What Encryption Errors Can Reveal: Cross-Cipher Errors in Mary Queen of Scots' Letters

Norbert Biermann
Universität der Künste Berlin
mail@norbertbiermann.de

Satoshi Tomokiyo
Cryptiana
verlat@hotmail.com

George Lasry
The DECRYPT Project
george.lasry@gmail.com

Abstract

In the recently deciphered letters from Mary Queen of Scots, a large number of systematic encryption errors were found and attributed to confusion as a result of concurrently using at least one other cipher key to communicate with a different recipient. In this paper, we further analyze such cross-cipher errors in those letters and identify additional cipher keys involved. This analysis also reveals valuable insights on the secret communications of Mary, Queen of Scots. We employ several techniques including statistical analysis, which may be applied to the analysis of encryption errors in other collections of historical enciphered documents.

1 Introduction

Lasry et al. (2023) found an unexpectedly high number of encryption errors in Mary Queen of Scots' ciphered letters to Michel Castelnau, the French ambassador to England, and concluded that a large part of them is explainable by the assumption that the encipherer unintentionally[1] used symbols from another cipher which was in use at the time, most probably to communicate with another recipient. They provide a hypothesis to explain this cross-cipher phenomenon (op. cit., Appendix B):

> When a cipher secretary starts using a new cipher he is not yet familiar with, it is likely that at first, he will be looking up the correct graphic symbol(s) for each letter of the alphabet he wants to encipher, in the cipher table. But after a while, the secretary is likely to start memorizing symbols representing the most frequent letters of the alphabet so that enciphering becomes faster, without always having to consult the cipher table. If the same secretary needs to encipher another letter on the same day with a different cipher, and if he continues to rely on his memory rather than on the cipher table to encipher the most frequent letters of the alphabet (e.g., i or e), he may subconsciously recall the symbol for that letter from the wrong cipher table, which he may have used recently to encipher another document. This theory is also consistent with such errors being systematic, and since the secretary might not be aware of them, he repeats them throughout the document he is currently enciphering.

We assume that the likelihood of such errors, which we call *cross-cipher errors* (CCEs), was increased by the fact that Mary lived in captivity and her secret letters not only had to be written in cipher, but the entire process of writing and sending them had to be concealed, putting her secretaries under mental stress. Moreover, the coming and going of secret messengers dictated the cryptographic work rhythm: Sometimes, a secretary was forced to work in a hurry or all night long so as not to miss the messenger's departure.[2]

[1] The opposite assumption, that such errors were made deliberately to increase the security of the cipher is worth discussing, but there are several points against it. One is that those errors were quite often corrected by the encipherer in ways that cannot be explained by cryptological sophistication, such as striking out the wrong symbols or simply overwriting them with the correct ones.

[2] Nau, Mary's French secretary, writes in a postscript to a letter to Archbishop Beaton that he was up all night deciphering letters that had arrived late in the evening (Labanoff, 1844, v, 13, *J'ay veillé toute ceste nuict pour deschiffrer voz lettres et aultres qui furent apportées hier au soir bien tard*). In the letters to Castelnau, we find wordings such as *la haste du (*or *de ce) porteur* (F89, F113) or *je suis si pressée du partement inopiné de ce porteur* (F34), and, again to Beaton, *ce qui suit est escript en fort grande haste* (Labanoff, 1844, v,

Lasry et al. identified the cipher between Mary and her ambassador to France, James Beaton, Archbishop of Glasgow, as the source for a large part of the cross-cipher errors in Mary's letters to Castelnau. The authors also hinted at the possibility that other ciphers were involved in this intriguing cross-cipher phenomenon.

In this paper, we confirm this latter hypothesis by systematically analyzing thousands of cross-cipher errors in Mary's letters to Castelnau, and we present various insights that this research produced.

In Section 2, we introduce the ciphers Mary used to communicate with Castelnau and Beaton, and use examples to demonstrate how cross-cipher errors manifest themselves. This reflects the state of knowledge as described in Lasry et al. In Section 3, we provide statistics of the systematical survey of all potential cross-cipher errors in the examined letters from Mary to Castelnau. We then introduce in Section 4 two additional ciphers we found to have caused such errors. In Section 5, we use different methods to corroborate our findings. Further intriguing insights resulting from our analysis are presented in Section 6. We conclude our examination in Section 7.

2 Two ciphers – occasionally mixed up

In this section, we describe two ciphers Mary Queen of Scots used, one with Michel de Castelnau and the second with James Beaton, and demonstrate CCEs, i.e., how symbols of the latter cipher occasionally appear unintentionally in the letters written in the former cipher.

16th century nomenclators were typically composed of:

- An alphabet section with one or more symbols (homophones) representing each letter of the alphabet
- A nomenclature with symbols representing common words, parts of words, persons, or places
- Special symbols like *nulls*, which should be ignored, a *deleter* symbol indicating that the last symbol should be erased, a *repeater* symbol for doubling the last symbol, or symbols for punctuation marks.

22). Note that we follow the naming convention outlined in Lasry et al. (2023, 109) to refer to individual letters discussed therein (F89 etc.).

Both ciphers we introduce in the following are based on this structure, and they have a further similarity in that they use *diacritics*, i.e., marks that alter the meaning of specific symbols when added to them.

2.1 Mary-Castelnau cipher (MC)

This is the cipher Mary Queen of Scots used (at least) between 1578 and 1584 to communicate with Michel Castelnau, as reconstructed by Lasry et al. (2023). All newly discovered letters to the French ambassador (op. cit.) are written in this cipher which we refer to as the *Mary-Castelnau cipher* (MC). Table 1 shows an extract of the cipher key.

Table 1: Mary-Castelnau cipher (MC), extract

a	b	c	d	e	f	g	h	i	k	l	m	n	o	p	q	r	s	t	u	x	y	z
o	q	ω	//	c	∂	θ	ẟ	α	ʄ	n	ŋ	x	ʒ	+	y	s	ε	ʌ	ı	ϯ	8	℘
	π	4	7	ɥ	ʕ		9	¤	\	℘		2		ꝛ	ʃ	ж	⊥	T				ǂ

q,	advis	ω·	ent	n/	monsieur
\,	affaires	ʒ̣	est	e·	on
α·	ant	℘̣	et	℘̣	par
ɓ·	com	ω,	faire	ʕ̣	plus
6·	con	c·	ion	7̣	pour
ỵ	de	9·	ite	ϯ̣	que
ʜ,	depesche	ε̣	je	π,	service
ʃ̣	des	ж̣	la	2̣	tout
ʕ̣	dict	ʌ̣	les	∂̣	vous
+·	en	!̣	leur	9̣	voz
4·	endre	ç̣	mais		

K	Monsieur de Mauvissiere
ˀ delete last symbol	*ɟ repeat last symbol*

2.2 Mary-Beaton cipher (MB)

Apart from the newly discovered letters which are held by the Bibliothèque nationale de France, most of Mary's letters in ciphertext are found in TNA SP,[3] BL,[4] and SCA JB.[5] Of these, SP 53/22,

[3] The National Archives, State Papers. The letters we used are from SP 53/10 and SP 53/18.

[4] The British Library. The letters we used are from Cotton MS, Caligula C III.

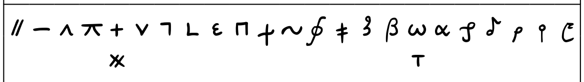
[5] The University of Aberdeen, Scottish Catholic Archives, Archbishop James Beaton's Papers. Mary's letters in SCA JB are mainly those addressed to James Beaton, and their plaintext is printed in Labanoff (1844) as "Déchiffrement. — Collection du docteur Kyle, à Preshome." Bishop Kyle deciphered the letters in cipher and provided his decipherments to Labanoff (1844, i, 399).

SP 53/23, and SCA JB 3/4 include collections of keys. Several of these ciphers were used between Mary and James Beaton. The particular cipher that we refer to as the *Mary-Beaton cipher* (MB)[6] is found in TNA SP 53/23 no. 38 (dated 1577) and the verso of no. 34 (crossed out), as well as SCA JB 3/4 p. 40 (no. VIII).

Based on extant letters, MB was used at least from 20 February 1576[7] to 10 September 1582.[8]

Table 2 shows an extract from the MB cipher key.

Table 2: Mary-Beaton cipher (MB), extract

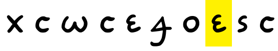
a b c d e f g h i k l m n o p q r s t u x y z								
								
o.	advis		ɘ	est		ŋ	on	
s.	affaires		ꞩ	et		s	par	
ȣ	ant		m	faict		8	plus	
ɱ	com		ꞩ.	force		o	pour	
ŋ	con		ꞓ	ite		n	que	
ʂ	de		X	je		2.	service	
\\.	depesche		y	les		6	vous	
6	des		+	leur		e	voz	
ꞓ	endre		ʒ	mais				
ꟻ	ent		∂.	monsieur				
8̈ Monsieur de Glasgo								
⧺ *repeat last symbol*								

2.3 Examples of cross-cipher errors

All the examples given below are taken from Mary's letters to Castelnau, which means they should be decrypted using the MC key (Table 1).

Example 1.

n e c e s s a s r e



Any contemporary decipherer would have tacitly corrected this[9] to the French word *necessaire*. The symbol ε appears in this word with two different meanings: First, it stands for *s*, which is correct according to MC, but it was also wrongly used to encipher *i*. For the latter case, we assume that the secretary unconsciously made use of the Mary-Beaton cipher in which the symbol ε is assigned to the letter *i*. We denote this specific cross-cipher error as ε=*i*, indicating that in MB, the cipher that causes this CCE, the symbol ε represents *i*.

We distinguish between *uncorrected* and *corrected* errors. In many cases, an encryption error was spotted and corrected by the secretary enciphering the letter. Such corrections could have been made by crossing out the wrong symbol or by overwriting it with the correct one. In such cases, it is often difficult to determine what the corrected (wrong) symbol looked like, so we have not included them in the analysis presented here. But most often, to invalidate a wrong symbol, the *deleter* symbol was used, and written right after the original (deleted) symbol which is still legible, and therefore useful for our analysis. While our first example represented an uncorrected error, Example 2 shows corrected ones:

Example 2.

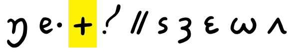
Monsieur de M a u v i s s i e r e

Here,[10] the scribe made two errors while enciphering *Monsieur de Mauvissiere*. The first error is using the symbol ʂ to encipher *de* (ʂ=*de*). This symbol appears only in the key of MB, not in MC. The second error is the same one as in example 1 (ε=*i*). This time, however, both erroneous symbols have been invalidated and corrected.

Both cross-cipher errors presented so far, ε=*i* and ʂ=*de*, can be attributed to the Mary-Beaton cipher. This is the case for a large number of errors appearing in the letters to Castelnau, but is not true for all such errors, as demonstrated in Example 3:

Example 3.

m on d r o i c t



Here,[11] we assume that the secretary made – and corrected – the cross-cipher error +=*d* in *mon droict*. This time, MB cannot explain the error: although the symbol + is also part of MB, it stands for *e* in this cipher. This is why we assume that there was another, unidentified cipher also used by Mary's secretary at the time, in which + enciphers *d*, causing this specific cross-cipher error.

Errors like those in our three examples recur systematically and quite often. For example, in

[6] It is called "chiffre de Nau" in the endorsement of a letter to Beaton (18 September 1581, SCA JB 2/5/7, printed in Labanoff (1844, v, 254)). MB may have been given to Nau by Beaton when Nau came over from France to enter Mary's service in 1575. Before this cipher came into use, another cipher (SCA JB 3/4 No. II) called "chiffre de Raulet" (Labanoff, 1844, v, 89) was used.

[7] SCA JB 2/3/13

[8] BL Cotton MS, Caligula III, f. 426

[9] F87, line 10

[10] Beginning of F87

[11] F87, line 14

a single letter (F87), we observe +=*d* 7 times, and ɛ=*i* no less than 64 times. This means they cannot be purely random errors which are expected to some extent in any historical ciphered letter.

3 Systematically analyzing cross-cipher errors

In the previous sections, we summarized and exemplified the findings described by Lasry et al. (2023, App. B). In this section, we record all potential cross-cipher errors that occur in the letters from Mary to Castelnau (not only those we can assign to MB), drawing new insights from initial statistical observations.

3.1 Determining potential cross-cipher errors

For our examination, we systematically compared deciphered and edited versions of the letters with the raw transcripts of the symbols. Non-matching symbols found in this way were considered potential CCEs unless another plausible explanation for their appearance could be found, such as:

- The secretary missed a letter in the middle of a word while enciphering
- The secretary forgot to add the correct diacritic near a symbol, or wrote a wrong or misplaced diacritic
- What seems to be an error may be an acceptable spelling of a word according to the – sometimes inconsistent – orthography of 16th-century Middle French

Similarly, we took into account errors corrected by the secretary only if other plausible causes of enciphering errors could be excluded.

3.2 Overview of the results

Examining all the 57 ciphered letters from Mary to Castelnau, we counted 2,740 occurrences of potential cross-cipher errors. The examined letters count up to a total of around 176,900 symbols, of which the 2,740 potential CCE occurrences make up a percentage of 1.55 – a remarkably high ratio.

Table 3: Cross-cipher errors by criteria

	uncorrected	corrected	Σ
intrinsic symbol	1,487	911	2,398
extrinsic symbol	80	262	342
Σ	1,567	1,173	2,740

Table 4: Cross-cipher errors by occurrence

CCE	occur.	cipher	CCE	occur.	cipher
o=*i*	525	?	ſ=*t*	17	MB
ɛ=*i*	234	MB	ƀ=*vous*	17	MB
∧=*c*	161	MB	c=*c*	16	?
ı=*e*	105	?	τ=*c*	15	?
+=*e*	73	MB	τ=*t*	15	?
ɑ=*a*	68	?	∠=*t*	14	?
τ=*r*	68	MB	♪=*u*	14	MB
//=*a*	67	MB	s=*par*	13	MB
ω=*r*	67	MB	ƒ=*n*	12	MB
ȝ=*s*	64	?	ı=*s*	12	?
s=*s*	55	?	ɑ=*n*	10	?
ƃ=*et*	48	MB	ŧ=*o*	10	MB
m=*r*	48	?	ƃ=*r*	10	?
x=*[x2]*	44	?	ϰ=*r*	10	?
π=*d*	43	MB	o=*n*	9	?
ƃ=*de*	42	MB	++=*[x2]*	9	MB
ŋ=*con*	39	MB	6=*d*	8	?
o=*pour*	36	MB	+=*d*	8	?
∝=*s*	33	MB	ϑ=*est*	8	MB
ϰ=*e*	32	MB	—=*l*	8	?
o=*o*	31	?	ȝ=*p*	8	MB
∧=*i*	29	?	⊬=*leur*	7	MB
∧=*n*	28	?	ƃ=*r*	7	?
ϑ=*p*	28	?	ʊŋ=*on*	6	MB
c=*a*	25	?	4.=*que*	6	?
∼=*m*	25	MB	//=*s*	6	?
+=*u*	25	?	x=*s*	6	?
s=*p*	23	?	o.=*advis*	5	MB
ı=*a*	21	?	—=*b*	5	MB
ƃ.=*dict*	20	?	ɛ=*e*	5	?
+=*r*	20	?	⊬=*l*	5	MB
ꝓ=*y*	20	MB	ɛ·=*la*	5	?
ȝ=*n*	19	?	e=*n*	5	?
n=*que*	18	MB	|=*r*	5	?
//=*u*	18	?	ƈ=*z*	5	MB
4=*t*	17	?			

Table 3 shows how these occurrences distribute

across criteria corrected/uncorrected and whether the wrongly used symbol is part of MC (intrinsic) or not part of it (extrinsic).

One can see that CCEs with intrinsic symbols are much more frequent than those with extrinsic symbols. This suggests that an overlapping vocabulary of symbols between two cipher tables was likely to lead to confusion and increase the occurrence of cross-cipher errors. Errors involving intrinsic symbols were more likely to be overlooked by the secretary, while the majority of CCEs with extrinsic symbols were fixed during the encryption process (probably because an extrinsic symbol stands out visually, usually causing the secretary to spot it and correct it right away).

The CCEs can be attributed to over 140 distinct errors. To exclude random errors, in our analysis, we ignore those with less than 5 occurrences across all 57 letters, which reduces the above number to 71. These 71 errors, which we consider to be potential cross-cipher errors, are listed in Table 4, sorted in descending order of occurrence. The errors that can be attributed to the Mary-Beaton cipher are marked (MB in the third column).

4 Identifying other ciphers causing CCEs

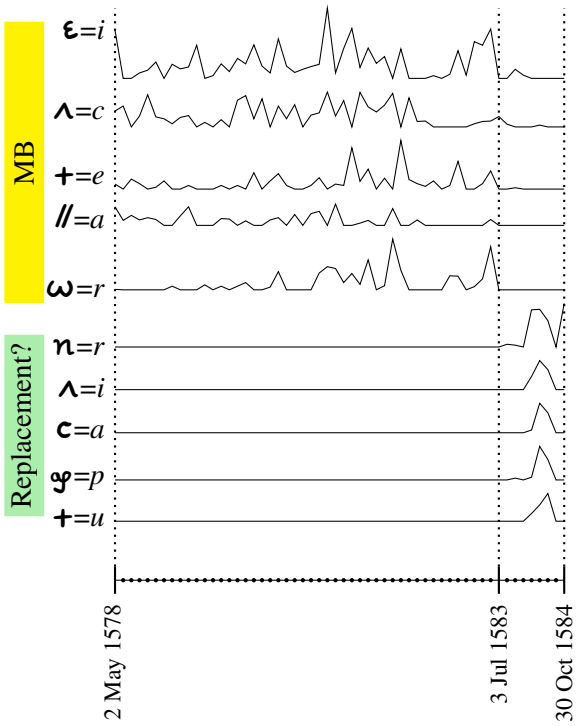
In this section, we present three other ciphers, two of which we found to cause cross-cipher errors in Mary's letters to Castelnau, matching the vast majority of the 71 CCE candidates shown in Table 4.

4.1 Late Mary-Beaton cipher (LMB)

When we examined the incidence of CCEs per individual letter, we observed that the CCEs attributed to the Mary-Beaton cipher practically disappear from a certain date (3 July 1583). Specific other CCEs, on the other hand, only seem to begin to appear after this date. Figure 1 illustrates this phenomenon showing the frequency of selected CCEs over time, per letter from Mary to Castelnau.

This led us to suspect that the cipher between Mary and James Beaton was replaced from this point onward by a new cipher which the newly-appearing CCEs may be attributed to. Indeed, we found such a cipher between Mary and Beaton, perfectly matching the CCEs that occurred after 3 July 1583. We refer to it as the *Late Mary-Beaton cipher* (LMB). The key of this cipher, however, could not be found in archives, and rather, it was reconstructed from three extant letters from 1586

Figure 1: Frequencies of selected CCEs per letter



enciphered with it.[12] The reconstructed key is given in Tomokiyo (2023).

Table 5 shows an extract of the cipher key of LMB, and Table 6 lists CCEs that we can attribute to LMB.

Table 5: Late Mary-Beaton cipher (LMB), extract

a	b	c	d	e	f	g	h	i	k	l	m	n	o	p	q	r	s	t	u	x	y	z
c	ϙ	⊤	6	ı	⊼	⊦	⊬	∧	ƒ	⍙	ȝ	ƨ	ϥ	⍙	ϙ	m	/	⊓	+	4	⊼	s
ɔ		⊥	ϥ	⊬	⊼	⊧	⊭	∨	⊼	⊧	ε	4	⍙	ϕ	n	⊻	⊔	—	4	⊻	ϥ
		⌐	ϥ		⊻				ϵ	⊤		ϥ	ā	ϥ	m̄	⊻		×			
		▾							⍙	⊦					n̄	⌿					
															ϥ						

∞· ant	s· est	o· pour
ƒ· ble	⊼: faire	4· que
⍙· con	ω· ion	//· tout
⊬· de	ε· la	∧· vous
ϕ· en	4· on	
	ō Monsieur de Glasgo	
⊻ o J	*delete last symbol*	

[12] TNA SP 53/18/60: Mary to Beaton, 12 and 16 July 1586, and to Monsieur de Mondevis, Cardinal de Laurea, 30 June 1586, printed in Labanoff (1844, vi, 362, 381, 347).

Table 6: CCEs attributed to LMB

CCE	occur.	cipher	CCE	occur.	cipher
ı=*e*	105	LMB	ȝ=*n*	19	LMB
m=*r*	48	LMB	⊤=*c*	15	LMB
∧=*i*	29	LMB	/=*s*	12	LMB
ϑ=*p*	28	LMB	6=*d*	8	LMB
c=*a*	25	LMB	4.=*que*	6	LMB
+=*u*	25	LMB	ε·=*la*	5	LMB

In the following, we provide examples of encryption errors that can be explained by LMB.

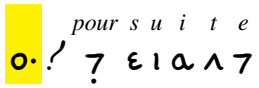
Example 4.

r est e b l t s s e m ent

This[13] should read *restablissement*, but symbols from LMB are used for the wrongly enciphered *a*, *i*, and *m*. Only the last error, involving an extrinsic symbol for *m*, has been spotted, deleted with a deleter symbol, and corrected by Mary's secretary.

Example 5.

pour s u i t e



This word[14] should read *poursuite*. The first symbol, o·, was wrongly taken from LMB to encipher *pour*. In this case, the secretary spotted the mistake, erased the wrong symbol with the deleter symbol, and wrote the correct 7 from MC.

Example 6.

des e d v o u e r de oit c e

This[15] should read *desadvouer de tout ce*. Three symbols are wrongly taken from LMB, of which only the last one is corrected. Without knowledge of the LMB cipher, the decipherer has no chance to conclude that //· is meant to encipher *tout*.[16]

4.2 Mary-Aubigny cipher (MA)

Still, the LMB cipher could not account for large numbers of other potential CCEs, including o=*i* which accounts for 525 occurrences, from a total of 2,740. We conducted a systematic search for another cipher that could explain those unattributed CCEs, and found a promising candidate, which we refer to as the *Mary-Aubigny cipher* (MA).

[13] F96, line 26

[14] F42, line 20

[15] F96, line 48

[16] This letter had been leaked to spymaster Walsingham and is printed in Labanoff (1844, v, 458). Indeed, we read there only *désadvouer de ce*, implying that Castelnau's secretaries ignored the symbol //· they could not make sense of.

The table for this cipher is the second of the two ciphers written on SP 53/22 f. 21. The endorsement in one word on the verso is illegible, but this cipher must have been used in Mary's correspondence with Esmé Stewart, sixth Seigneur d'Aubigny and Duke of Lennox, because the cipher symbol assigned for "daubugny" (O:) is the same as the enclosure marking for the Duke of Lennox in the letters from Mary to Castelnau.[17] Table 7 shows an extract of the key, and Table 8 lists the CCEs that we can attribute to it. Note that

- ı stands for the letter *e* in both LMB and MA, so the CCE ı=*e* is included in both Table 6 and Table 8
- the symbol ∠ (Table 8) does not appear in the cipher key (Table 7). However, we assume that it is a half-finished version of 4, both symbols enciphering *t*.[18]

Table 7: Mary-Aubigny cipher (MA), extract

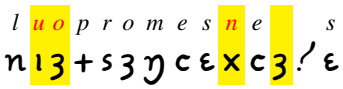
a b c d e f g h i k l m n o p q r s t u x y z
z 4 c q ı # □ ⊓ o 8 n ϑ ∧ v s ## + 3 4 6 7 8 9
q· de
O: (Monsieur) d'Aubigny
⁄ *delete last symbol* x *repeat last symbol*

Table 8: CCEs attributed to MA

CCE	occur.	cipher	CCE	occur.	cipher
o=*i*	525	MA	s=*p*	23	MA
ı=*e*	105	MA/LMB	+=*r*	20	MA
3=*s*	64	MA	4=*t*	17	MA
x=*[x2]*	44	MA	c=*c*	16	MA
∧=*n*	28	MA	∠=*t*	14	MA

Example 7.

l u o p r o m e s n e s



This example[19] should read *les promesses* but includes no less than four symbols wrongly taken

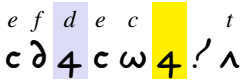
[17] For enclosure markings, see Section 5.2. Moreover, next to the name "Daubugny", there is a special mark apart from a cipher symbol. The same mark is found in several other cipher tables in SP 53/22 next to the names of the respective correspondents such as Châteauneuf (f. 22), Cherelles (f. 23), and Claude Hamilton (f. 33). The first clue that attracted our attention to Aubigny was a symbol later added to represent Cavaillon, who was a secretary to the Duke of Lennox (Bossy, 2001, p. 91, fn. 17).

[18] In the letters from Mary to Castelnau, ∠=*t* is always a corrected error, implying that the secretary spotted the mistake before finishing to write the symbol.

[19] F165, line 24

from the MA cipher, only the last being corrected by the secretary.

Example 8.



This[20] should read *effect*. The first error is not a CCE: We would expect the repeater symbol 4 from MC, to repeat the previous *f*. Instead, we have the wrong but similar-looking 4, uncorrected. The second error is a CCE: The second 4 is an MA symbol for *t* that has been deleted and corrected by the secretary.

4.3 The Nau brothers' cipher (NN)

The key (reconstruction) is found in TNA SP 53/23 no. 43.

This is a candidate we have considered as a potential source of CCEs, but dismissed after further analysis. We mention it here to illustrate that a set of potentially matching CCEs may not lead to a decisive conclusion that a specific cipher is causing CCEs. On the one hand, several likely CCEs, ∧=*c*, ı=*e*, ſ=*p*, and 4=*t*, could have been explained by the concurrent use of the NN cipher table. On the other hand, we also found that ∧ can be better explained by MB, and the other three are covered by MA. There is therefore not enough evidence to conclude that those CCEs in Mary's letters to Castelnau can be attributed to NN.

5 Corroborating our findings

We validated our hypotheses by statistical analysis, by examining the mentions of enclosed letters intended for other recipients, as well as by analyzing CCEs in the opposite direction, i.e., CCEs in Mary's letters to Beaton (encrypted with the MB cipher) that can be attributed to the MC cipher.

5.1 Statistical analysis

Our main assumption was that the secretary who enciphered the letters to Castelnau was affected not only by his level of concentration at the particular hour of the day or night but also by the specific mixture of other ciphers he may have used on that certain day (or week) to encipher letters to other recipients. Accordingly, to validate our CCE hypothesis, we expected certain sets of CCEs to occur together in particular letters. Thus, in one letter the errors induced by a certain cipher would predominate, resulting in an individual CCE "fingerprint" of each letter. Indeed, we found a high correlation between specific CCEs, and using a K-means clustering analysis, those correlations were found to fairly match the set of CCEs attributed to the MB, LMB, and MA ciphers (see Section 4). Those findings are described in more detail in Appendix A.

5.2 Evidence from mentions of enclosed letters to other recipients

Mary's secret letters, after being folded and sealed, were usually marked externally with a cipher symbol that stood for the recipient. This would allow Castelnau to identify which letters were addressed to him, and to whom the other letters should be forwarded. A ciphered postscript in the letter addressed to Castelnau, usually written by Nau, Mary's secretary, would list the names of the recipients of enclosed letters, together with the marking symbols identifying each recipient. During the relevant period 1578–1584, this external marking symbol was often the one representing the name of the recipient in the cipher used to encipher the letter. Thus, the letters to Michel Castelnau de Mauvissière were usually marked with the symbol K, which stands for *Monsieur de Mauvissiere* in MC (cf. Table 1).[21]

Based on those postscripts, we can confirm that the MB, LMB, and MC ciphers were indeed used concurrently with the MC cipher. For instance, F87 from Mary to Castelnau, enciphered with MC, ends with *Le chiffre cy-encloz marqué* 8̈ *est a Monsieur de Glasgo* (the enclosed ciphered letter marked 8̈ is for the Archbishop of Glasgow), and the marking symbol represents the Archbishop of Glasgow (James Beaton) in the MB cipher (cf. Table 2).

In the 57 letters from Mary to Castelnau, we find 20 mentions of "Monsieur de Glasgo" in conjunction with the symbol 8̈ from 2 May 1578 to 1 June 1583 and three further mentions of the same person associated with the symbol 8̄ from 3 July 1583 to 3 September 1583 (cf. Table 5). Both symbols correspond to James Beaton, the archbishop of Glasgow, in the MB cipher and the LMB cipher, respectively, and the latter marking symbol appears at the same time we estimate the MB cipher was replaced by the LMB cipher (cf. Section 4.1).

Similarly, the symbol for d'Aubigny in the MA

[20] F98, line 2

[21] See, for example, the verso of F87, `https://gallica.bnf.fr/ark:/12148/btv1b9059908w/f122.item`

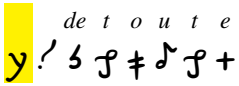
cipher, 0:, is used in association with enclosures for "le duc de Lenox" (i.e., d'Aubigny) twice in letter F225 from 31 December 1582 (cf. Table 7).

5.3 Reciprocal cross-cipher errors in letters from Mary to Beaton

Another way to corroborate our findings is based on the following idea: If such a large number of CCE occurrences in Mary's letters to Castelnau are caused by the concurrent use of the Mary-Beaton cipher, then we should also find the phenomenon in the opposite direction, i.e., in her letters to Beaton enciphered with the MB cipher, we should find cross-cipher errors that can be explained by concurrent use of the Mary-Castelnau cipher. And indeed, we found numerous occurrences of those kinds of errors in the letters from Mary to Beaton. For example, in the letter from 10 September, 1582,[22] we find two CCEs in the phrase *le principal autheur de toute l'entreprise*:

Example 9. *p r i n r i p a l*

Example 10. *de t o u t e*



Apparently, the wrong symbols for *c* (uncorrected) and for *de* (corrected) are taken from MC (although the dot in y:=*de* is missing).

6 New historical insights from cross-cipher errors

Based on our analysis of CCEs, we were able to gain highly interesting insights into Mary's secret communications:

- While the earliest of the newly deciphered letters from Mary to Castelnau is dated May 2, 1578, CCEs in letters from Mary to Beaton that can be attributed to MC (see Section 5.3) start to appear earlier. For instance, a significant amount of those CCEs is found in a letter to Beaton from 5 November 1577,[23] which shows that the Mary-Castelnau cipher was in use as early as this date.
- The CCEs attributed to the MA cipher show that Mary's secret communications with d'Aubigny were more extensive than previously known. The analysis of CCEs also shows that at the beginning of September 1582, Mary temporarily loses contact with d'Aubigny following the Raid of Ruthven and the pro-English coup in Scotland,[24] but that soon afterward the communication channel is re-established.[25]
- Previously, F308, one of the 57 letters presented by Lasry et al. (2023) could not be dated. We found several MB CCEs but no LMB CCE in this letter, allowing us to determine with high confidence that F308 was written before July 3, 1583.
- We can establish that the MB cipher between Mary and James Beaton was replaced by LMB in mid-1583 when MB CCEs disappeared and LMB CCEs began to show up, and based on the enclosure markings (see Section 5.2), we can even narrow down the timing of the replacement to between June 1 and July 3 of that year. This is remarkable as no letter encrypted with the new cipher before 1586 seems to have survived.

7 Conclusion

The systematic examination of ciphering errors and cross-cipher errors in particular in Mary's letters has led to valuable insights. Those errors were frequent, as she was communicating with multiple recipients, her secretaries using a different cipher for each recipient, and working under time pressure, so that multiple letters could be handed over to a trusted courier visiting Mary.

It is difficult to assess whether our method can be successfully applied to other collections of ciphered letters. Cipher secretaries around the world might also have had a stressful job, handling a high volume of communications using different ciphers at the same time, so those kinds of errors could have happened as well. If there is enough ciphertext material to analyze, it may be worthwhile to take a closer look at the errors they made, applying the techniques described in this paper.

[22] BL Cotton MS Caligula C III, f. 426, reproduced in Labanoff (1844, v, 309)

[23] SCA JB 2/4/3

[24] F229 of 2 September, 1582, in which hardly any CCEs can be assigned to MA. This fits in with her stating in the same letter that she would like to write to Aubigny if only she could (*j'eusse tres volontiers escript au duc de Lenox* [i.e., d'Aubginy] *si j'eusse eu aulcun moyen de ce faire*, line 10–11).

[25] E.g., F151 of 10 September 1582, with a significant number of CCEs that can be assigned to the MA cipher.

Acknowledgments

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

List of Abbreviations

BL The British Library

CCE Cross-cipher error

LMB Late Mary-Beaton cipher

MA Mary-Aubigny cipher

MB Mary-Beaton cipher

MC Mary-Castelnau cipher

NN Nau Brothers' cipher

SCA JB The University of Aberdeen, Scottish Catholic Archives, Archbishop James Beaton's Papers

TNA SP The National Archives, State Papers

References

John Bossy. 2001. *Under the molehill: an Elizabethan spy story*. New Haven: Yale University Press.

Alexandre Labanoff, editor. 1844. *Lettres, instructions et mémoires de Marie Stuart, Reine d'Écosse, 7 vols.* London: Charles Dolman.

George Lasry, Norbert Biermann, and Satoshi Tomokiyo. 2023. Deciphering Mary Stuart's lost letters from 1578-1584. *Cryptologia*, 47(2):101–202, March.

Satoshi Tomokiyo. 2023. Ciphers of Mary, Queen of Scots. http://cryptiana.web.fc2.com/code/mary.htm.

Appendix A. Clustering analysis

We first examined the correlations between certain pairs of CCEs, by computing their frequencies in each of the 57 letters, and computing the correlation according to the Pearson correlation formula. Let f_{ik} be the frequency of CCE C_i in letter k, and $\bar{f}_i$ the arithmetic mean of f_{ik} across all k. Then for a pair of CCEs C_i and C_j, the Pearson correlation coefficient r_{ij} is computed as follows:

$$r_{ij} = \frac{\sum_k (f_{ik} - \bar{f}_i)(f_{jk} - \bar{f}_j)}{\sqrt{\sum_k (f_{ik} - \bar{f}_i)^2 \sum_k (f_{jk} - \bar{f}_j)^2}}$$

Some correlations were exceptionally high between CCEs we had attributed to a particular cipher, and could not be the results of purely random errors.

Next, we aimed at grouping all types of CCEs into subsets of highly correlated CCEs, likely to have been caused by the concurrent use of the same cipher (other than the Mary-Castelnau cipher). For that purpose, we employed a common clustering algorithm, k-means,[26] based on the frequency of CCEs in the 57 letters from Mary to Castelnau. We limited our analysis to the types of CCEs that occur at least five times in the letters, i.e., the 71 CCEs given in Table 4.

The distance between two elements is (inversely) measured using the Pearson correlation coefficient mentioned above. The number of clusters – the k in the k-means – was found to be $k = 3$ using the "elbow method".[27] The results are shown in Figure 2.[28] We observe the following:

- Cluster 1 is mostly composed of errors attributed to LMB, which are strongly correlated, or moderately correlated with one another.
- Cluster 2 is mostly composed of errors attributed to MA.
- Cluster 3 is mostly composed of errors attributed to MB.
- A few errors attributed to MB are also associated with cluster 2, but with a low correlation.
- Some errors for which we did not identify a matching cipher table (UC) are assigned to the various clusters, but with a lower correlation.

Those results provide solid statistical evidence for the identification of the MB, LMB, and MA ciphers as being the root cause of almost all those errors.

[26] *See* Wikipedia, *k-means clustering*, `https://en.wikipedia.org/wiki/K-means\_clustering` (as of December 1, 2023, 19:00 GMT)

[27] *See* Wikipedia, *Elbow method (clustering)*, `https://en.wikipedia.org/wiki/Elbow_method_(clustering)` (as of December 1, 2023, 19:00 GMT)

[28] We show each such CCE type in the plot using a colored circle, whose size is proportional to the number of occurrences of the error (using a logarithmic scale). We also tie each CCE to the cluster identified by the k-means algorithm, and the length of the edge between the circle and the cluster is inversely proportional to the Pearson correlation between the frequencies (in the 57 letters) of the error and the average frequencies of the members of the cluster.

Figure 2: Clustering of the cross-cipher errors.