Key Design in the Early Modern Era in Europe

Beáta Megyesi and Crina Tudor
Dept. of Linguistics and Philology
Uppsala University
Sweden
first.last@lingfil.uu.se

Benedek Láng and Anna Lehofer
Budapest University of Technology
and Economics, and ELTE
Hungary
(lang|lehofer.anna)@filozofia.bme.hu

Abstract

We present an empirical study on historical keys in their original form from Early Modern Times (1400-1800) in Europe. We describe the internal structure of keys, and specify what was encoded and how. We present some trends of the construction of historical keys over time. Some of these trends have been sensed but never systematically documented by crypto historians, some other trends however are revealed here for the first time.

1 Introduction

Many studies in historical cryptology have been published on the cryptanalysis of single ciphers but systematic studies on the development of ciphers and the way encryption was carried out are rather few. Studying a large number of keys from various time periods and geographic areas gives insights into the evolution of encryption. To study original keys over time in a systematic way requires a significantly large sampled set of original keys, collected from archives and libraries. Large-scale studies have not been possible due to the lack of infrastructural resources and tools for historical cryptology.

The DECODE database (Megyesi et al., 2019) developed recently for the collection of historical ciphertexts and keys contains over 1 000 keys, of which ca 41% have been transcribed with publicly available transcriptions at the time of writing. The transcribed keys allow us to carry out large, quantitative studies to investigate and compare the internal structure of keys.

Relying on materials published by other scholars and on the basis of the DECODE collection containing many different types of keys, in many languages and from various European territories, we provide some insight into the evolution of encryption, describe some trends, along with a structural description of keys to present their typology.

The study described in this article seeks to get insights into answers to the following research questions:

- What types of keys were used in Europe between the 15th and 18th centuries? What were their specific characteristics?
- What was encoded and how?
- How did encryption evolve over time?
- Can we apply simple statistical methods to large-scale analysis of transcribed historical keys?

We focus on original keys from the Early Modern times, ca 1400-1800 found in European archives and libraries.

We start with an overview of previous studies on encryption methods with the main focus on key structure and an overview of the morphology of keys. We continue with a description of the data collection used in our study and the automatic structural description of keys. Then, in Section 5, we present results about what is encoded in keys and how, and describe some trends in key design over the centuries. Lastly, we discuss some issues and conclude our findings.

2 Historical Cipher Keys

In classic cryptography, a key defines the transformation of the plaintext units (characters, words, phrases, etc) into ciphertext to encrypt the plaintext message, and vice versa, to decrypt ciphertext. The plaintext units are replaced with a code as specified by the key. The code can be represented by symbols from alphabetic characters and digits to many kinds of graphic signs.

While large-scale systematic studies on historical keys are missing, we can find a few late

19th and mid 20th century text editions of cipher keys that did not go beyond simply publishing the tables, see e.g. Rockinger (1892) and Devos (1950). The most well-known studies on keys were performed by Aloys Meister in the beginning of the twentieth century, who first offered systematic analyses of this kind of source. In two volumes, he focused on the cipher system of the Vatican (Meister, 1906) [p. 69], and other Italian city-states (Meister, 1902)), not only publishing, but also classifying the keys. Meister collected keys from the 14th to the 17th centuries from various archives in the Vatican and identified 12 types of keys using digits, and described an advanced system of cryptography carried out by professionals involving training in both the creation and the cryptanalysis of ciphers. Meister focused on keys and did not publish ciphertexts so we cannot draw any conclusion from the actual usage of keys.

The Vatican ciphers were revisited in a recent study (Lasry et al., 2020) aiming at the decryption of ciphertexts and the recovering of keys, originated from the papal correspondence in European countries between the 16th and the 18th century. The study gave unique insights into papal cryptographic practices and showed that in the 16th century, and in accordance with Meister's study, there is strong evidence for diversity, innovation, and sophistication in the development and use of (papal) cipher methods and keys. The cipher types from that period include simple (one plaintext entity – one code), homophonic (one plaintext entity – several codes), and polyphonic (several plaintext entity – one code) substitutions with or without nomenclature elements, i.e. codewords, the cipher equivalents of proper names, geographical entities, common words, etc. Most of the homophonic ciphers use variable length codes for various plaintext entities, making codebreaking much harder. In the 17th and 18th centuries, on the other hand, shorter or longer nomenclatures were standard and the ciphers were homophonic with codes of fixed-length, thus easier to use, but also easier to break, allowing deterministic parsing and decoding.

To our knowledge, the only study that systematically described early modern code keys was carried out by David Kahn published in his famous *Codebreakers* (Kahn, 1996).

Not to mention here a great number of useful case studies published in the following half century (more often than not in the journal *Cryptologia*) that did not exceed local relevance, in 2018 Benedek Láng (Láng, 2018) chose a fairly large, but still limited territorial scope, that of East-Central Europe. On the source material of this territory, he carried out a systematic analysis. He mapped the many small steps stages through which monoalphabetic ciphers evolved first into large homophonic systems, which finally gave the floor to code-booklets. On this rich, but geographically well defined area, he managed to match cipher keys with the corresponding encrypted documents. In this matching process, such structural features as we present in this article, were of great help.

To quote David Kahn again, he emphasised first that a systematic research is to be done in the historical evolution of nomenclators. Note that Kahn uses the word nomenclator in a more general sense than we defined nomanclatures above: he refers to the whole cipher key. Kahn writes:

> "At first, the substitution symbols were neither letters or numbers but fanciful signs like % or . But nobody has looked into when, in the later evolution, as nomenclators ran out of easily distinguishable symbols and began using numbers, the cipher secretaries began forming two-part nomenclators. This research requires merely examining the many nomenclators in the archives of Italy and France and timing and quantifying the change. I suppose it will be tough, living in Europe for a year and having an aperitif after a day examining antique manuscripts. But somebody should do it!" (Kahn, 2008) [p.58].

And this is exactly what the authors of this paper are up to.

3 The Morphology of Keys

A key defines how each entity in the original plaintext shall be encrypted. Keys contain a mapping between the plaintext entities and their corresponding codes used for encryption. There are some basic elements in historical keys that can be structurally described. We introduce the term "morphology" to describe the form and structure of keys with respect to codes and their corresponding plaintext entities.

Entities that can be encrypted range from characters in the plaintext alphabet and space to hide

word boundaries, to nomenclature elements that are plaintext entities with two or several characters, such as syllables, morphemes, common words, and/or named entities, typically referring to persons, geographic areas, or dates. Punctuation marks or capital letters might also occur in keys while diacritics are often not encoded. A key might also contain nulls, i.e. symbols without any corresponding plaintext characters to confuse the cryptanalyst and make decryption even harder.

Each type of entity to be encrypted might be encoded by one symbol only, two symbols, three symbols, and so on. The codes in a key might be of fixed or of variable length. For example, one key might contain only two-digit codes while another key might contain two-digit numbers for the encryption of the characters in the plaintext alphabet, three-digit numbers used for the nomenclature elements, one-digit numbers for space, and four-digit numbers for the nulls. To make decryption difficult, the most frequently occurring plaintext characters in a language might have several corresponding codes.

Figure 1 illustrates a key based on homophonic substitution with nomenclature from the second half of the 17th century. Each letter in the alphabet has at least one ciphertext symbol represented as a two-digit number or a symbol, and the vowels and double consonants have one additional graphical sign (e.g. A – 18, m; B – 20; C – 19). The key also contains encoded syllables with two-digit numbers or bigram characters (e.g. BA – 65; BE – 66), followed by a nomenclature in the form of a list of Spanish words encoded with three-digit numbers or symbols (e.g. ajustiamento – 106).

Given a transcribed key, we can automatically derive the key's morphological structure. Next, we describe our method for the empirical study on historical keys using computational methods.

4 Analysing Keys

4.1 Key Collection

Finding original keys in archives and libraries is a time-consuming and frustrating endeavor as these manuscripts are rarely indexed as keys. The DECODE database (Megyesi et al., 2019) provides a collection of encryption keys with information about their origin and other relevant documents. At the time of writing, the database contains over 1 116 original cipher keys originating from the 15th to the 18th centuries. They have been collected in libraries from European countries, mainly from Austria, Belgium, Germany, Hungary, Italy, the Netherlands, the UK, and the Vatican. 41% of the keys have been manually transcribed, following the transcription guidelines developed for historical ciphers (Megyesi, 2020). The distribution of keys throughout the centuries in this study is shown in Figure 2.

The digitized and transcribed cipher keys allow us to make large-scale studies of the morphology of ciphers, and make comparisons across time periods, geographic areas, and other information of interest. In order for our analysis to be as accurate as possible, we must first establish a transcription standard. This way, we ensure a stable and uniform basis to provide a reliable comparison across keys.

Our method makes use of plain text files (".txt") containing the transcription of the original key document. The transcription replicates the original document as closely as possible, both in terms of its structure as well as its content. In large terms, we follow the same guidelines (Megyesi, 2020) as those used in the DECODE database (Megyesi et al., 2019), and expand on them in order to adapt to the specific key structure.

Next, we describe the automatic process of the structural description of keys.

4.2 Automatic Structural Description of Keys

We provide automatic description of keys based on their transcription and extract statistical information from the transcription file b y u tilising a Python script that analyses the text file and returns a detailed analysis of its content, as described in Tudor (2019) and Tudor et al. (2020).

The first major section of our output focuses on the analysis of ciphertext symbols, beginning with the type of symbols used for encryption. Here we differentiate between 3 major types, namely Latin alphabet, digits, and graphic signs.

The next section of the output looks more in-depth into the internal structure of the ciphertext symbols, which we will refer to as unigraphs, bigraphs, trigraphs, and 4+graphs. What counts as unigraphs are usually digits, isolated letters or graphic signs.

We then move on to investigate plaintext units. Similarly to ciphertext, these are separated in unigrams, bigrams, trigrams, and 4+grams. We do

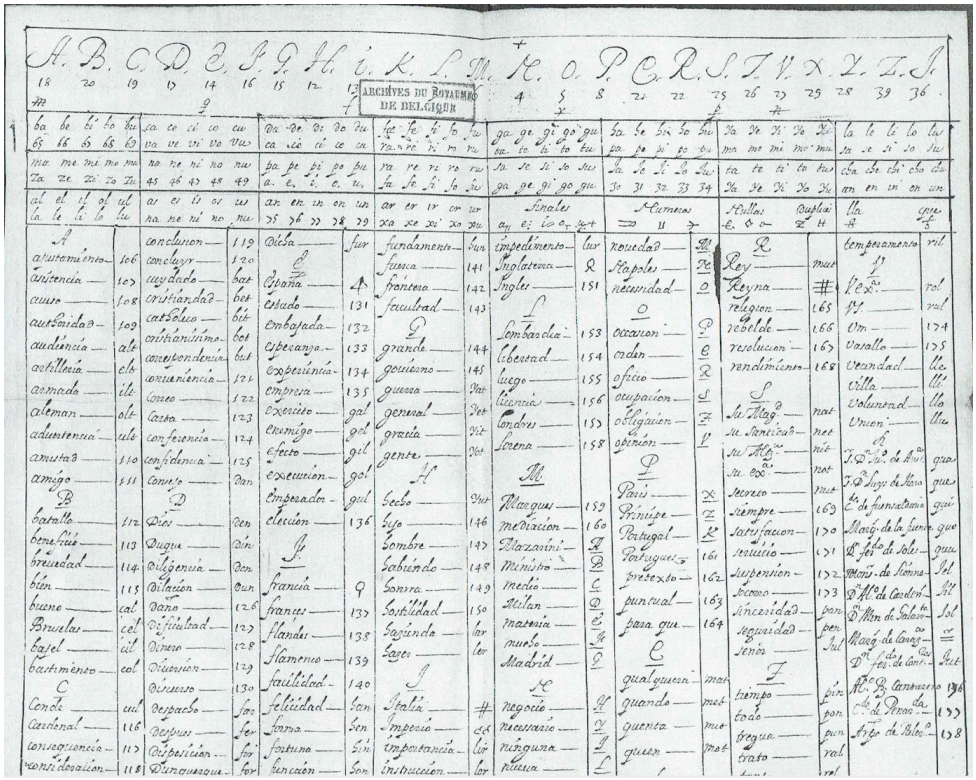

Figure 1: Example of homophonic key with variable length code (ARA Brus SEG inr.2chiffres1647-98 key3, 2018).

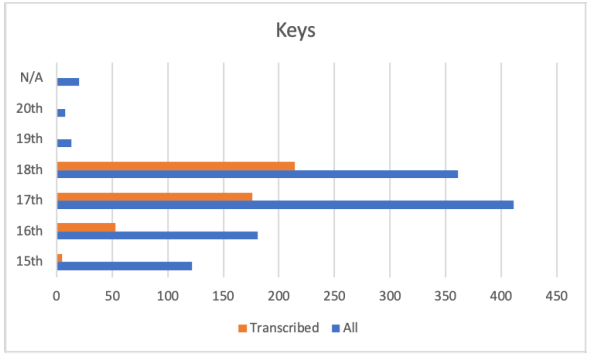

Figure 2: Key distribution throughout centuries in the DECODE database

add 3 additional ones, namely nulls, empty and cancellation signs.

For the most part, the type of plaintext unigrams that we find in keys are either letters or digits, even punctuation in some cases. Bigrams and trigrams are commonly either non-lexical units (e.g. double letters that occur frequently in the language of encryption, such as “ll” or “ee” in English, syllables, morphemes etc.), or short function words (“at”, “for”, “to”, “and” etc.). Under 4+grams we include those units that consist of 4 or more elements, such as longer function words or nomenclature entries, which can consist of names, places, common words. Nomenclatures can also include words that are specific to the lingo used in the topic the key was designed for, such as army terms in military correspondence.

Even though nulls and empty elements might sound the same in theory, we differentiate between them in terms of their purpose; we look at nulls as entities that have been purposefully inserted by the author of the key to hinder the decryption process, while ”empty” entities are unintentional. The latter usually occurs in preset tables of codes that are later filled in with plaintext unit, but some codes are not assigned semantic significance, as shown in Figure 3.

The last category, cancellation signs, refers to those codes that not only do not carry significance, but also negate a certain number of codes in their vicinity, rendering them null as well, which we exemplify in Figure 4.

Once we described the code and plaintext structure, we can analyze the distribution of ciphertext symbols to plaintext elements from several different perspectives.

First, we establish the cipher type, such as simple, homophonic or polyphonic substitution, or a

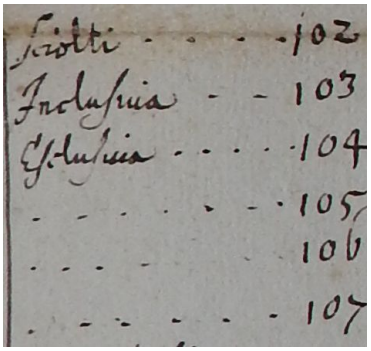

Figure 3: Excerpt from key containing empty entities.

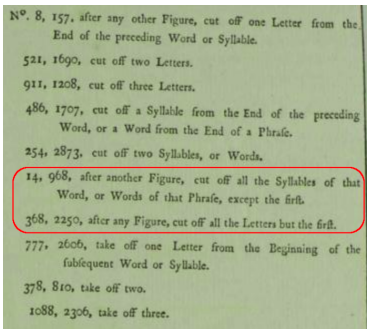

Figure 4: Excerpt from key containing cancelling signs.

mix of these. Then, we also look into the length of the codes used for encryption, be it fixed or variable. We also indicate how many of the codes are used for encrypting each level of n-grams and similar plaintext representations.

The last significant portion of the automatic analysis looks into the specific distribution of ciphertext to plaintext units for each section of the key, separated into alphabet, nomenclature, nulls, empty, and cancellation signs.

The final step is to output all of the specifications for each key into a global csv file.

5 Results

In the transcriptions and their structural descriptions, we study the entities that were chosen to be encoded, the codes themselves, and the relation between the codes and the plaintext entities.

5.1 What is encoded

Given the plaintext entities, we analyze them with respect to the number of characters and their types as well as the language(s) they represent.

5.1.1 Plaintext

Plaintext entities, such as characters, syllables, words, or sentences that are described to be coded in the keys, can be rather short, like a size of the alphabet of ca 20-30 entities, to several hundred like a long list of a nomenclature. 72% of the keys contain over 100 different plaintext entities, of which all contained the plaintext alphabet and an additional list of word-like elements, such as syllables, function words, frequent content words, and named entities. We present the distribution over the keys on the basis of the length of plaintext divided into unigrams of length 1, bigrams of length 2, trigrams of length 3, and 4+ grams of length 4 or more in Figure 5.

5.1.2 Languages

The involved languages that we find among the plaintext elements in the transcribed keys are: German (DE), English (EN), Spanish (ES), French (FR), Hungarian (HU), Italian (IT), and Latin (LA). See Figure 6 for an overview. Keys may encode entities not only in one but also in several languages. The involved languages depend on the time period, the geographic area of the corresponding people, and the lingua franca of that time.

Almost 30% of the keys contain several languages, which is hardly surprising due to the well-known property of code-switching in historical texts. Latin occurs in almost half of the keys, followed by English, French, and Italian.

Figure 6 illustrates the distribution of the languages, occurring as the only language, or as one of several languages.

5.1.3 Nulls

Keys might also contain nulls, elements that are fake codes without any underlying plaintext. Ca 32% of the keys contains one or several nulls. How many nulls are used vary across keys, as illustrated in Figure 7. Nulls can be listed as a finite set of numbers, or defined in cleartext corresponding to several hundred codes.

5.1.4 Empty plaintext

Keys are not necessarily complete, sometimes we find a list of codes in some structural manner without any corresponding plaintext. In fact, 19% of the keys contained some empty plaintext elements ranging from 1 up to 2500 empty places.

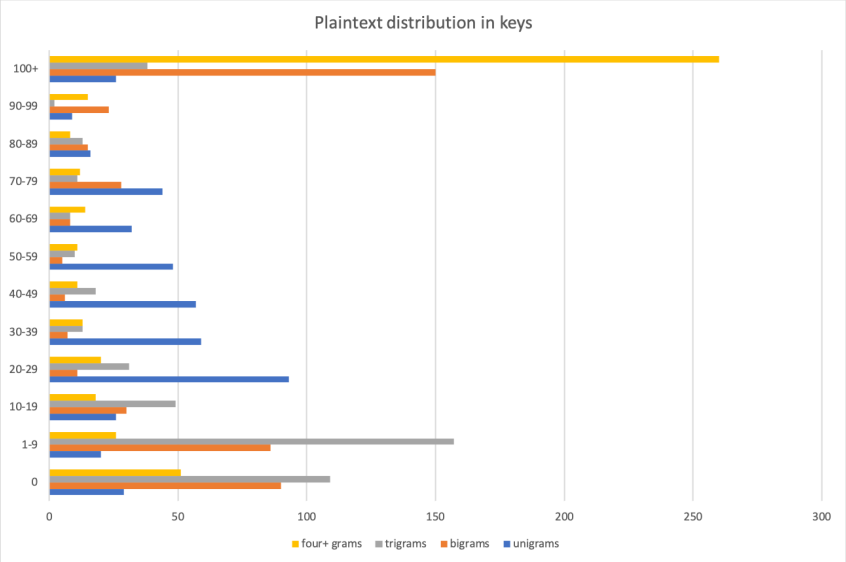

Figure 5: The distribution of plaintext entities of variable length: unigrams, bigrams, trigrams and four+grams.

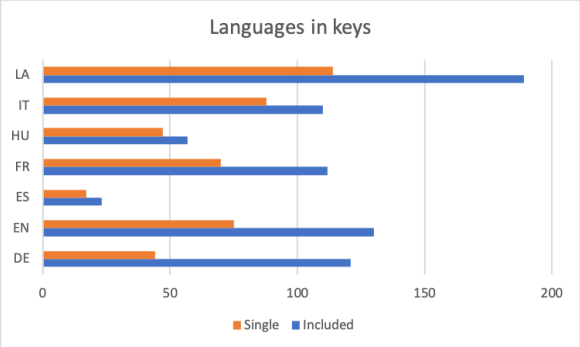

Figure 6: The distribution of languages in keys: blue marks the number of keys the language occurs, and orange marks the number of keys where the language is the only one used.

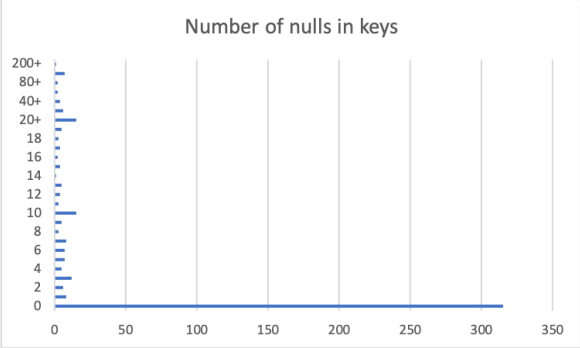

Figure 7: The number of nulls in keys.

5.2 How it is encoded

Encoding systems have been varying over time, and here we try to summarize the encoding systems in terms of symbols and code types.

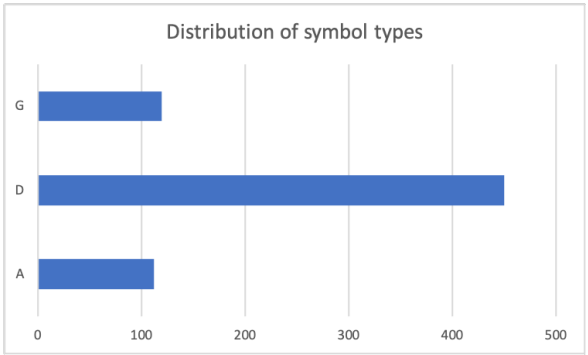

Figure 8: Symbols in keys: A=alphabet, D=digit, G=graphic sign.

5.2.1 Symbol systems

We distinguish between alphabets such as Latin and Greek, digits, and graphic signs such as alchemical symbols or Zodiac signs. The great majority, 98% of the keys contain digits (0-9) and only 25% use codes expressed as alphabetical characters or graphic signs, as show in Figure 8. In 72% of the keys, the only symbols that are used are digits. The remaining ones combine digits with alphabets, oftentimes Latin letters. Graphic signs occur only in few keys. The distribution of symbol sets across keys is illustrated in Figure 9.

5.2.2 Code types

85% of the keys contain codes of variable length, and only 15% have a fixed length code, mostly 2-

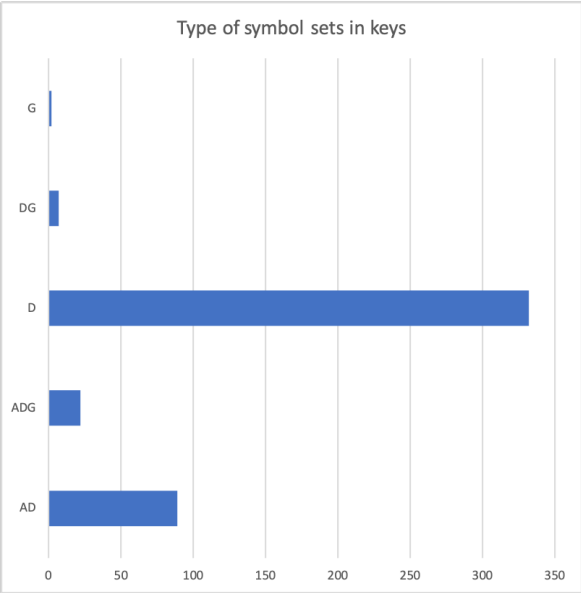

Figure 9: The combination of symbols in keys: A=alphabet (Greek or Latin), D= digits, G=Graphic signs.

digit codes.

Code types vary across the plaintext entity types not only in length but also in type. For example, it is common that the alphabet is encoded as 2-digit homophonic codes while nomenclatures have 3-digit simple substitution code system. Thus, the distribution of code types vary not only across but within a single key. In Figure 10, we show the code types for alphabets, nomenclatures as well as for nulls. Typically, while several characters in alphabets are often encoded with two or more codes resulting in a homophonic substitution, elements in nomenclatures tend to have one code only.

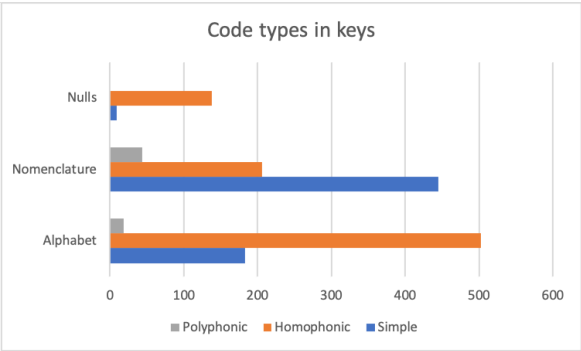

Figure 10: The number of nulls in keys.

Given the various code types in a key, we analyze the type given their components, see Figure 11. Homophonic substitution is far most popular either on its own or combined with simple substitution. Purely polyphonic or simple substitution occur seldom, and if they do they are often combined with homophonic codes. In a partly homophonic, partly polyphonic cipher key, for example, some elements of the plaintext alphabet are substituted by several cipher text characters (that is the homophonic component), while some elements of the nomenclature are substituted by the same code (that is the polyphonic part).

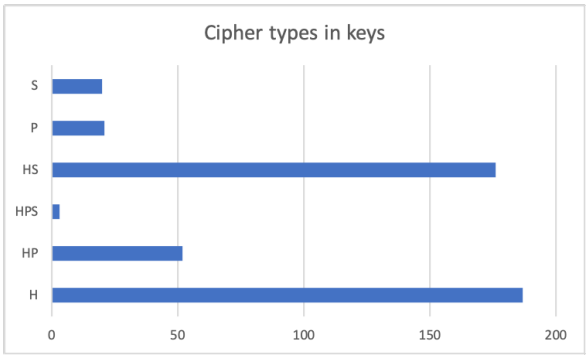

Figure 11: The distribution of cipher types in keys.

5.2.3 Cancellation

Cancellation, i.e. codes that define elements that should be removed in the plaintext, are not very common but appear in ca 4% of the keys, and not until the 18th century. Cancellation can be defined in many different ways, not only as codes but as in cleartext describing how cancellation is performed, which can be seen in Figure 4.

5.3 Trends

Given the keys' structural description, we can investigate the trends throughout the centuries concerning what has been chosen to be encoded and how. Since the set of structurally described keys that have been automatically extracted from transcriptions originate from the 17th to 18th centuries, (see the orange bars in Figure 2), we manually extracted structural information from 251 keys without any transcriptions originating from the 15th and 16th centuries. In total, we investigate 700 keys. In the subsequent paragraphs, we report some of our findings about the main trends of key structure over the centuries.

The usage of the types of symbols that have been chosen for encoding varied over the centuries, as illustrated in Figure 12. While alphabetical characters, digits, and graphic signs were evenly distributed in the 15th century, we can see a clear increase in tendency to use digits as the main

encoding at the expense of Latin letters or graphic signs, which we can hardly find in keys from the 18th century.

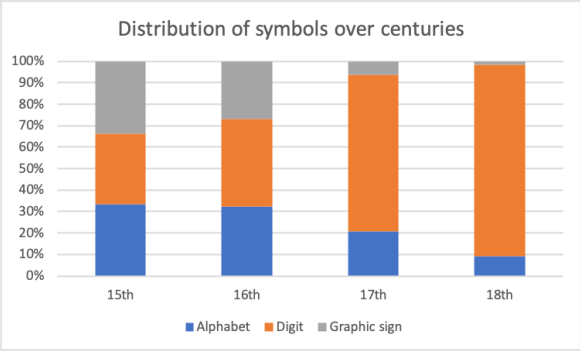

Figure 12: The distribution of symbols over time.

The symbol systems used in keys often contain a combination of digits, letters, and graphic signs. In the 15th century, all three types of symbols were combined in almost all keys, but this eclectic symbol set have been reduced in the 16th and 17th centuries in favor of digits in combination with Latin letters. The distribution of various symbols sets over centuries is shown in Figure 13.

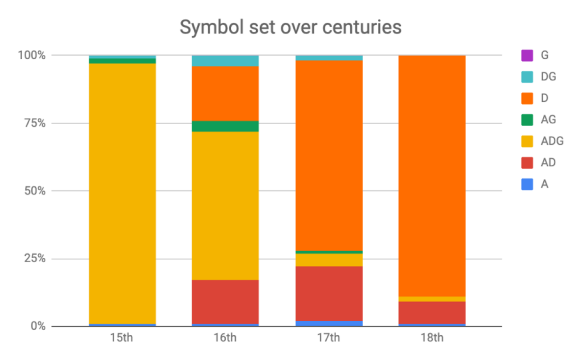

Figure 13: The distribution of symbols set containing (a combination of) Latin alphabet (A) digits (D) and/or graphic signs (G) over time.

The usage of the length of the codes also varies over time, as illustrated in Figure 14. The great majority of keys contain codes of variable length and the length typically differ between alphabetical elements, nomenclatures, as well as nulls.

To investigate the type of codes in more detail, we analyzed the type of codes used for alphabets and nomenclatures separately, distinguishing between simple, homophonic, and polyphonic distributions.

Encodings of alphabetical signs were mostly homophonic, as shown in Figure 15. Quite surprisingly, however, we can see a decrease in favor of simple substitution which became more frequent in the 17th and 18th centuries. This might be due to the increase in the size of the nomenclatures over time.

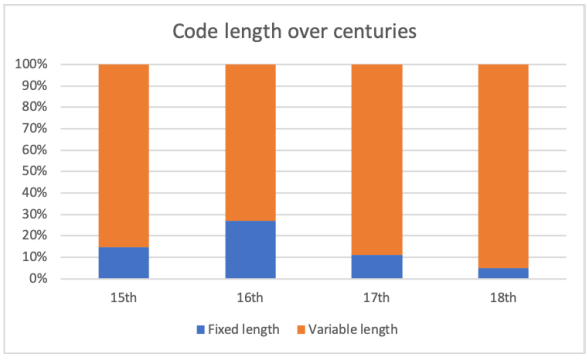

Figure 14: The distribution of fixed vs variable length codes over time.

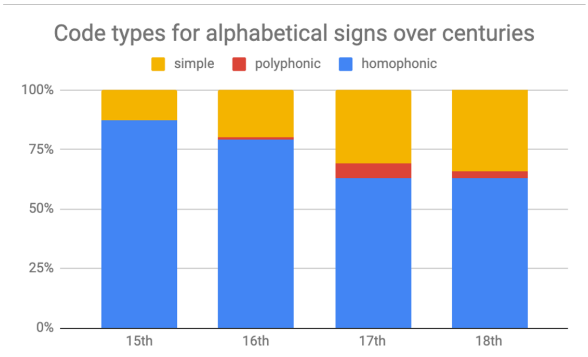

Figure 15: The distribution of code types for alphabetical signs over time.

Encodings of nomenclatures, on the other hand, are mostly simple substitution, but homophonic and even polyphonic encodings become standard in the 17th and 18th centuries, see Figure16.

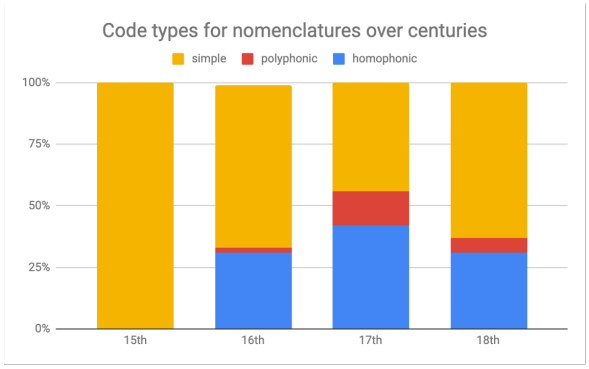

Figure 16: The distribution of code types for nomenclatures over time.

The usage of nulls in keys also varied over time, as illustrated in Figure 17. While nulls have been frequently occurring in keys, i.e. 96% of keys included nulls in the 15th century, we find nulls in 27% of the keys in the 18th century. The nulls were in the great majority of the cases (94%) encoded with at least two possible codes.

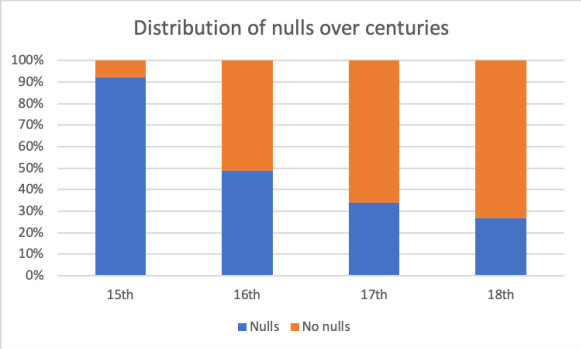

Figure 17: The distribution of nulls over time.

Clearly the usage of nulls decreased over time, and codes for cancellations have not been used until the 18th century.

6 Discussion

One surprising result that emerged after looking more in-depth into the structure of keys was a rather large amount of nomenclatures that use homophonic substitution. This was particularly interesting to investigate as the phenomenon was mostly visible in the keys that were automatically analysed by our script, and not nearly as much in those that passed through a manual analysis. Upon further inspection, we were able to isolate two main factors that cause this phenomenon.

- Frequent bigrams that can occur in a language, such as ”ae”, ”oe”, ”au” in Latin (NAH G15 CAPS C FASC 43 18, 2018), or ”gy”, ”cz, ”sz” in Hungarian (NAH G15 CAPS C FASC 43 40, 2018), can often be encoded by means of homophonic substitution. In our analysis, we consider bigrams to be part of the nomenclature, whereas some keys include them on the same level as the alphabet. For example, if at first sight it seems like a key is using homophonic substitution at alphabet level and simple substitution at nomenclature level, we may discover that the author included some bigrams which are encoded by 2 or more codes at alphabet level, which in turn makes the nomenclature homophonic as well.

- Some very large tables (100+ ngrams) can use homophonic substitution only for a few entities in the nomenclature table, oftentimes those that are used most frequently in the language (e.g. ”aller” - *to go*, ”peu” - *few* in French (KHA_ A29_ PWIV_ inr301_ B, 2019)) or for the purpose of the correspondence (e.g. titles, such as ”The King”, ”His Majesty” (ÖStA HHStA Stk Int Chiffrenschlüssel fasc 20 kt14 152, 2020)). These tend to be rather hard to spot with the naked eye among the multitude of plaintext entries.

This only goes to show that automatic methods are a lot more reliable when it comes to picking up subtle elements of key structure.

All in all, given the results presented above, we cannot draw certain conclusion about how the keys have been used — we can only see what the intentions of the key creators have been. More ciphertexts and systematic studies would be needed about the actual usage of the keys.

7 Conclusion

In this paper, we investigated 700 cipher keys from the 15th to the 18th centuries, all originating from European archives and libraries. We described the keys' internal structure and their morphology looking at what has been chosen to be encoded and how over four centuries. In particular, we described the type of the symbol set and the code structures used, and the changes and trends of each century.

Not surprisingly, we found that keys evolved over time, and their structure changed in various ways. While codes with various symbols including alphabets, digits, and graphic signs were dominating in the 15th century, using digits only became more frequent to became the standard in the 18th century. The codes varied in length for alphabetical signs and nomenclatures throughout all centuries while codes with fixed length seemed to be most popular in the 16th century. Coding alphabetical signs were mostly homophonic, but simple substitution of letters became more frequent as the length of the nomenclatures increased over time. Nomenclatures, however, were mostly encoded as simple substitution. Nulls have been frequently used in the 15th century and decreased signifi-

cantly over time. Cancellation as phenomenon became popular in the 18th century.

Our results presented in this paper are based on 700 original keys from four centuries, but the dataset is rather opportunistic — we took what was available to us — the data is not evenly distributed across geographic areas, countries, or senders/receivers. In the future, we intend to extend our collection with more keys from a large number of places, and make in-depth analyses of the nomenclatures and the involved plaintext languages.

Acknowledgments

We would like to thank Anne-Simone Rous and Karl de Leeuw for generously sharing their keys with the public through the DECODE database.

This work has been supported by the Swedish Research Council, grant 2018-06074, DECRYPT – Decryption of Historical Manuscripts.

References

ARA Brus SEG inr.2chiffres1647-98 key3. 2018. Reproduced image from Algemeen Rijksarchief, Secretairerie d'Etat et de Guerre, inv.nr. 2, DECODE link: https://cl.lingfil.uu.se/decode/database/record/960 .

J. P. Devos. 1950. *Les chiffres de Philippe II (1555-1598) et du Despacho Universal durant le XVIIe siècle.* Brussels: Académie Royale de Belgique.

David Kahn. 1996. *The Codebreakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet*. Scribner, New York, NY.

David Kahn. 2008. The future of the past—questions in cryptologic history. *Cryptologia*, 32:56–61.

KHA_A29_PWIV_inr301_B. 2019. Reproduced image from Koninklijk Huisarchief (KHA), Prins Willem IV, inv.nr. 301 B, DECODE link: https://cl.lingfil.uu.se/decode/database/record/1024.

Benedek Láng. 2018. *Real Life Cryptology: Ciphers and Secrets in early modern Hungary*. Atlantis Press, Amsterdam University Press.

George Lasry, Beáta Megyesi, and Nils Kopal. 2020. Deciphering Papal Ciphers from the 16th to the 18th Century. *Cryptologia*.

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, June.

Beáta Megyesi. 2020. Transcription of Historical Ciphers and Keys. In *Proceedings of the 3rd International Conference on Historical Cryptology, HistoCrypt20*, Budapest, Hungary, June.

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

Aloys Meister. 1906. *Die Geheimschrift im Dienste der Päpstlichen Kurie von Ihren Anfängen bis zum Ende des XVI. Jahrhunderts*, volume 11. F. Schöningh.

NAH G15 CAPS C FASC 43 18. 2018. Reproduced image from National Archives of Hungary, G15 Caps. C. Fasc. 43. 18., DECODE link: https://cl.lingfil.uu.se/decode/database/record/600 .

NAH G15 CAPS C FASC 43 40. 2018. Reproduced image from National Archives of Hungary, G15 Caps. C. Fasc. 43. 40., DECODE link: https://cl.lingfil.uu.se/decode/database/record/581 .

Ludwig von Rockinger. 1892. Über eine bayerische Sammlung von Schlüsseln zu Geheimschriften des sechzehnten Jahrhunderts. *Archivalische Zeitschrift*, pages 18–92.

Crina Tudor, Beáta Megyesi, and Benedek Láng. 2020. Automatic Key Structure Extraction. In *Proceedings of the 3rd International Conference on Historical Cryptology, HistoCrypt20*, Budapest, Hungary, June.

Crina Tudor. 2019. Studies of Cipher Keys from the 16th Century: Transcription, Systematisation and Analysis. Master thesis in Language Technology, Uppsala University, Sweden.

ÖStA HHStA Stk Int Chiffrenschlüssel fasc 20 kt14 152. 2020. Reproduced image from Österreichisches Staatsarchiv, Haus-, Hof- und Staatsarchiv, Staatskanzlei Interiora, Chiffrenschlüssel, Kt. 14. Fasc. 20. f 152., DECODE link: https://cl.lingfil.uu.se/decode/database/record/1397.