards in steady constructive work, despite the unfavourable economic circumstances." The failure of the production of the writing machine is interpreted by the capitalists' group as progress:

"The manufacturing contract concluded with the Schiele & Bruchsaler industrial group in February 1924 proved to be unfeasible in view of the changed conditions of the time and was placed on a different, considerably more favourable basis for the company at the end of 1924, with a complete change in the manufacturing model and thus possible very sharp price reductions."

On 20th April, Post Minister Höfle died in pre-trial detention, which triggered heated discussions in the Weimar Republic about the necessity of detention and about its conditions.

On 3rd July a months-long dispute began over the dismissal of Crilaers, the chairman of the N.V. Ingenieursbureau Securitas appointed by Hermkes. Koch had been appointed in his place, who only succeeded in driving Crilaers out of this office after legal disputes.

The capitalists' group had already left the ChiMaAG. Bernstein had been dismissed, Elsbeth Rinke had procuration, she was close to a major remaining shareholder, the Drahtseilwerke (Wire rope factory) Adolf Deichsel. Koch was at the head of Securitas in the Netherlands, Scherbius and the company Scherbius & Ritter were redesigning the machines, the inventors' group was supported by the shareholders, there were negotiations with the Schiele & Bruchsaler group for the compensation of the broken contracts. First steps to rehabilitate the company had begun.

The actions of Hermkes and Eberbach, especially their relationship with Postminister Höfle, had not yet become public, allegedly they were not known to the inventors' group either. The annual report for 1923 as well as that for 1924 showed a profit - albeit a relatively small one - the relationship with Schiele & Bruchsaler was still being negotiated, the purchase of the 20 machines by the Reichspost as well as the agreed purchase of shares were not recorded in ChiMaAG's books. It looked as if ChiMaAG had problems, but that was nothing special in 1925.

11 The Showdown

On 15 July, the Barmat Committee of Inquiry of the Reichstag also dealt with the business dealings between ChiMaAG and the late Anton Höfle resp. the Reichspost. The minutes of this meeting were published in nearly all national German newspapers, partly also commented (e.g. RAnz 1925/164, p. 1).

ChiMaAG was ruined.

The 1925 business year ended with considerable debts. Until then, the various models of the Enigma – in contrast to the publications of the "capitalists' group" – had only been produced and sold in very small series. The introduction of the Postmaschine, which was supposed to bring the breakthrough, had finally failed.[3]

How and by whom ChiMaAG was restructured and new Enigma models were constructed is the subject of another publication which also covers how ChiMaAG became a supplier to the Reichswehr and paid off its debts from 1925 by early 1930s.

A detailed history of the Enigma and biographies of some of its protagonists are planned.

Acknowledgements:

The author is most grateful to all the persons whose private archives greatly contributed to all his research, not only about 1917-1925, especially Frode Weierud for his advice during many years and for the materials he generously shared with me. Olaf Ostwald, who is researching the technical and cryptological side of the first Enigmas, corrected my non-professional judgements and errors about cryptography. I also thank the staff of all the archives I used: Their employees helped an amateurish researcher to find the key information about companies and persons.
And I thank my neighbour Jim Rawlinson for correcting the translation.

References

Reference works:

Deutscher Reichsanzeiger und Preußischer Staatsanzeiger. Digital edition: University of Mannheim (referred to as RAnz year/No.) https://digi.bib.uni-mannheim.de/periodika/en/imperial-gazette/

„Akten der Reichskanzlei. Weimarer Republik" online. Die Kabinette Stresemann I/II./Band 2/Dokumente
(Especially: No. 179. Besprechung mit den Vertretern der besetzten Gebiete im Kreishaus in Hagen vom 25. Oktober 1923), https://www.bundesarchiv.de/aktenreichskanzlei/1919-1933/0000/index.html

Patents: depatis.net; entries in the RAnz; cryptomuseum.com

Claus Taaks: Chronology (will be continued). PrivArch CT, partly containing Twitter-entries of Frode Weierud. Unpublished manuscript (in German, partly in English)

ZEFYS - ZEitungsinFormationssYStem of the Staatsbibliothek zu Berlin.

Wikipedia (de. and en.) - *Normally not a reliable source - but the entries on cryptography and cipher machines are trustworthy.*

Books, Articles and Archival documents:

Eugen Antal and, Pavol Zajac, 2021. The first Czechoslovak cipher machine. In: Cryptologia online, accessed 28.12.2021.

BArch, 1919, R 4701 / 8665, Reichspost, Erfindungen

BArch, MA RM 5 / 3566, Geheimschriftmittel, Angebote, Verschiedenes vom Januar 1918 bis Juli 1919. Az. II 8.-12., Bd. 1.

BArch, R8127, 6937. BHG. *(Newspaper clippings on the ChiMaAG)*

Conférence internationale pour l'amélioration des communications postales et ferroviaires,

[3] Only one person had been able to exploit the ruin of ChiMaAG for his own purposes: The reference to the "completely useless Barmat machines" and the claim that his machine, invented in 1918, could become the better and cheaper postal machine, helped Alexander (von) Kryha to get the first of his investors, who all lost a lot of money. But that is another story.

télégraphiques et téléphoniques et radiotélégraphiques, Paris, 7-13 juillet 1920, Paris, Imprimerie nationale, 1920.

Jane Desborough, Curator of Scientific Instruments, Science Museum, London: Correspondence about Sidney Hole, 2022

Donald W. Davies, 1984. Sidney Hole's Cryptographic Machine, Cryptologia, 8:2, 115-126, DOI: 10.1080/0161-118491858881.

John Robert Ferris, 2005: Intelligence and Strategy: Selected Essays. Studies in Intelligence Series. Routledge, Taylor & Francis. London & NY 2005, pp. 138-181.

Hagelin Crypto, in-house newspaper, 1992: 100 years of Boris Hagelin https://www.cryptomuseum.com/crypto/hagelin/files/100_Jahre_Boris_Hagelin.pdf

Handelsregister (Commercial register) (HR) of ChiMaAG and H&R: Landesarchiv Berlin, A Rep 342-02 No.21576 and A Rep. 342-02, Nr. 21577

HR of Scherbius & Ritter: Landesarchiv Berlin, A-Rep. 342-02, Nr. 42878

HR of N.V. Ingenieursbureau Securitas and N.V. Damaraland. Noord-Hollands Archief, Z-15-07718Z-15-07718

David Kahn, 2012. Seizing the Enigma - The Race to Break the German U-Boat Codes. 1939-1945. Naval Institute Press, Annapolis. (First edition 1991.)

David Kahn, DK 110_02. National Cryptologic Museum, Ft. Meade *("Arthur Scherbius and the Early Days of the Enigma", remarks by Jürgen Rohwer – preparatory work for "Seizing the Enigma".)*

David Kahn, DK 110_04 National Cryptologic Museum, Ft. Meade *(Containing copies of Articles from and about Scherbius.)*

Gottfried Korella, 1976: Über die Zusammenarbeit der deutschen Post mit Heer / Wehrmacht im Fernmeldewesen von 1900 bis 1945. In: Archiv für deutsche Postgeschichte. Issue 2/1976, p. 25-45

Louis Kruh, Cipher Deavours, 2002: The Commercial Enigma: Beginnings of Machine Cryptography. Cryptologia. January 2002 Volume XXVI, Number 1.

Karl De Leeuw, 2003. the Dutch Invention of the Rotor Machine 1915-1923, Cryptologia, 27:1, p. 73-94

Meyer, Joseph A. WICHER. Der Fall WICHER. German Knowledge of Polish Success on ENIGMA. (n.d., 1975?). https://www.nsa.gov/portals/75/documents/news-features/declassified-documents/tech-journals/der-fall-wicher.pdf (p. 19, fn 124)

Olaf Ostwald, 2023: Cryptographic design flaws of Early Enigma. https://cryptocellar.org/enigma/files/enigma-design-flaws.pdf

Politisches Archiv des Auswärtigen Amts, folder "Enigma Unterlagen und Bearbeitung" Bestand Rückgabe TICOM, S8: T-3342, Box 165.

Kurt Reiner, 1988: Reiner. 1913-1988. 75 Jahre Firmengeschichte. Company publication Furtwangen

Claus Taaks, 2021: Arthur Ludolf Jacob Scherbius. Ein leidenschaftlicher Erfinder. Unpublished manuscript (in German)

The National Archives: Early correspondence relating to the ENIGMA cypher machine, including the patent specifications and photographs. HW 25/6

TICOM DF-202. Wilhelm Fenner: The History of the Cryptologic Agency (Translation of Fenner 1945: "Die Geschichte der Chiffrierstelle").

Frode Weierud, 2018: Enigma Utvikling (Lecture about the development of the Enigma. In Norwegian.) ENIGMA_Utvikling_v1.pdf - Google Drive

Anders Wik, 2018: The First Classical Enigmas. Swedish Views on Enigma Development 1924-1930. In: Proceedings of the 1st Conference on Historical Cryptology, pages 83- 88, Uppsala, Sweden, 18-20 June, 2018.

“We just did it!” – Female employees in Swedish sigint during the Second World War

Fredrik Wallin
FRA
Box 301, 161 21 Bromma, Sweden
fredrik.wallin@fra.se

Abstract

When the Swedish sigint agency, FRA, was formed in 1942, the civilian personnel consisted of 67 percent women. This paper explores the roles and duties performed by women at the FRA during the wartime years and their working conditions.

1 Credits

This paper is based on documents in the FRA archives. Most of the material used is official documents, but parts consist of material collected by personnel at the FRA interested in preserving the history of the agency, including taped interviews with wartime employees. I want to mention in particular Bengt Beckman, author of the book ‘Codebreakers’[1] , and Sven Wäsström[2], who helped collect and preserve the “softer” parts of the history of the agency in the archives.

2 Introduction

The Swedish Defence Radio Establishment, Försvarets Radioanstalt, abbreviated FRA in Swedish, is the Swedish government agency for Signals intelligence. The FRA was created in 1942 from the parts of the Swedish Armed Forces High Command performing signals intelligence, mainly the Crypto Department.

Sweden had started modest preparations for signals intelligence already in the 1930s, by various exercises and training of suitable conscripts as cryptanalysts. In 1939 there were about a dozen people employed with these duties at the Crypto Department. They were all officers, and one woman, a secretary named Eva Löfvenmark who was the first civilian employee of the future FRA.[3]

Sweden was not an active participant in the Second World War, but the outbreak of war naturally led to a ramping up of intelligence activities. There was a very rapid growth of the organisation from about a dozen people in 1939 to almost 400 persons in 1942, when the FRA was created as a separate agency.

When the FRA was formed as a separate government agency on 1 July 1942, the agency had 287 civilian employees, of which 193 were women and 94 men[4], which gives a percentage of 67 percent women. In addition, there were 113 military personnel seconded to the FRA, so all in all the agency consisted of 400 persons at the time it was created. Why so many women? Many men were called up for military service, or were expected to be so, which led to an ambition to use women in those defence related duties where it was possible.

Another reason for the prevalence of women was that much of the work at the agency consisted of various kinds of entering data or typing up reports, or doing routine statistical or mathematical compilations[5]. Those were duties that were traditionally performed by women in the office workplaces of the 1930s.

[1] Bengt Beckman: “Codebreakers: Arne Beurling and the Swedish Crypto Program During World War II.”

[2] Sven Wäsström was a former head of the analysis section at FRA

[3] FRA. Löfvenmark.

[4] FRA. List of FRA personnel 1942.

[5] FRA. Descriptions of duties performed in certificates of employment for female personnel.

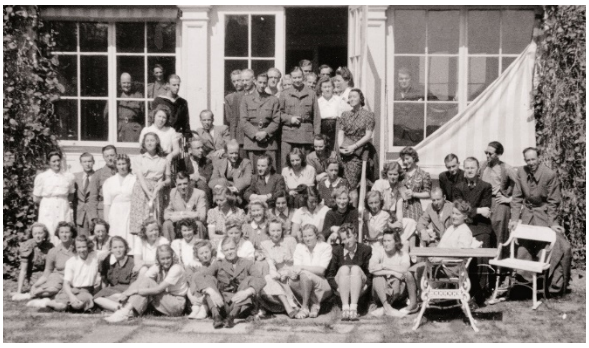


Fig. 1: Personnel outside an FRA site, c. 1942. The large proportion of women is apparent.

The large number of women in a Sigint agency was far from unique for Sweden. When one looks at the sigint organizations in other nations during the war, one can see that in Britain, the USA and Germany there was also a substantial number of women working in the sigint services, presumably for the same reasons.

3 Duties performed by women

Women at the FRA were mostly found in the lower paid grades with titles like office assistant, typist and similar[6]. Some middle positions at the agency had both men and women in the same grades, and the higher grades at the agency were dominated by men. This condition was of course very similar to the situation in the working life in general in those times.

The personnel at FRA were divided into working groups of varying size depending on the type of duties. Most often groups consisted of between three and ten people. It was common for employees to be moved between different groups, partly because of the changing demands for workers, but also as a conscious policy to try out employees in different kinds of work[7].

The most common type of work performed by women concerned various kinds of typing. This could be of reports, compilations or connected to codebreaking[8].

In some cases, women were trained by the agency to perform work normally typically performed by men. For example, in 1943 a number of women were trained to be telegraphists, a traditionally male occupation[9]. There was also a drive to use women as operators in interception of teleprinter traffic and automatic Morse.

Below are more detailed descriptions of some of the duties performed by women in the wartime FRA.

3.1 German Geheimschreiber traffic

In the spring of 1940, Germany asked Sweden to use telegraph lines through Sweden for traffic to occupied Norway. Permission was given by the Swedish authorities, but the traffic was intercepted. Expecting this, the Germans encrypted their traffic with the Siemens & Halske T52, the so called Geheimschreiber.

The Swedish armed forces crypto department started analysis of the code, and relatively quickly the code was broken in principle through work led by Professor Arne Beurling, one of the conscripts trained in the 1930s. However, even if the workings of the code is known, it is a very painstaking and slow process to decode a machine cipher by hand. After the daily key was broken, the practical decryption work was initially done by young women by hand. It took three weeks for seven "girls" to decrypt the traffic from one day. This was obviously too slow for the material to be of use as an intelligence source.

In cooperation with an engineer from the LM Ericson company, a machine was developed to decrypt the traffic. This decryption machine was called "App" at the time, an abbreviation of "Apparatus". The apps were manufactured in great secrecy by the Swedish Cash Registry Company, a subsidiary of LM Ericsson[10].

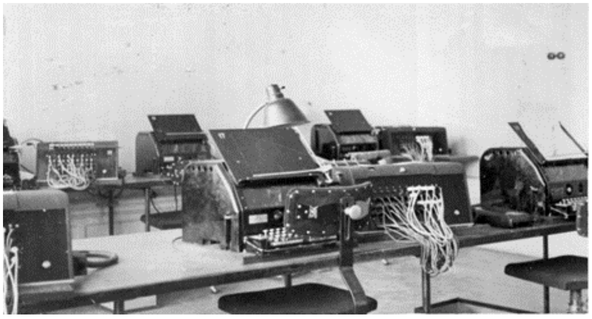


Fig. 2: Apps and teleprinters. Each station would be operated by a young woman typing in the encoded text.

[6] FRA. Lists of FRA personnel 1942 and 1945.
[7] FRA-order.
[8] FRA. Work description.
[9] FRA. List of FRA personnel 1945
[10] FRA. Receipt for the manufacture of equipment.

The machines were not fully automatic. The key had to be set, and then the encrypted text was entered on a teleprinter connected to the app. The codetext was entered on the keyboard of the teleprinter and the signals passed through the app, where they were decoded, and back to the teleprinter, where the plaintext was written out on a tape. This typing of the encrypted traffic was done by women, and it was not simple work restricted to just the typing. The app had to be continually monitored to make sure it did not go out of sync, in which case it produced only gibberish.

The whole workflow included interception and pasting teleprinter tapes with encoded text on papers, then breaking the daily key, done by male mathematicians[11]. Then typing it into the apps for decryption, pasting the resulting plaintext on papers, cleaning up the language and lastly typing out the end reports. In total 175 persons were employed with the process, the majority of them women. The work was continuous around the clock, and was in shifts. In total, more than 100.000 messages were decoded, printed out and delivered to recipients during the war.

3.2 Interception of voice radio traffic

Interception of voice radio was a relatively new field for Swedish sigint during WW2. Telegraphy was established since a long time, and Morse interception was a typically male area of work. In contrast, voice interception was seen as typical female work, to the extent that the job description “voice interceptress” was commonly used in writing. A document recommends the use of “female personnel with good language skills”.

In a study from 1943 outlining the organisation of a voice intercept detachment tasked with intercepting German Luftwaffe traffic, 20 female interceptresses were needed, working five shifts for around the clock coverage, with four being on duty at any given time. They were supervised by a male telegraphist. The men worked in four shifts, while the women worked in five shifts, it being assumed that women were less resistant to the rigors of shift work.

3.3 Teleprinter and automatic morse interception

Interception of radio teleprinter traffic and automatic Morse traffic was another field that was developed during the war. Initially, the interception was performed by male conscript students. The output was enormous amounts of tapes, either with text or printed Morse code. These tapes had to be typed out, and this work was performed by women. To make the work efficient, a machine called a puller was used that propelled the tape across the top of a typewriter. The speed of the puller was regulated by a foot pedal. This was very monotonous work.

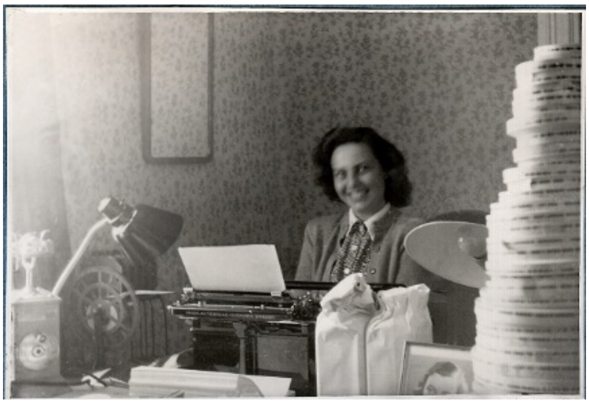


Fig. 3: Typist with rolls of tapes awaiting typing, c. 1942.

A problem with the interception was that once the conscript interceptors had been fully trained, there wasn’t much time left of their callup period, which led to a repeated needs to train new interception personnel. To alleviate this, trials were made with training female personnel as interceptors. Normally anything technical was performed by men, so this was a marked innovation for the time. The trials were a great success, and number of female operators were trained, completing an examination at the end of the course[12].

This activity was expanded, and in 1943 it was described as consisting of 20 female assistants with a male engineer as head of department. As in the example with voice interception, the women were organised in five shifts of four women each. A perhaps even more radical innovation was that one of the four women in each shift was shift manager; in other words, the

[11] FRA. Borelius.

[12] FRA. PM on the use of women in automatic morse.

women were expected to be able to lead themselves, with no male involvement A problem with the interception was that once the conscript interceptors had been fully trained, there wasn't much time left of their callup period, which led to a repeated needs to train new interception personnel. To alleviate this, trials were made with training female personnel as interceptors. Normally anything technical was performed by men, so this was a marked innovation for the time. The trials were a great success, and number of female operators were trained, completing an examination at the end of the course .

This activity was expanded, and in 1943 it was described as consisting of 20 female assistants with a male engineer as head of department. As in the example with voice interception, the women were organised in five shifts of four women each. A perhaps even more radical innovation was that one of the four women in each shift was shift manager; in other words, the women were expected to be able to lead themselves, with no male involvement.

3.4 Codebreaking assistant

Code breaking was performed in a number of different areas. Diplomatic and military codes were broken to a smaller or greater extent from a number of countries. The actual mathematical analysis in breaking the code was in almost all cases done by men, usually with an academic degree. However, as we have seen in the case with the Geheimschreiber, even if the workings of the code is figured out, the daily decryption of encoded messages is very labour intensive. Most of the routine decryption work was done by female codebreaking assistants. The work often involved writing up the encrypted text on large sheets of paper mounted on artist's worktables. This led to one of the rooms where this work was performed being called "the gallery".

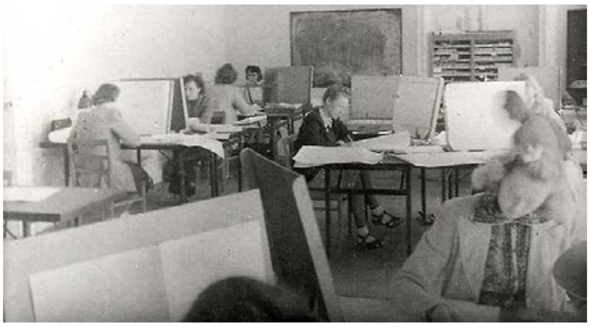


Fig. 4: The "Gallery".

One area where it was popular to work was the French code department, led by the Franco-Swedish cryptologist Yves Gyldén. He had many French mannerisms, and when a new code was broken he served champagne to the entire group, with glasses being specially acquired for the purpose[13].

A not so well known field of decryption was weather reporting. During the war, the belligerent nations encrypted their weather observations, to prevent the enemy from using the information. This led to problems with accurate weather forecasting in Sweden, which hampered the activities of the Swedish Air Force.

To alleviate this, the FRA started a cooperative project with the Swedish Meteorological Bureau to decrypt and use foreign weather observations[14]. German and Soviet weather observations were regularly broken and read, but the encryption of British weather observations proved too difficult to crack. The bulk of the routine work was done by female codebreaking assistants, supervised by male mathematicians. The activity took place at the offices of the Swedish Meteorological Bureau.

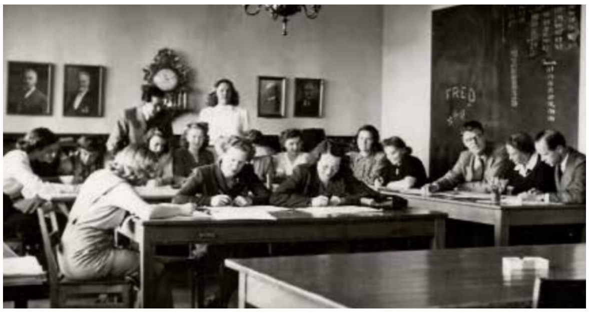


Fig. 5: Female codebreaking assistants at the FRA office at the Swedish Meteorological Bureau in 1945.

4 Background and education

Most of the women hired by the FRA during the war years were very young. A majority were in their early 20s, and some were as young as 16 years. In 1945 the average age for female employees of the FRA was 24 years[15]. For men it was 30 years. As a comparison, the average age for employees at the FRA today is 47 for women and 46 for men.

[13] FRA. Löfvenmark.
[14] FRA. PM H 153/44.
[15] FRA. List of FRA personnel 1945.

Most of the women employed by the FRA were unmarried, and indeed one reason for the low ages of the female employees is that at that time women worked until they married, and after that they normally quit working life.

The skills wanted in female applicants were mainly typing and language skills. The main languages sought were German, English, French and Russian. Knowledge of Russian was relatively uncommon in Sweden, so applicants with knowledge in Russian were highly sought after. German was the most common second language in Sweden in the 1940s, with English and French also being fairly common.

The majority of women working at FRA had some kind of secondary school education. In 1945 about 10% of both men and women in Sweden completed an examination from secondary school. Thus, the women working at FRA had a considerably higher level of education than both women and men in general. The proportion of employees with a secondary school degree was actually higher among women than among men at the FRA at that time. Many of the men, being telegraphists, had only primary school and some kind of military or civilian telegraphist training[16].

If we look at university level education, 12 women at FRA in 1942 had some level of degree from university, compared to 21 men. An interesting observation is that the proportion of women among all employees with university education at FRA was more than 30%, while in general only 20% of university students in Sweden at the time were women[17].

So, in general, the women working at FRA were much better educated than women in general at the time, and indeed better educated than the general male population.

5 How women were recruited

As mentioned, the FRA was formed out of the sections at the Armed Forces High Command that were employed with signals intelligence and codebreaking, and took over the existing personnel that were employed there. Under the High Command there was policy to recruit daughters and other relatives of officers or existing personnel to female sigint positions, as it was considered an advantage from a security standpoint[18]. As an aside, this led to rather large proportion of female employees having names from the nobility, as it was not uncommon for higher officers to be of noble families.

Some of the women came from the Women's Voluntary Defence Organization[19], others came as a result of a course in cryptography for female students organized by the Armed Forces High Command at the University of Uppsala in 1938.

These sources of personnel were not sufficient with the rapid expansion during the early years of the war, and the agency also resorted to advertisements. At that time, advertisements for jobs were divided into male and female in the newspapers. Accordingly, the FRA had adverts for female typists. As both the existence and the activities of the agency were secret in those days, the advertisements did not give the name of the eventual employer. Instead there was just an anonymous signature like "Personnel manager" or "Work for typists". This might seem strange today, but in fact this was a rather common practice in Swedish newspaper ads of the 1940s.

Lastly, the FRA also cooperated with the Public Employment Service, who recommended suitable female work applicants to the agency [20].

Applicants to the FRA were not taken to the offices of the agency for interviews. Instead, offices were borrowed at other places, for example at one of the larger Swedish insurance companies of the time, Skandia-Freja. Applicants were interviewed and tested at these offices, and only when they had been accepted were they told who their real employer would be [21].

Tests for female employees in almost all cases included a typing test. In many cases that was all, but in 1942, the agency developed a number of "Psychotechnical tests", in practice IQ tests with a strong tendency towards aspects of codebreaking [22]. These were used for applicants, but also to find suitable talents in codebreaking

[16] Ibid.

[17] Swedish Higher Education Agency.

[18] FRA. Krybo order June 5 1941.

[19] FRA. Certificate of employment Nr 140/1943 and FRA. Krybo-order June 23 1941.

[20] FRA. Personnel file Bojan Frykholm.

[21] FRA. Interview with Maj-Britt Skoglund.

[22] FRA. Psychotechnical tests.

among the already employed personnel [23]. As a result, some already employed women who scored well were given courses in cryptanalysis.

6 Working conditions and health issues

Apart from pay grades, there were also in the 1940s several different subgroups of government employment, with the main difference being job security, pensions and sickness benefits. A majority of the personnel at FRA were employed as "extras", which meant a minimum of benefits and not much job security.

In 1936 a common pay grade scale was introduced for both men and women in Swedish government service. In practice, women tended to cluster in the lower pay grades. However, the pay was fixed for each grade, so if a man and a woman had the same position, they were paid exactly the same.

Working hours were 40 hours per week if working shift and 42 hours a week if working regular office hours. However, from June to September, work hours per week were reduced to 34.5 hours per week, at that time called "summertime". This reduced the average hours per week over the year to something close to modern hours worked. Overtime work appears to have been common.

Women working shifts had a working time of 35 hours per week, this being explained in a period document as "shift work has proven to have an adverse effect on the psyche, with female personnel being less resistant than men to the pressures of working shifts".

Some duties entailed working until 22.00 in the evening, then sleeping at the workplace and starting again at 06.00 the next morning. In those cases, 25% of the sleeping time was counted as time worked.

The new site for FRA at Lovön was completed in 1943. As there was need for some personnel to sleep at the workplace, there were barracks arrangements for both male and female personnel. Human nature being what it is, there were strict rules against male personnel being in sleeping areas reserved for women and vice versa.

Being absent due to illness was common, with around 10% of the total personnel being ill at any one time not being uncommon [24]. Women had four times as many days absent from illness as men, counted per employee.

The high level of absenteeism among the female employees was noted by management, and several investigations were conducted in the causes. It was concluded that the number of workdays lost from sick leave were considerably more in the lower pay grades, while women in higher positions had only slightly more illness days than men in the same grade.

The phenomenon of high female rates of absenteeism was by no means unique to the FRA. A government study from 1953 came to the conclusion that if men and women with the same work duties were compared, sick leave for the women were invariably much higher[25]. The numbers for the women were highly influenced by the fact that married women were far more likely to be absent than unmarried women. The report concluded that this had to do with "responsibility for home and care for the children".

As a comparison, sick leave today at the FRA is much lower than in the 1940s, both for men and women, but still slightly higher for women than for men [26].

Accidents and mishaps occurred from time to time. In the 1940s as today, government employees having an accident while on their way to work could report that as a workplace accident.

Accidents befalling women at the FRA during this time included[27]:

- Falling in a stair while leading a bicycle on the way to work.
- Fainting during work in the kitchen and knocking the chin on an open cupboard door and the head in the floor.
- Slipping on a newly polished floor and hurting the knee so that "swelling

[23] FRA-order, March 27 1942.

[24] FRA-order.

[25] SOU 1953:18.

[26] FRA annual accounts 2019.

[27] FRA, Reports of accidents at the workplace.

appeared and absolute stillness had to be maintained"

- On one occasion the bus to the FRA collided with a truck and a woman working in the kitchen was hurt so badly that her leg had to be amputated.

7 Retention issues

During the first years of the Second World War, Swedish sigint underwent a very rapid expansion, from a dozen persons to around 400, but quick expansion has its drawbacks. All were not content at their new place of work. We have already discussed the relatively high levels of sickness absence. Another issue probably related to the quick expansion was rapid turnover of personnel. During 1943 over 50 percent of the women working at FRA quit their jobs and had to be replaced. Among the men, the corresponding ratio was 33 percent[28]. This was of course a far from an ideal situation for a secret organization and studies were made to find the causes of the problem.

Among the women, one obvious cause was marriage. The women working at FRA were mostly in their early twenties, and thus very much of marriageable age. As mentioned before, at the time it was common for women who married to quit working, if not immediately, at least when children arrived. In 1943 21 women working at the agency got married[29]. Compared to the total of 98 women ending their employment at the FRA, that is only a fifth, but it still accounts for a large part of the difference between women and men leaving the agency.

Other causes found were that the work, while perceived as adequately paid, was considered repetitive and boring. Prospects for advancement were not good, and as the activities at the agency were expected to be cut down after the war, many employment contracts were of the less permanent type with lower levels of benefits for the employee. Many were not happy to work shifts.

These problems were not unique for the FRA. In 1944, a government enquiry was initiated to find solutions to the problem of high turnover of female personnel among government employees.

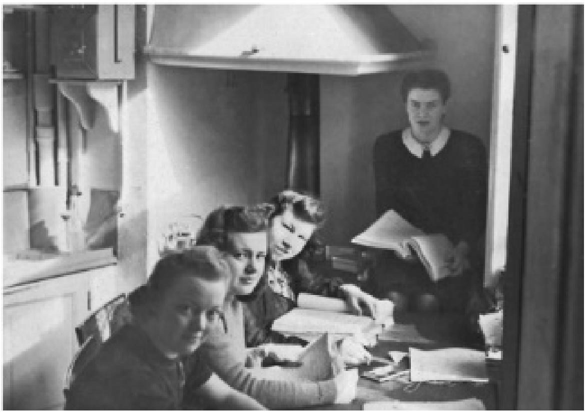


Fig. 6: Women working in the archive at the main FRA site in 1943.

Recommendations from the study included better conditions of employment with more job security, and clearer regulation of the possibilities of promotion[30].

Another cause for the great turnover in 1943 and 1944 was that the agency moved to the new site at Lovön outside Stockholm. The previous offices were to a large extent more centrally located in Stockholm. Lovön was felt to be very far away in those days.

This was alleviated by instituting chartered buses that ran from convenient spots in the centre of Stockholm directly to the agency at Lovön[31]. However, these buses were often full, and many had standing room only. A system of reserved seats was instituted, were personnel senior in age had their own seats. Interestingly, women became seniors at the age of 30, while men had to wait until they turned 40 for the coveted reserved seat[32].

All in all, these measures seem to have reined in the rapid turnover of personnel. In the post-war year 1953, turnover of personnel was down to 10%. As a comparison, today the turnover at FRA is 9.3% for women and 7.4% for men[33]. It should also be mentioned that a substantial part of the women hired during the war liked their work at the FRA, and continued their employment until they retired.

[28] Numbers compiled from FRA-order.
[29] Ibid.
[30] SOU 1946:66.
[31] FRA. Missive to the government re buses.
[32] FRA Tjänstemeddelande June 14 1944.
[33] FRA annual accounts 2019.

8 Women as managers

During the time in question the FRA was organizationally divided into three departments or bureaus, the Signals bureau, the Analysis Bureau and the administrative bureau. The departments in turn consisted of sections, and each section contained a number of working groups. As mentioned, a working group normally consisted of three to ten persons working on a clearly defined subject or target, for example Soviet Baltic Navy ciphers or German Air Force voice traffic. The total number of groups varied, but were usually around 30 to 40.

There were women serving as managers of working groups, but there are no records of women managing sections or departments. In total, the records mention at least nine female managers of working groups at various times [34]. Groups with a female manager could contain male employees or conscripts.

There appears to have been no bias against female managers at the working group level, though it must be admitted that is doubtful if such sentiments at the time would have been recorded for posterity in the archival material. It seems that group managers were selected on ability and previous education level, which is confirmed by the fact that most of the female group managers were among the women with highest education levels.

9 Examples of women working at the wartime FRA and their careers

5.1 Marina Löfström

Marina Löfström was born in 1906 and was daughter of the Finnish general Ernst Löfström, of Finland-Swedish descent. At the time of her birth, Finland was part of Czarist Russia, and General Löfström served in the Czarist army. His wife was from the Russian nobility and Marina was born in Saint Petersburg, where the family lived until the Russian revolution, after which they moved to Finland.

Marina knew a number of languages: Russian, Swedish, Finnish, French, German, English and Italian. She gave Russian as her mother tongue, and knowledge in the other languages as very good. She had an education in business correspondence in Helsinki[35].

When the Winter War broke out in 1939, Marina, her sister and mother rapidly fled to Sweden (the general had died in 1937). We do not know the background to this, but it is fair to guess that the family, having lived through the Russian revolution, feared a communist takeover of Finland.

Marina and her sister were more or less immediately hired by the crypto department of the Swedish Armed forces high command. This is remarkable, as there were strict rules for employees against socializing with foreigners, and here we see two foreign citizens going straight into the most secret part of the Swedish defence efforts. We can only assume that General Löfström had good connections with his Swedish colleagues, and that the sisters excellent command of Russian was of great interest to the agency.

Marina, who was described as the more gifted of the sisters, initially worked with Soviet ciphers. In July 1941 she was head of a working group at the Rabo site[36]. From June 1943 she was head of group 53g, working with Soviet Naval encrypted telegrams[37]. From November 1944 she became head of group 55f, which worked on Soviet diplomatic ciphers. The group consisted of three male employees and one conscript. In May 1945, Marina is listed as “independent cryptanalyst”, where she is the sole woman [38].

Marina and her sister worked at the FRA until their retirement. Colleagues described the sisters as always having something Russian about their dress and style. Even after many years in Sweden, they still spoke Swedish with a slight Russian accent.

5.2 Bojan Frykholm

Bojan Frykholm was born in 1920. Her education was “realexamen”, which was the lower grade of Swedish secondary education at the time. She had several shorter office employments around 1940, among them at Åhlen & Holm, a Swedish department store chain today

[34] FRA-order.

[35] FRA. Personnel file for Marina Löfström.

[36] FRA. Rabo-order July 22 1941.

[37] FRA-order.

[38] FRA. List of FRA personnel 1945.

known as Åhlens. In January 1942, she was laid off from work as a typist at the Government Industrial Commission. She was referred to the FRA by the employment office, applied for a post in February and was hired as a typist. She initially worked with typing out German Geheimschreiber traffic. The fact that she did not pass her school exam in German was apparently not a problem [39].

From January 1943 she worked at group 53f, decoding Russian Naval telegrams [40]. In 1945 she is listed as codebreaking assistant. She apparently had a natural talent at solving ciphers. Her colleagues say that she had an incredible ability to solve complex transposition ciphers, and reputedly could solve double transposition ciphers by just taking a glance at the codetext [41]. Later in the 1960s she worked on encoded traffic relating to the Biafran war.

Bojan Frykholm is an example of a woman that started at the FRA more or less by accident, but once there her natural talents were recognized, and she was given the opportunity to develop them. She continued to work at the FRA until her retirement, and finished her career as a respected specialist in her field.

5.3 Eva Löfvenmark

Eva Löfvenmark was trained as a typist at the Bar Lock institute in Stockholm, a well-known school for secretaries in Stockholm at the time. She finished the course in 1937[42], and by the Bar Lock employment service she was referred to the Armed forces High command, who were looking for a secretary. She was hired, and ended up at the crypto department. At the time, it consisted of her and five officers. She was initially employed with typing up various reports and examples used in cryptologic training exercises. She took part in preparations for a number of exercises, and typed out results and conclusions. The employees were encouraged to bring any relatives or acquaintances who were interested in cryptology.

Early in the war, she worked in the group decoding French telegrams. She could well remember champagne being served on the successful breaking of a new French code. Later she worked on Soviet Naval ciphers and typing and pasting German Geheimschreiber traffic. In 1945 she worked on decoding weather observations at the Government Meteorological Office. This description of working in several different groups is fairly typical, and a good example of how employees were shifted around between various tasks.

Eva Löfvenmark describes her time at the wartime FRA as a “wonderful time”. In an interview in 1976, she said “This fantastic feeling of taking part, of doing something useful. We didn’t care that we had to work overtime, we didn’t think about that. When something needed to be done, we just did it!”[43]

10 Women at the FRA, then and now

When the FRA was created, 67% of the employees were women. Since then the proportion of women in the agency has declined steadily. Already in 1945, the proportion of women was 52%[44], and in 1949 it was down to 33%. Today the FRA has 25% female employees.

What is behind this decline? Doubtless there are several factors. During the war years, there was a tendency to use women in headquarters and staff duties in the military to free up men for frontline duty. After the war, when defence was no longer a self-evident duty for everyone, there was a reversal to more traditional roles.

Signals intelligence is a highly technical activity, today as in the 1940s. However, the part played by technology has increased more and more since the 1940s, and focused on the use of computers. Computer technology tends to be a male area, and the increased use of computers has contributed to an increase in the proportion of men at the FRA.

The technology of the 1940s demanded much manual work in the form of typing, routine decryption, compilations, statistics and other supporting jobs, that were at that time typically performed by women. The increased use of computers has automated many of these jobs and made them redundant.

[39] FRA. Personnel file, Bojan Frykholm.
[40] FRA-order.
[41] Wik.
[42] FRA. Personnel file Eva Löfvenmark.
[43] FRA. Löfenmark
[44] FRA. List of FRA personnel 1945

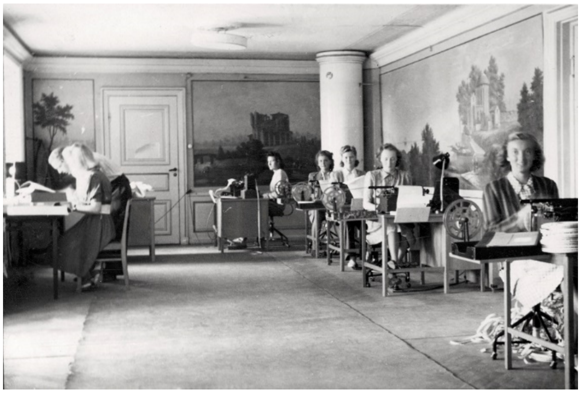


Fig. 7: Female typists at an FRA site during the war.

If we look at differences in pay, in the 1940s women were predominantly found in the lower pay grades at the FRA, even though there were exceptions. However, pay was fixed in the pay grade, and if a man and a woman had the same job, they had the same salary.

Today, wages are individually set depending on performance at work. However, when men and women are performing the same duties, men typically have higher salaries. Differences in salary between women and men remain, but the mechanisms are different.

References

Borelius, Carl Gösta c 1980, *CGB*. Memories of a mathematician who worked with the Gehemschreiber traffic. FRA archives.

FRA. *FRA-order*, FRA weekly administrative regulations 1942 - 1943. FRA archives, vol BII:1.

FRA. *Rabo-order*, weekly administrative regulations for the Rabo site, 1940-1943. FRA archives, vol B III:1.

FRA. *Krybo-order*, weekly administrative regulations for the Krybo site, 1940-1943. FRA archives vol B III:1.

FRA. *Tjänstemeddelanden* 1944-1945 (replaces FRA-order as weekly administrative regulations from 1944). FRA archives, vol BII:2.

FRA. *List of FRA personnel 1942*. Attachment to, FRA-order July 27 1942, FRA archives, vol BII:1.

FRA. *List of FRA personnel 1945*. FRA archives

FRA. 1992. *FRA 50 year anniversary memorial publication.* Interview with Maj-Britt Skoglund. FRA archives.

FRA. *PM H 153/44 Interception and decryption of foreign weather telegrams.* FRA archives.

FRA: *Receipt for the manufacture of equipment by the Swedish Cash Register Company 1943*. FRA archives.

FRA. 1942. *Psychotechnical tests I-IV*. FRA archives.

FRA. *Certificates of employment.* Outgoing correspondence 1942-1945, FRA archives.

FRA. *PM on the use of women as automatic Morse operators 1943*. FRA archives.

FRA. *Work description for personnel at FRA sites on Lidingö, 1942*. FRA archives.

FRA. *Missive to the government requesting financing of buses to FRA site, Nr 72/43*. FRA archives.

FRA. *Personnel file Marina Löfström*. FRA archives.

FRA. *Personnel file Eva Löfvenmark,* FRA archives.

FRA. *Personnel file Bojan Frykholm,* FRA archives.

FRA. *FRA annual accounts 2019*.

FRA. *Reports of accidents at the workplace.* FRA archives.

Högskoleverket (Swedish Higher Education Agency): *Study of higher education and research 1945-2005*, Högskoleverket, ISSN 1400-948X

Löfvenmark, Eva. 1976. Taped interview made by Bengt Beckman. FRA archives

Swedish government. *SOU 1946:66. Study of working conditions for government employees.* Official reports of the Swedish government.

Swedish government. *SOU 1953:18. Study of equal pay for women and men in government service*. Official reports of the Swedish government.

Wik, Anders. Former head of FRA crypto department, interviewed by the author.

Runic cryptography in early epigraphic period (200-700)

Sebastien Zimmermann
Université de Lorraine
zimmermann-sebastien@pm.me

Abstract

Runic script is an alphabetic system based on a non-alphabetical order row. The oldest row contained 24 letters but was reduced to 16 in Scandinavia around 800, and enlarged to 28 then 31 letters in England between 7th and 10th century (See Düwel, 2008). Runic script also adopted various original encryption systems during Viking and Middle Ages. Recent and very complete studies help us to understand the way they worked, both on social and technical levels. Nevertheless, the very first uses of cryptography are probably far much older and could have even occurred since runic script is first attested, that is in the end of 2nd century. It should be emphasized that it was based on a visual effect and riddles which goal was to make guess names rather than magical formulae. Indeed, the use of magical formulae is only attested lately, in the medieval times in Scandinavia and England inscriptions, and it is clearly related to Christian prayers and Kabbala (Bauer, 2020). This article aims to provide a chronology and a typology of possible cryptography techniques used with runic script in its earliest period.

1 Introduction

Due to its various particularities, and its non-standardised letters, runic script has, for long, been seen as being either originated from a cipher alphabet or being a cipher alphabet *per se*. That is, a secret alphabet. Nevertheless, as it featured many linguistic and grammatical tools since the very beginning, we can assume it was clearly created as a real script rather than a cryptographic system, *i.e.* a dissimulating and concealing technique, which doesn't need such a complicated structure. Furthermore, runic script was constantly improved on very modern technical basis of adaptability, since it is in fact a kind of open-source system updated by its users through time. In this way it is very similar to Irish ogham script. Ogham script also has a non alphabetical order but is organized around sound values, and its users also developed cipher systems with high closeness to those used for runic script around the same period, that can be found in manuscripts such as *In lebor ogaim* written between 8th and 12th century (Derolez,1952 ; 1954, pp. 146-156). During that period, concerning the reasons why cryptography was used in runic script, we acquired new insights based on the recent and exhaustive J. K. Nordby's study (2018). He has established an inventory of runic encryption systems and detailed how they worked (substitution, permutation, visual…) with an impressive catalogue. But to conclude, he underlines that use of cryptic runes was rather a cognitive process included in the learning of script with support of ciphers and riddles. Use of cryptography was also a mark of knowledge and proficiency among carvers, with a striking effect undoubtedly based on its visual role (Nordby, 2018, pp. 229-239). Actually, runic cryptography has been developed on various levels, both in England and Scandinavia, in a literate context. Hiding texts or playing with words is a worldwide and ancient activity that took many forms (Blake, 2010). Thus, the idea of mind games remains undoubtedly valuable in the time span studied here.

Therefore, relevant inscriptions have been selected here and briefly analysed to understand the chronology and typology of possible use and development of graphic effects if not cryptography, from 3rd century until attested cryptography in 8th century. We must note that in comparison, first centuries Roman carvers often used various techniques of abbreviations, ligatures and acrostics on many supports and that alphabetical learning process included order changes in letters (Mees, 2006). Furthermore, some attested encryption techniques were already employed in the Antique world (Nordby, 2018, pp.39-43) and several new cryptography systems appeared between 4th and 8th century, mainly in the Western Christian world (Nordby, 2018, pp.44-48). Another important fact is that latin script was gradually spread among high ranking and christian germanic populations in France, England and Southern Germany since 5th century (Fischer, 2005). Although these influences were limited, they could have played a role for adopting new cryptography techniques (Düwel, 1994).

2 First attempts of cryptography ?

In epigraphy, and particularly in runic epigraphy, meaningless inscriptions are often categorized as resulting from errors and illiteracy, or as magical formulae. Several inscriptions or words from the ancient period appears to be impossible to understand either because they are not wholly legible or only written with consonants. In some cases we could find ligatures or even workers abbreviations similar to Roman ones (*f* for *fecit…*) like possibly in Northern Germany on the Thorsberg bronze umbo inscription from 3rd century, **ansgzh**. Thus, *Ansgiz* would be here the smith name and aberrant ending **h** could fit as an abbreviation for PG **handuz* adapted from latin *manu*, “has made from his hands” (Imer, 2014, pp. 72-73), as this surely happens on the Femø gold bracteate (Denmark, 5th century) with **ekfakaʀf**, “I, Fakar”, ending with **f** for *fāhi*, “painted” (Nordby, 2018, p. 64). Especially, we must note that other early inscriptions, such as these from Vimose deposit (Denmark, 3rd century) are for most very puzzling compared to the contemporary deposit from Illreup. Some have only one word, mainly a name, but one of them with just consonants is read as **ttnþ** (?) on a bronze chape. Another inscription on a silver scabbard suspension fitting possibly features pseudo-runic characters. Both items are properly not luxury goods, but are related to sword equipment from high ranking and professional soldiers (Stoklund, 1995a). The Tørvika B stone (Norway, 5th-6th century) used as slab in a grave has a small and short inscription which also features ligatured or mirrored letters still not deciphered and a zigzag line as ornamentation (MacLeod, 2002, pp.117-120). These kind of para-scripts or meaningless texts continuously appeared in every regions where runic inscriptions are found, with no explainable motives for that, but as they often come with ornaments, the social context and decorative process should be considered (Graf, 2010 ; Waldispühl, 2013) rather than only the ‘magic’ explanation (Antonsen, 1980).

Legible early inscriptions sometimes contain owner names but have high proportion of maker names, thus alleging that workers were owning script and developed creative stylistic variants and designs they possibly shared with others. For instance, one lancehead of the Vimose deposit, with the name **wagnijo**, has an interesting design, known as mirrored letters, only adequately fitting with some characters (*e.g*: ᚹ, w = ᛰ mirrored). These having no practical uses, just aesthetics ones, that could be, in some ways, related to encryption. And this mirror effect is also found in other Danish deposits. Among nine Illerup items from the same period, we have two silver shield handles, with name **laguþewa** and worker signature **niþijo tawide**, “Niþijō made”, fig.1, and two lanceheads bearing the same inscription as on the Vimose one (Stoklund, 1995a). This goes on more recently, on Nydam lance shafts 8, 9 and 10, the latter, fig.2 (5th-6th century), the only legible one, with another worker signature, **tauiteka**, *tawide eka*, “I made” (Rau & Nedoma, 2014). We also find them on the Spong Hill cinerary urn (5th-6th century), with a stamped inscription **alu** (see 3.4) and on the ending letter, **a**, of the Broadley brooch (6th-7th century), with the name **liota**, both from England (Parsons, 1999, pp.46-47, 60-62). Mirrored runes are not the only particular effects from this period. Another distinctive feature comes from the Skovgårde silver fibula inscription (Denmark, 3rd century), with two separated words, **talgida:omal**, that is “Lamo made”, in opposite directions of writing which can create a step between art and encryption techniques (Stoklund, 1995b, pp. 213-214) knowing that several fibulae of Rosette type also have similar decorated patterns and inscriptions (Przybyła, 2018, pp. 29-144). A last type of very low encryption, is found on arrow shafts from Nydam deposit (Denmark, 4th century) where the word **alu** (see 3.4) is presumably written **lua**, **la** or **l** (Imer, 2014, pp. 80-81). Nearly 55 other arrows have various letters, signs or symbols that could indicate ownership (Bemmann & Bemmann, 1998, pp.416-418, fig.15). Finally, use of elided vowels could appear on the Ethelhem fibula (Sweden, 5th-6th century) which inscription has mostly consonants : **mkmrlawrtaa**, with no probing interpretation (Antonsen, 2002, p186-187 ; Imer, 2014, pp. 114, 117-118), and **rnʀ** standing for *runor*, on the Nebenstedt gold bracteate 1 (Germany, 5th century) but with no certainty (Imer, 2015, p. 187).

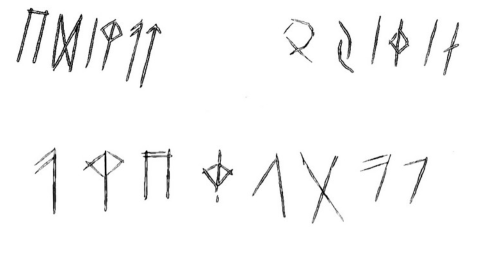


Figure 1 : mirrored runes on the Illerup silver shield handles

It is then difficult to establish when and if the use of cryptography really occurred with runic script in elder period, but we encounter here the very first evidences of abbreviations, ligatures and mirrored runes in the earliest inscriptions possibly only related to ornamentation or visual effects (MacLeod, 2002).

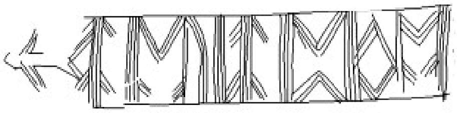


Figure 2 : mirrored runes on the Nydam lance shaft (illustr. Pia Brejnholt)

3 A mere visual effect or real cryptography ?

This chapter provides a selection of the most relevant inscriptions that could be related to the use of cryptography in order to subsequently give a chronology of its possible appearance. A first step was to define inscriptions fitting this category. The following inscriptions cannot clearly be considered as using cryptography, but they all show odd features that could not have been carved by mistake or by chance, but only with a particular purpose that must be clarified. Therefore, the possible encrypted elements are gathered by style and likeness for further studies. A noticeable fact is that the various effects appear almost all at the same period in various regions. Furthermore, we can wonder if the visual effect implied by mirrors and ligatures, decorated letters and repeated letters was not limited to ornamentation but was already suggesting riddles or mind games that gradually lead to more and more elaborated effects.

3.1 Decorated letters

At first glance not significant, this technique of carved letters, simply highlighted with two or three lines, started with the golden Gallehus horn (Denmark, 4th-5th century), but is more relevant on few other objects such as the Nydam lance shafts 9 and 10 (Denmark,5th-6th century). It worth being noticed that in this deposit, 34 other lance and arrow shafts and 3 knife handles, are decorated with interlaced motives as well as 43 lance shafts in the contemporary Kragehul deposit (Denmark, 5th-6th century), but also archery material, such as arrows, (see 2), have ornaments and scripts or para-scripts (Iversen, 2010, pp. 65-70 ; Petersen, 2020). Such a decorative technique was also applied to various runic objects of bone or wood dated around 5th-6th century, which function remains unknown and are then labelled as 'amulet' : Sorte Muld (Denmark), illegible Ødemotland (Norway) and Lindholm (Sweden) alike with two others from the Netherlands, difficult to date precisely (4th- 9th century), Britsum and Wijnaldum A (Kaiser, 2021, pp. 319-333, 395-401; Stoklund, 2005, p.362).

3.2 Meaningful sentence with meaningless words

We have here, possibly, a similar system as in the medieval period with substitution of letters (St John's College, Oxford, MS 17, 5V, ca. 1100), thus giving legible but meaningless words (Saltzman, 2018), that are found on stones which function as memorial is the more accurate. The oldest monument of this kind is the Hogganvik stone (Norway, 4th-5th century), fig. 3, with an inscription dedicated by relatives to a local leader, "Kelbaþewas stone, …, I, Naudigastiz, I Erafaz" : **kelbaþewas s[t]ainaz | aaasrpkf aarpaa inananaboz | ek naudigastiz | ek erafaz**. The two meaningless words **aaasrpkf aarpaa** are seen as a magical formula (Schulte, 2013) or at least as a coded text (Knirk, 2011). Here the repeated letter **a** and the use of the letter **p**, very scarce in the Nordic language, is of course the most puzzling and precisely indicates an unusual language.

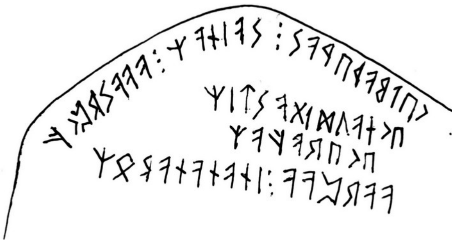


Figure 3 : Hogganvik stone

The same process occurs on the second monument, the Noleby stone (Sweden, 6th-7th century), fig.4, with a short alliterated dedication : **runofahiraginakudotojea | unaþou: suhurah: susih -atin | hakuþo**. Here the second sentence divided in 3 parts by separation marks with aberrant ending **h**, is seen as a magical formula since the preceding and last sentences in *scriptio continua* are translated as "I paint the rune provided by the powers, I prepare… Hakuþo" and thus interpreted as some kind of prayer or curse possibly similar to the textual content of the contemporary Blekinge stones (Hellstam, 2014). Nevertheless, the social context should be considered as the most important and key factor for the inscriptions (Antonsen, 2002, pp.180-185 ; Imer, 2015, pp.191 ; Marold, 2012, pp. 83-84). The coded part could explain the reason why the stone was erected.

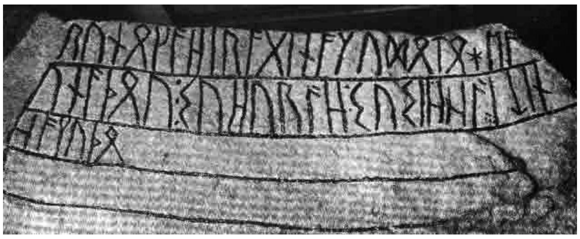


Figure 4 : Noleby stone

3.3 Alphabet with meaningless text or symbols

In early Christians times, Roman alphabet was, on occasions, inscribed as a consecration on churches floor or on grave stones. Eastern germanic populations were christianized since 4th century as could indicate the Breza inscriptions (Bosnia, 6th century). A latin alphabet is carved on a column from what was probably a church and on a second column is carved a runic alphabet with a five pointed star below (Looijenga, 2003, pp.50-62, 232-234). In Greek alphabet, we are aware that first and final letters, A and Ω, have a deep significance with Christ himself. Does runic alphabet refer to a similar concept when it is carved on memorial stones or jewellery ? The oldest evidence of both alphabet carved with other meaningless words is the Kylver stone (Sweden, 4th-5th century), fig. 5, used as a slab but which could be perceived as a learning tool rather than a memorial (Antonsen, 2002, pp.176-179). Furthermore it features a meaningless palindrome with variant **s** letters and a tree shaped letter or symbol after the alphabet : **fuþarkgwhnijpïʀstbemlŋdo sueus** (Scholma-Mason, 2016). On the Rök stone (see 4), a different tree shape fits for letter **þ** on line 21 (Nordby, 2018, pp.199-202).

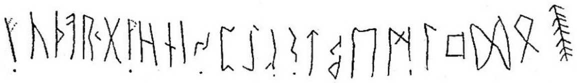


Figure 5 : Kylver stone

The Vadstena gold bracteate (Sweden, early 500) is very particular because it is considered as the first evidence of 3 separated rows of 8 letters thus allowing binary substitutions similar to those explained in the *Isruna tract* (see 5). These pendant types are luxurious objects worn by rich women. Some feature a male figure riding a horse, inspired from Roman coins and medals iconography. Of the one thousand found only two hundreds have inscriptions, some being wholly legible, but most are totally meaningless. However it is not related to encryption but rather to problem with reproduction of text models. Here the inscription reads : **tuwatuwa· fuþarkgw· hnijïpʀs· tbemlŋo** [**d**]. The repeated word *tuwa* is still unexplained (Imer, 2014, pp.107-110).

Another kind of luxurious jewel, the silver fibula, is a highly decorated typical female object but instead of visible inscriptions as on bracteates, those on fibulae are carved on the back side and unseen. The Acquicum fibula (Hungary, 6th century) has an incomplete alphabet (**fuþarkgw**), maybe indicating the first row, and the naming of the object, **klain kiŋia**, that is "fine fibula" (Düwel et al., 2020, pp. 9-19). On the Charnay fibula (France, 6th century) we have an incomplete alphabet with separated and rather unclear words : **fuþarkgwhnijïpzstblem : uþfnþai:id dan : liano**. The possibility of a dedication with two names (*Iddan* and *Liano*) is retained, but with no certainty (Antonsen, 2002, pp.152-153, 179). Another particular case is included here with the Fonnås fibula (Norway, 6th century) which inscription is somewhat similar to Charnay but without alphabet. It is not clear if some of the characters are variants used as encryption, however the inscription remains undeciphered : **wh : b/widulti wkhu Alklʀ þArbe : iAʀ** (Birkmann, 1995, pp. 87-89). Finally, the Beuchte fibula (Germany, 6th century), fig. 6, has no real textual problems with an incomplete and incorrect alphabet row (**fuþarzj**) and a name (**buirso**), but it features an intriguing and quite unusual ornament of repeated geometric patterns with no other attested finds (Düwel et al., 2020, pp. 76-85 ; Waldispühl, 2013, pp.260-261).

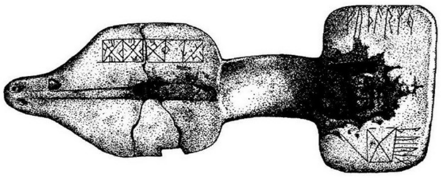


Figure 6 : Beuchte silver fibula

The following inscription is only mentioned for the record and for its similarities and continuity with previous ones, despite its time period is overlapping our selected chronology. The Malt stone (Denmark, 800-900) is one of the first stone inscriptions with the brand new 16 letters alphabet (**fuþarkhniastbmlʀ**), a tree shape symbol and human face with X on the forehead. The text, mentioning "runes of gladness and runes of eternal friendship" (**taitirunaʀ u-aivinrunaʀ**) is supposed to be related to magic but knowing "runes" means "letters" or "text", it could be a riddle or some kind of mind game as well, with meaningless words (**titultitul**) reminding the Vadstena **tuwatuwa**. (Birkmann, 1995, pp.361-372),

3.4 Repeated letters with use of ligatures

Two very different artefacts should retain our attention here because of their particular likeness in the process and in the word used as a code or just as a visual and sound effect with rhythm and alliterations (Marold, 2012, pp. 79-80) but with no apparent meaning (MacLeod, 2002, pp. 105-112 ; MacLeod & Mees, 2001). The first is a weapon, the Kragehul lance shaft (Denmark, 5th century), fig. 7, with a long inscription of decorated letters, still unclear and debated, but naming the **erilaʀ**. It is possibly the early word, only used in Migration Period, for Old Norse *jarl* and Old English *earl*. It would be then a high ranking chief warrior. However, the inscription, **ek erilaʀ asugisalas muha haite gagaga ginu gahe … lija … hagala wiju big**, has long been asserted as being a magical formula with the three times repeated **ga** letters decorated with a ligature forming a X shaped cross since X is rendering letter **g** (Antonsen, 2002, pp. 230-231 ; Imer, 2011, 2014, pp. 113-114 ; Iversen, 2010, pp. 68-69 ; Parsons, 1999, pp.18-19). And we can find exactly the same process on a jewel from England with typical Roman iconography including Romulus and Remus, the Undley gold bracteate (5th century) : **gagãga • maga • medu**. The last word, could mean "mead" and be related to a similar Old Nordic word found only in older inscriptions and often seen on the gold bracteates, **alu**, probably meaning *ale* and considered as a "magical word" (Parsons, 1999, pp.18-19, 62-67).

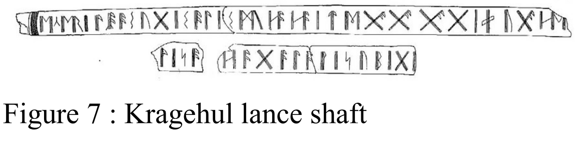


Figure 7 : Kragehul lance shaft

3.5 Meaningless repeated letters

On the Lindholm bone "amulet" (Sweden, 5th-6th century), fig. 8, we have the word **erilaʀ** mentioned again, with his name, *Sawilagar*, and there is a quite innovative feature in this unclear inscription carved on three sides also with decorated letters : **ekerilaʀsa**[**wil**]**agaʀ hateka** | **aaaaaaaaʀʀʀnn**[**n**?][**b**]**muttt alu**. The word **alu** (see 3.4) concluding a series of repeated letters **a**, **ʀ**, **n** and **t** is obviously assumed to point towards the idea of a magical formula. Instead, these letters could be arranged to mean something strictly related to the owner, the high status **erilaʀ** (Antonsen, 2002, pp. 187-188 ; Imer, 2014, pp. 113-115). That is probably the case on the Chessel Down I brass bucket (England, 6th century), an object originating from Byzantine workshops and found in a rich woman grave with a rather short but confusing inscription : **bwseeekkkaaa**. Here, the repeated letters added to the three first letters could be a reminding of Scandinavian personal pronoun, "I", *eka*, or could give three attested Old English names (*Becca*, *Wecca*, *Secca*), maybe relatives. It would work with a system so far considered as being imported from Scandinavia, the -*istil* code (Nordby, 2018, p.104-111), first appearing on the Gørlev memorial stone (Denmark, 9th century) and found until 12th century in epigraphy on stone monuments and various objects including graffito on churches. And it is even known from a 14th century Icelandic tale called *Bósa saga* where it concludes a curse chanted by a witch. This code is based on repeated letters, **þmk:iii:sss:ttt:iii:lll**, creating the suffix -*istill* added to each first three letters, **þmk**. Thus, these give the three following Old Norse words : *þistill*, *mistill*, *kistill*, that is "thistle, mistletoe, coffin". A riddle that echoes with the *horn*, *þorn*, *korn* formula from the Gotland Island (Birkmann, 1995, pp.356-360 ; Hines, 1991 ; Looijenga, 2003, pp.280-281 ; Parsons, 1999, pp.51-52)

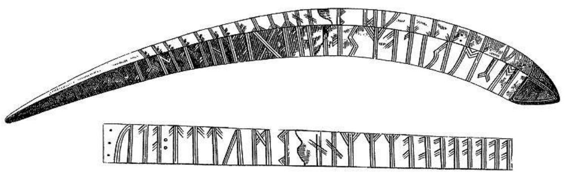


Figure 8 : Lindholm bone "amulet"

Although it is dispatched in this category, the Ellestad stone (Sweden, 6th-8th century) seems a bit different from the others. The inscriptions is carved on a memorial stone and the lower line is ending with groups of letters ᛉ, **k** and ᛁ, **i : ekA sigimArAʀAfs… kA rAisidokA |stAinA | kk.. kiiii kkk…** , “I Sigimar… raised the stone” (Imer, 2011 ; 2015, p.59). These repeated letters are looking like those on sticks from Bryggen (Norway, 1170-1250) (Nordby, 2018, pp.318-319, 332, 337-342) and from Trondheim (Norway, 1175–1275) (Nordby, 2018, pp.362-363). This kind of sequence could be a substitution system (Nordby, 2018, pp.72-73, 158-163) such as those suggested in old literature in Jón Ólafsson’s *Runologia* (1752, pp.160-166) and Joh Liljegren’s *die nordischen Runen* (1848, pp. 34-40). It is possibly the carver name.

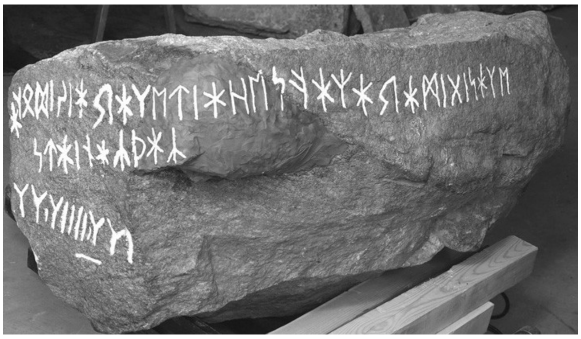


Figure 9 : Ellestad stone

3.6 Only one repeated letter

Two female objects have a particular letter repeated several times either at the beginning or at the end of the inscription. The Gjersvik bone scraper (Norway, 6th century) has an incomplete and still not translated inscription which ends with letter ᛚ, **l** : **d--fioþillllllllllll**. This letter could be related to the function of the object, used to crush fibres of flax, as this word translates as *lina* in Old Germanic and also appears on a similar object from Fløksand (Norway, 4th century) as **linalaukaR**. On the opposite, the letter on the Skabersjö bronze fibula (Sweden, 7th-8th century) has the letter ᛦ, **ʀ**, at the beginning of the inscription : **ʀʀʀʀʀʀʀʀʀʀʀʀʀʀʀʀ raþi tuk fauka fiaʀ sis in a iak asu þui launat | … auab-ksuafakat**, possibly meaning “Hráði took (fauka?) from his wealth, but therewith have I rewarded Ása”. But this letter should be used to conclude words not to start them. This reminds the famous golden ring found in Danish archbishop Absalon’s grave and dated around 11th-12th century, with inscription **þorKair**, *Þorger*, followed by five letters ᛦ, each inside brackets (Imer, 2011 ; 2015, pp. 94, 232).

3.7 Organized crossed letters

Reminding more of other contemporary monograms or maybe christograms (Hilberg, 2000 ; Schwab, 1998) than of cryptography, this system with four or five letters organized and crossed around a ‘X’ shape, or maybe letter ᚷ, **g**, is only found on two inscriptions from Germany (6th century) but was used more recently in Sweden on Södermanland Christian stones with more letters (Bianchi, 2010, pp. 142-143; MacLeod, 2002, pp. 165-166). The first one is related to a high ranking warrior, the Schretzeim IV silver sword ring (Düwel et al, 2020, pp.574-580 ; Waldispühl, 2013, pp. 304-305) and the second one, the Soest gold fibula comes from a rich woman grave (fig.10), but includes another inscription with two female names (Düwel et al, 2020, pp.587-596 ; Graf, 2010, pp.114-119 ; MacLeod, 2002, pp. 101-105 ; Waldispühl, 2013, pp. 153-171, 306-309). It is possibly the carver name or its signature.

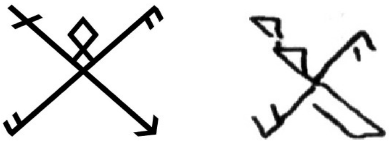


Figure 10 : Soest and Schretzeim inscriptionsʀ

4 Attested cryptography around 8th century

The very first evidences of cryptography emerged in England when literacy is promoted by Christianity and manuscripts. It is also linked to art and high technical level of craft. In this way, the Auzon casket (8th century), is a very unique object, combining runic script with Old English and Latin languages as well as Roman letters. The whole object is then conveying mixed cultural influences, narratives and images, enhanced by riddles and mind games, even codes with letters substitution (Page, 1999, pp. 87-88 ; Schwab, 2008, pp.73-75 ; Webster, 2012).

The Hackness cross (8th-9th century) is, of course, a Christian monument, with a Latin inscription dedicated to abbess Ethelburga, and both incomplete and worn runic inscription followed by a cryptic text using the *hahal-rune* system, which makes runes looking like tree shapes (Page, 1999, pp. 83-86). Again, we have cross-cultural techniques and scripts designed for highly literate people.

At these times, runic cryptography was preserved through manuscripts written by clergymen and copied by monks as language investigations (Zironi, 2011). The *Isruna tract* is the first tutorial with three runic codes and binary substitutions found in several European manuscripts (9th-11th century) which are closely related to the whole literacy of this period (Derolez,1954, pp. 120-169) including riddles in Old English as well as in Scandinavian so-called runic poems (Bauer, 2003; Halsall, 1981) and various riddles based on runic letters like in Exeter book (*Codex Exoniensis*, MS 3501, 8th-10th century) or poetry like with Cynewulf's signatures (Symons, 2016).

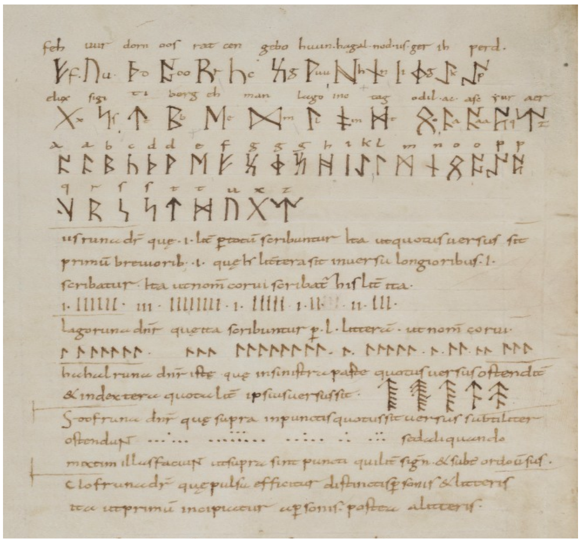


Figure 11 : *Isruna tract* fom St Gallen codex

Since manuscripts and Christianity expanded more recently in Scandinavia, we have here only monumental evidences of cryptography. As we have seen, 9th century stones offer interesting texts in Denmark, but the most striking evidence of cryptography comes from Sweden, with the famous Rök stone, which text is probably an epitaph. A question that could be answered very soon since a new hypothesis about it was proposed recently. And if it is still not offering the complete solution, however it is very close to it (Holmberg et al., 2020). The whole text, one of the longest with more than 760 letters carved on every sides, is made of riddles and cryptic runes. Here, even the older runic script is used for cryptography since the main text is carved with the brand new 16 runes row (Nordby, 2018, pp. 222-228). This indicates that the stone carver was a high skilled worker and highly literate. On a mere scriptural level, this monument enlightens us on the variability of concomitant letters, be they old or new, and the relevance of several encryption techniques.

5 Conclusion

The context of creation of runic script is presumably related to the Roman world. Since the 2nd century, qualified Germanic people have been serving in Roman army as spies (*exploratores* or *speculatores*). Thus, they should have been familiar with encryption and dissimulation of sensitive informations (Austin & Bankov, 2014, p. 191), and obviously with latin script and techniques (abbreviations, ligatures...) employed in military or civilian workshops. A knowledge that some could have shared with other people from free *Germania*. These combined techniques were possibly used in war times during conflicts in the Danubian regions and subsequently applied to the creation of the runic system so that such a script couldn't be understood by foreigners and especially Roman troops (Lund Hansen, 2003).

That is why some kind of cryptography could have been developed and enhanced since 3rd century. Nevertheless, concerning the previously selected inscriptions, innovative encryption techniques would probably start in 4th-5th century and only hide worker name in most cases and not magical or sensitive matters. Valuable objects or monuments related to high ranking people are mostly concerned and it is to be noticed that inscriptions are continuously improved as well as ornaments (Bianchi, 2010, pp. 115-164).

Although high status people, both male and female, are in some ways involved with script, this one clearly remains in the hands of craftsmen who modified and adapted it to language evolution. As the same techniques of cryptography appear at the same periods and at different locations, it could means acculturation and exchanges (Lebecq, 1997 ; Moore, 2016) rather than separated development as would prove spread of 8th and 9th century cryptography. Moreover, various goods and art styles were circulating in large areas including Scandinavia and Eastern regions since the 3rd century (Lund Hansen, 1994). And the same could be true for workers and traders who were in some cases highly mobile even before Viking period (Beghelli, 2022). However, concerning the oldest inscriptions, and with no parallels, it still remains difficult to evaluate precisely the use of cryptography and to establish which system was employed. Despite various elements available both for a cautious typology and chronology provided here, further studies will be necessary to understand and decipher potential encryptions contained therein. But to achieve this, it would be essential to consider both visual and sound effects provided by the craftsmen of ancient societies still based on orality, remaining very different from our modern views.

References

Antonsen Elmer H., On the Typology of the Older Runic Inscriptions, in : *Scandinavian Studies*, Vol. 52, No. 1 (Winter 1980), pp. 1-15, University of Illinois Press, 1980

Antonsen, Elmer H., *Runes and Germanic Linguistics,* Berlin/New York, de Gruyter, 2002

Austin N.J.E. & Rankov N.B., *Exploratio, Military and political intelligence in the Roman world from the Second Punic War to the battle of Adrianople*, Routledge, London, 2014

Beghelli Michelle, Artisanal mobility, artisanal sedantism, and the economic context. Some examples from the 7th-9th centuries, in : *Un monde en mouvement. La circulation des personnes, des biens et des idées à l'époque mérovingienne*, pp.133-148, AFAM, Saint-Germain-en-Laye, 2022

Bemmann Güde & Bemmann Jan, *Der Opferplatz von Nydam*, Bd 1 & 2, Wachholtz, Neumünster, 1998

Blake Barry, *Secret Language*, Oxford University Press, Oxford, 2010

Bauer Alessia, *Runengedichte, Texte, Untersschungen und Kommentare zur gesamten Überlieferung*, Studia Medievalia Septentrionalia, 9, Fassbaeander, Wien, 2003

Bauer Alessia, Biblical Magic as a Manifestation of Folk Belief in the North, in : *Faith and Knowledge in Late Medieval and Early Modern Scandinavia*, pp. 269–296, Brepols, Turnhout, 2020

Bianchi Marco, *Runor som Resurs. Vikingatida skriftkultur i Uppland och Södermanland*, Runrön, 20, Uppsala Universitet, 2010

Birkmann Thomas, *Von Agedal bis Malt, Die skandinavischen Runeninschriften vom Ende des 5. bis Ende des 9. Jahrhunderts*, de Gruyter, Berlin/New-York, 1995

Derolez René, Richtingen in de Runenkunde, met enkele beschouwingen over het probleem : Ogam-Runen. In: *Revue belge de philologie et d'histoire*, tome 30, fasc. 1-2, pp. 5-49, 1952

Derolez René, *Runica Manuscripta, The english tradition*, De Tempel, Brugge, 1954

Düwel Klaus, Runische und lateinische Epigraphik im süddeuteschen Raum zur Merowingerzeit, in : *Runische Schriftkultur in kontinental-skandinavischer und -angelsächsischer Wechselbeziehung*, pp. 229-308, de Gruyter, Berlin/New-York, 1994

Düwel Klaus, *Runenkunde*, Metzler, Stuttgart, 2008

Düwel Klaus, Nedoma Robert & Oehrl Sigmund, *Die südgermanischen Runeninschriften*, de Gruyter, Berlin/New York, 2020

Fischer Svante, *Roman Imperialism and Runic Literacy. The Westernization of Northern Europe (150-800 AD)*, Uppsala University, Uppsala, 2005

Graf Martin Hannes, *Paraschriftliche Zeichen in Südgermanischen Runeninschriften*, Chronos, Zürich, 2010

Hilberg Volker, Monogrammverwendung und Schriftlichkeit im merowingischen Frankenreich, in: *Arbeiten aus dem Marburger Hilfswissenschaftlichen Institut*, Elementa diplomatica 8, p. 63-122, Institut für historische Hilfswissenschaften, Marburg an der Lahn, 2000

Halsall Maureen, *The Old English Rune Poem : a critical edition*, University of Toronto Press, 1981

Hellstam Antonia, *Fyra blekingska runstenar i social kontext*, Lunds Universitet, 2014

Hines John, Some observations on the runic inscriptions of early Anglo-Saxon England, *Old English runes and their Continental Background*, pp.61-83, Bammesberger, Heidelberg, 1991

Holmberg Per, Gräslund Bo, Sundqvist Olof & Williams Henrik, The Rök Runestone and the End of the World, *Futhark: International Journal of Runic Studies,* 9–10 (2018–2019), pp. 7-38, 2020

Imer, Lisbeth M., The oldest runic Monuments in the North, Dating and Distribution, in : Language and Literacy in Early Scandinavia and Beyond, *NOWELE*, 62/63, pp.169-212, University of Southern Denmark, Odense, 2011

Imer Lisbeth M., *Jernalderens runeindskrifter i Norden*, Aarbøger for nordisk Oldkyndighed og Historie, 2013 & 2014, Det Kongelige Nordiske Oldskriftselskab, København, 2014 & 2015

Iversen Rasmus Birch, *Kragehul Mose, Ein Kriegbeuteopfer auf Südwestfünen*, Jysk Arkaelogisk Selskab, Moesgård, 2010

Kaiser Livia, *Runes across the North Sea from the Migration period and beyond*, De Gruyter, Berlin/Boston, 2021

Knirk James E., Hogganvik-innskriften: en hard runologisk nøtt, *Viking* LXXIV, pp.25-40, Oslo, 2011

Lebecq Stéphane, Routes of change : Production and distribution in the West (5th-8th century), in : *The Transformation of the Roman World, AD 400-900*, pp. 67-78, University of California Press, Berkeley/Los angeles, 1997

Looijenga Tineke, *Texts and Contexts of the Oldest Runic Inscriptions*, Brill, Leiden/Boston, 2003

Lund Hansen Ulla, Skandinavien und der Kontinent zur Völkerwanderungs- und Merowingerzeit, in : *Runische Schriftkultur in kontinental-skandinavischer und -angelsächsischer Wechselbeziehung*, pp. 1-9, de Gruyter, Berlin/New-York, 1994

Lund Hansen Ulla, Die ersten Rünen, in : *Runica-Germanica-Mediaevalia*, pp.394-398, Berlin/New York, De Gruyter, 2003

MacLeod Mindy, *Bind-Runes, An Investigation of Ligatures in Runic Epigraphy*, Runrön, 15, Uppsala Universitet, Uppsala, 2002

MacLeod Mindy & Mees Bernard, The Triple Binds of Kragehul and Undley, *NOWELE*, 38, Issue 1, p. 17-35, Odense University Press, Odense, 2001

Marold Edith, Vers oder nicht Vers? Zum metrischen Charakter von Runeninschriften im älteren Futhark, *Futhark*, Vol. 2 (2011), pp. 63-102, University of Oslo/Uppsala University, 2012

Mees Bernard, Runes in the First Century, in : *Runes and their Secrets, Studies in Runology*, pp. 201-231, Museum Tusculanum Press, København, 2006

Moore Tom, Britain, Gaul, and Germany: Cultural Interactions, in : *The Oxford Handbook of Roman Britain*, pp.262-284, Oxford University Press, 2016

Nordby K. Jonas, *Lønnruner, Kryptografi i runeinnskrifter fra vikingtid og middelalder*, Universitetet i Oslo, 2018

Page, R. I., *An Introduction to English Runes*, Boydell Press, Woodbridge, 1999

Parsons David N., *Recasting the Runes, The Reform of the Anglo-Saxon Futhorc*, Uppsala Universitet, 1999

Petersen Peter Vang, Entrelac-ornamented wooden shafts and handles from Nydam, in : *Excavating Nydam*, pp. 269-302, University Press of Southern Denmark, Copenhagen, 2020

Przybyła Marzena J., *Pressblechverzierte spätkaiserzeitliche Trachtbestandteile in Südskandinavien*, Nordiske Fortidsminder series B 28, University Press of Southern Denmark, Copenhagen, 2018

Rau Andreas & Nedoma Robert, Eine Herstellerinschrift in Zierrunen auf einem Holzschaft aus dem Moor von Nydam, in : *die Sprache*, Bd. 50,1, pp. 63-82, Wiener Sprachgesellschaft, Wien, 2014

Saltzman Benjamin A., Vthkskdkxt : Early Medieval Cryptography, Textual Errors,and Scribal Agency, *Speculum*, October 2018, Vol. 93, No. 4, pp. 975-1009, The University of Chicago Press on behalf of the Medieval Academy of America, 2018

Scholma-Mason Nela, Shedding Light on the Kylver Slab, *Saga-Book*, Vol. 40 (2016), pp. 43-55, Viking Society for Northern Research, 2016

Schulte Michael, The Norwegian Hogganvik Stone as an Emblem of Social Status and Identity, in: Across the Sólundarhaf: Connections between Scotland and the Nordic World, *Journal of the North Atlantic*, Special Vol. 4, pp.120-128, 2013

Schwab Ute, Runen der Merowingerzeit als Quelle für das Weiterleben der spätantiken christlichen und nichtchristlichen Schriftmagie?, in : *Runeninschriften als Quellen interdisziplinärer Forschung*, pp. 377–433, de Gruyter, Berlin/New York, 1998

Schwab Ute, *Franks Casket, Fünf Studien zum Runenkätschen von Auzon*, Studia Medievalia Septentrionalia, 15, Fassbaeander, Wien, 2008

Stoklund Marie, Die Runen der römischen Kaiserzeit - in: *Himlingøje - Seeland - Europa*, pp. 317-346, Det Kongelige Nordiske Oldskriftselskab, København, 1995a

Stoklund Marie, Neue Runeninschriften um etwa 200 n. Chr. aus Dänemark: Sprachliche Gliederung und archäologische Provenienz, in : *Nordwestgermanisch*, pp.205-222, De Gruyter, Berlin/New-York, 1995b

Stoklund Marie, Chronology and Typology of the Danish Runic Inscriptions, in : *Runes and their Secrets,* Studies *in Runology*, pp. 355-384, Museum Tusculanum Press, København, 2005

Symons Victoria, *Runes and Roman Letters in Anglo-Saxon Manuscripts*, de Gruyter, Berlin/Boston, 2016

Waldispühl Michelle, *Schreibpraktiken und Schriftwissen in südgermanischen Runenschriften, zur Funktionalität epigraphischer Schriftverwendung*, Chronos, Zürich, 2013

Webster Leslie, *The Franks Casket*, British Museum, London, 2012

Zironi Alessandro, Marginal Alphabets in the Carolingian Age: Philological and Codicological Considerations, in : *Rethinking and Recontextualizing Glosses. New Perspectives in the Study of Late Anglo-Saxon Glossography*, pp. 353-373, Federation Intern. des Inst. d'Etudes Médiévales, Porto, 2011

Poster Abstracts

Simulating Cryptologic History

Teaching the history of cryptology depends on access to extant devices, not just their documentation and interpretation. Yet the devices that show major steps in the history of cryptology are rare. Replicas, emulators, and simulations are pivotal for the history of cryptology.

Peter Krapp
University of California, Irvine
krapp@uci.edu

The first volume of the Venetian ciphers written by Agostino Amadi

The secret manuscript of Agostino Amadi contains hundreds of cipher methods and examples. The total manuscript in ten volumes is translated from old Italian into English and published for the first time since 1588. Some examples from the first volume are presented here.

David Scheers
dscheers@webpoint.nl

Published by:

NEALT Proceedings Series No. 50
ISSN: 1736-8197 (print), 1736-6305 (online)
ISBN 978-91-8075-134-6 (PDF)
https://doi.org/10.3384/ecp195
Linköping Electronic Conference Proceedings No. 195
ISSN 1650-3686 (print), 1650-3740 (online)
NEALT Proceedings Series No. 50
ISSN 1736-8197 (print), 1736-6305 (online)
Cover images: Fritz Menzer - A Secret Life &
3D-Cipher, Deutsches Museum