Dutch Cryptanalysis of Four American Diplomatic Codes in World War I

Florentijn van Kampen
iHub, Radboud University
Nijmegen – The Netherlands
florentijn.vankampen@ru.nl

Abstract

During the First World War, the Netherlands carefully maintained a neutral position. To guard this neutrality, the Dutch authorities monitored the activities of the belligerent surrounding countries. International telecommunications via telephone and telegraph were closely monitored and censored by censor bureaus.

In 2019, the Dutch intelligence and security service released a dossier about these censor bureaus to the Dutch National Archive. In that dossier, a previously unknown history of two groups of pioneering codebreakers based at the censor bureaus in Amsterdam and Rotterdam, was uncovered. In 2024, a first publication appeared about this dossier, with particular emphasis on how the local staff successfully broke German codes.

Additionally, the Dutch codebreakers successfully broke four American diplomatic codes between June and December 1918. This breakthrough enabled Dutch intelligence to monitor secret diplomatic traffic between American officials in the Netherlands and Washington during and after World War I.

This paper examines the systematic cryptanalysis of U.S. Department of State communications by Dutch codebreakers. Through analysis of original documents and surviving codebooks, it identifies the compromised diplomatic codes and places these findings in a broader historical perspective.

1 Introduction

During World War I (WWI) the Netherlands served as a hub for spies, smugglers, and diplomats. As a neutral country with a stable government, it provided an environment where people could discreetly conduct their activities. The Dutch government sought to maintain an overview of these activities and installed so-called censor bureaus at various telephone and telegraph offices to monitor and censor unwelcome communications.

The censor bureaus in Amsterdam and Rotterdam were particularly well placed because they served as international telecommunication hubs. The local military staff had full access to multiple international diplomatic communication channels, including coded communications. These two groups started analysing and breaking German codes and soon they were achieving cryptanalytical successes on a par with their more widely known and much larger counterparts in France, England and the United States (Jacobs and van Kampen, 2024b).

While it is well documented that British cryptanalysts successfully decoded American diplomatic telegrams during this period (Larsen, 2017) (Zegart, 2022), the fact that at the end of WWI, Dutch codebreakers had fully compromised American diplomatic traffic is as yet unknown. This article will show how the Dutch codebreakers were able - without international cooperation - to intercept, analyse and solve the American diplomatic codes in use at that time.

This article begins with an introduction to American diplomatic codes in WWI, followed by the methodology section. Section 4 provides an overview of the recently disclosed Dutch archive material on American Diplomatic Codes. Section 5 will successively present the four American diplomatic codes, how they were used, how the Dutch analysed and broke them and which historical sources were used to verify this. Section 6 will conclude this article with a number of observations and conclusions.

Proceedings of the 8th International Conference on Historical Cryptology, pages 65-75
Poznań, Poland, 16-18 June, 2025

2 American Diplomatic Codes in World War I

The United States has a long history of codes and ciphers protecting sensitive military and diplomatic communications. The standard work on United States Diplomatic Codes and Ciphers (Weber, 1979), traces this history all the way back to 1775 when Charles William Frederic Dumas, one of America's first secret agents, designed and dispatched the first revolutionary secret diplomatic cipher to Benjamin Franklin. It masked the correspondence between the Continental Congress and its foreign agents in Europe.

2.1 Codes and Ciphers

The U.S. Department of State protected its diplomatic communications using a system called "The Cipher of the Department of State". In cryptology, however, it is common to make a distinction between "codes" and "ciphers" because they refer to different cryptographic techniques. It is therefore important to explore these terms in a bit more detail. For this, we use the terminology used by David Kahn (Kahn, 1996).

A *cipher* operates on one, or sometimes more, plaintext letters. A so-called "substitution cipher" replaces these plaintext letters with enciphered equivalents. In the case of a "transposition cipher", the order of the plaintext letters is changed in some way. In both cases, this process is performed in a pre-arranged way, referred to as the cipher key.

In the case of a *code*, words, expressions or names are replaced by *code equivalents*, in most cases, a series of digits. These series of digits are called codegroups. The system that describes which word or expression should be replaced by which series of digits is called a a codebook. Codebooks may also include elements like syllables, full sentences, punctuation, or grammatical instructions.

So, the "cipher" of the U.S. Department of State is really a code using a one-part codebook. It is called a one-part codebook because both encoding and decoding are performed using a single alphabetical ordered list, an example of which can been seen Figure 2. Note the characteristic property that plaintext words and phrases that are alphabetically close to each also have codenumbers and codewords that are close to each other. This is unlike a two-part codebook where the words and expressions are not in alphabetical order, but in random order. Therefore, these codebooks need two parts, one part for encoding with the plaintext entries in alphabetical order and one part for decoding with the codegroups in numerical order. This article will focus on the various editions of this cipher that were in use by the U.S. Department of State between 1914 and 1918. The origin of these books go back to 1876 and start with the adoption of standardised communication rules of the telegraph.

2.2 The Cipher of the U.S. Department of State

The adoption in 1870 of the five character group, numbers and letters, as the standard "word" to be used in telegraph communications served as a basis for the diplomatic code of 1876, designed by by John. H. Haswell, Chief of the Bureau of Indexes and Archives (Hove, 2011). Cost and efficiency for telegraph transmission was an important design criterion for this new code, and now, every English word or phrase would translate to a five digit codenumber, or one standard "word" in telegraph terms.

The first version of this new "Cipher of the Department of State" was introduced in 1876 and was commonly referred to as "The Red Cipher" because of the color of its cover. Updated versions of this codebook would be introduced in the years ahead, also designated by colors: The Blue Cipher in 1899, The Green Cipher in 1910, the Gray Cipher in 1918 and the Brown Cipher in 1938.

This article is about the four diplomatic codes that were intercepted and solved by the Dutch codebreakers: the Cipher of U.S. Department of State in the editions of 1876, 1899, 1910 and 1918.

3 Methodology

As a neutral nation, the Dutch codebreakers worked in international isolation and had no knowledge of American codebooks, colors or systems. They created their own classification system for organizing the codes that they intercepted and analysed. However, to understand the historical significance of their work, it is necessary to determine which American codes the Dutch actually broke and therefore to link the Dutch classification system to the official American code-naming system. For each of the four American codes, this paper will present the results in two steps.

The first step consists of the relevant information from the Dutch archive material with a focus on key insights into cryptanalytic methods, breakthrough moments and technical observations.

The second step involves the validation of all available observations against original, historical American codebooks. For each of these validations, a specific example will be presented alongside its corresponding entry from the original codebook.

Each section on a specific code will provided a reference to the specific historical sources and their location. This is both to provide historical evidence for the the Dutch codebreaking results but also to facilitate possible further research.

4 Dutch cryptanalysis of American codes

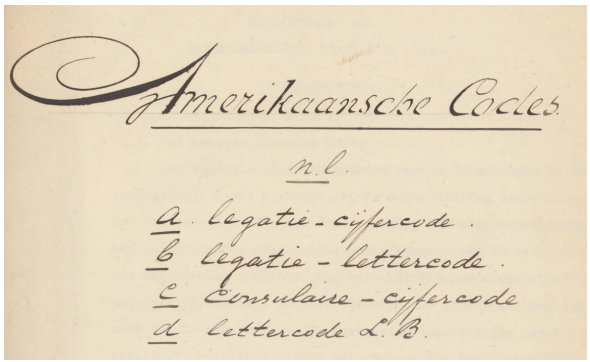


Figure 1: Title page of the Dutch collection of "notes" on American Codes (*Amerikaansche Codes* in Dutch). Source: Dutch National Archives.

The archive material of the Dutch codebreakers consists of two kinds of documents: "Reports" describing the day-to-day business of the censorship bureaus and "Notes" summarising the final cryptanalytical results of the codebreakers. The report[1] from Captain Nyweide, station chief in the Rotterdam telegraph office, describes the work on American codes. This report covers the timeframe between September 6, 1917 and September 30, 1918. On page 9 captain Nyweide writes[2]:

> Now that an English code had finally been broken[3], the American cipher telegrams were examined with even greater attention, and Lieutenant Vis of the Amsterdam bureau succeeded on June 14, 1918, in uncovering the American legation code, and on August 7 of the same year, he discovered the consular number code. Meanwhile, at the beginning of July, Lieutenant Berenschot managed to break an American letter code, to which Reserve Lieutenant de Vries subsequently dedicated all his attention and investigative skills. This code has also progressed to such a stage that the first telegrams could be submitted.[4]

The Dutch archive material includes a file[5] that contains four separate notes that deal with the topic of American codes. The title page of this file is shown in Figure 1. Not all of its contents has survived. The notes that are present are, in order of the American codebook versions that they cover.

- *"The Consular Number Code"*, Amsterdam, August 9, 1918, two pages. This note turns out to be about the solution of the Red Cipher of 1876.
- *"How the code was found"*, Amsterdam, June 18, 1918, two pages. As we will see below, this note is about the solution of the Blue Cipher of 1899 .
- *"Concerning the American Word Code"*, Rotterdam July 19, 1918, five pages. This note is about the solution of the Green Cipher of 1910.
- *"Concerning the American Letter Code B"*, Rotterdam , December 4 1918, five pages. This note is about the solution of the Gray Cipher of 1918.

In the section below, we will now explore the Dutch observations and analysis and link their results to the canonical names of U.S. Department of State codebooks.

5 The solution of four diplomatic codes

5.1 The Red Cipher of 1876

The codebook's central component consisted of almost 1000 pages of alphabetically ordered

[1] Dutch National Archives, 2.13.70 - 1939

[2] This translation from Dutch to English and all the subsequent translations from Dutch are by the author.

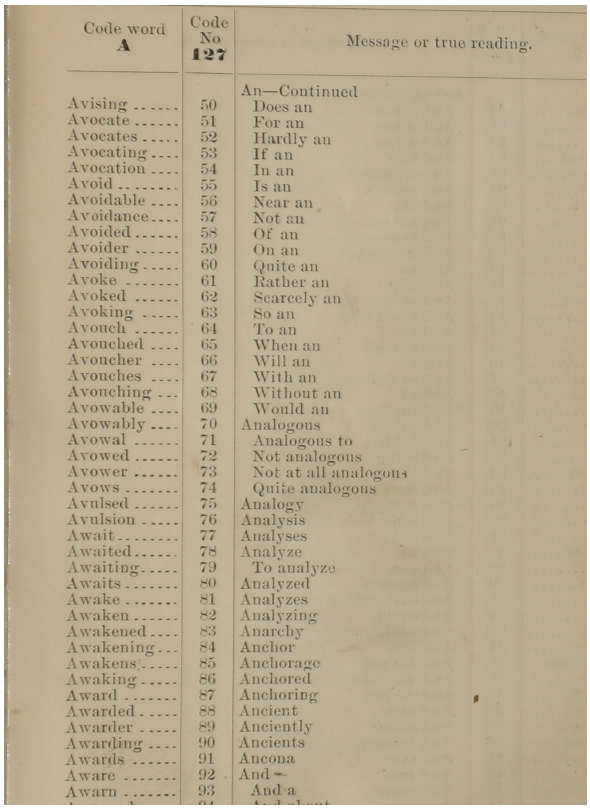
[3] Unfortunately, no information has survived about these results on English codes.

[4] Submitted here means: sent to the headquarters in The Hague to be used for intelligence purposes.

[5] Dutch National Archives, 2.13.70 - 1943

words and phrases. Each entry provided a choice between two encoding methods: A plaintext word or phrase could either be replaced with a code number or with a code word.

For example, the word "And" that is used in one of the archived diplomatic messages[6] is found on page 127 with number 92, as shown in Figure 2. The information from line 92 shows that the plaintext word "And" could be encoded with the codenumber "12792" or with the codeword "Aware". From the Dutch archives we can observe that most diplomatic communications that the Dutch intercepted from U.S. diplomats using this codebook was encoded with *codenumbers*.



Code word A	Code No 127	Message or true reading.
		An—Continued
Avising	50	Does an
Avocate	51	For an
Avocates	52	Hardly an
Avocating	53	If an
Avocation	54	In an
Avoid	55	Is an
Avoidable	56	Near an
Avoidance	57	Not an
Avoided	58	Of an
Avoider	59	On an
Avoiding	60	Quite an
Avoke	61	Rather an
Avoked	62	Scarcely an
Avoking	63	So an
Avouch	64	To an
Avouched	65	When an
Avoucher	66	Will an
Avouches	67	With an
Avouching	68	Without an
Avowable	69	Would an
Avowably	70	Analogous
Avowal	71	Analogous to
Avowed	72	Not analogous
Avower	73	Not at all analogous
Avows	74	Quite analogous
Avulsed	75	Analogy
Avulsion	76	Analysis
Await	77	Analyses
Awaited	78	Analyze
Awaiting	79	To analyze
Awaits	80	Analyzed
Awake	81	Analyzes
Awaken	82	Analyzing
Awakened	83	Anarchy
Awakening	84	Anchor
Awakens	85	Anchorage
Awaking	86	Anchored
Award	87	Anchoring
Awarded	88	Ancient
Awarder	89	Anciently
Awarding	90	Ancients
Awards	91	Ancona
Aware	92	And —
Awarn	93	And a

Figure 2: Original Red Cipher codebook, page 127. Entry 92 is "And". Source: the National Cryptologic Museum.

5.1.1 Dutch solution: Consulary Number Code

The Dutch codebreaker Lieutenant Vis wrote about this codebook in his note "about the Consular Number Code" dated, Amsterdam, August 9, 1918. His observations were brief and concise:

- This consular number code is alphabetical, ranging from 10000 "A" until 57371 "your".
- This code is used for the correspondence between the consulates

The Dutch considered the solution of this American code important, because of "America's increasing influence in European affairs". They noted that the representatives of the United States both in the Netherlands and in the Dutch Colonies found it impossible to maintain regular contact[7] with their principles in Washington. For time critical communication, the Americans had no choice but to use telegram or telephone communications, which were inherently more susceptible to interception.

The note includes four examples of intercepted encoded telegrams labelled "A, B, C, D" shown in Figure 3. The handwritten words above the codenumbers show some of the words that were recovered by the Dutch cryptanalyst: 12792 = "and", 31008 = "from", 57015 = "work" and 57371 = "your". The result for the word "And" can be verified in Figure 2.

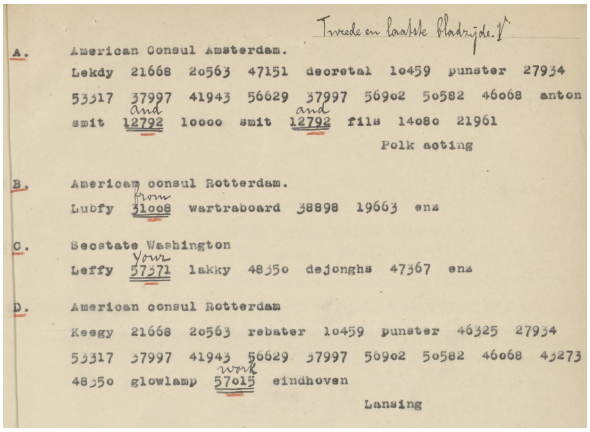


Figure 3: Analysis of the Consulary number code with Dutch handwritten annotations. Source: Dutch National Archive.

Careful examination of these telegrams show that each of these messages starts with a character sequence that is not related to the Red Cipher of 1876 (`Ledky`, `Lubfy`, `Leffy`, `Kecgy`, `Lakky`). It appears that these telegrams use an encoding system for dating telegrams which was designed in September 1900 and was first included in the Green Cipher of 1910 (Weber, 1979, page 247). This mix of codenumbers from 1876 and timestamps from 1910 is possible because these message were sent in 1918 after both editions were

[6] Telegram A in section 5.1.1

[7] This means by postal services.

Month	Day		Hour
	Tens	Units	
B January	O 0	B 1	B 1
C February	A 1	C 2	C 2
D March	E 2	D 3	D 3
F April	U 3	F 4	F 4
G May		G 5	G 5
K June		K 6	K 6
L July		L 7	L 7
M August		M 8	M 8
P September		P 9	P 9
R October		R 0	R 10
S November			S 11
V December			V 12
			i = forenoon noon y = afternoon midnight

Table 1: Date encoding system.

published. Apparently, these systems could be combined in daily usage. This date encoding system is shown in Table 1.

Since these four telegrams are the only examples of intercepted coded American diplomatic messages in the Dutch archives, it is worthwhile to demonstrate the decoding process and present the underlying plaintext messages. Using the date system from Table 1 for decoding the dates and the original American Red Cipher codebook from 1876 for the codenumbers and a few codewords, we can now decode the messages[8]:

Telegram A
American Consul Amsterdam `LEKDY 21668 20563 47151 DECRETAL 10459 PUNSTER 27934 53317 37997 41943 56629 37997 56902 50582 46068` *Anton Smit* `12792 10000` *Smit* `12792` *fils* `14080 21961 Polk acting`
July 26 3 pm: Confer colleagues report connections may not be accepted recommendations enemy Trading List or white list with specific reasons Anton Smit and A Smit and fils are they connected Polk[9] acting

Telegram B
American consul Rotterdam `LUBFY 31008 wartraboard`[10] `38898 19663` enz[11]
July 31 4pm: from War Trade Board you may certify etc.

Telegram C
Secstate Washington `LEFFY 57371 LAKKY 48350` *dejonghs* `47367` enz
July 24 4pm: your July 16 6pm : S de Jonghs reputation etc.

Telegram D
American consul Rotterdam `KECGY 21668 20563 REBATER 10459 PUNSTER 46325 27934 53317 37997 41943 56629 37997 56902 50582 46068 43273 48350` *glowlamp* `57015` *eindhoven* Lansing[12]
June 22 5pm: confer colleagues report recommendations enemy trading list or white list with specific reasons Philips[13] glowlamp works Eindhoven Lansing

5.1.2 Sources

Two sources for the original Red Cipher codebook of 1876 were used for this section. There is one copy in the collection of the National Cryptologic Museum. That version was used as the source for the image in Figure 2 and to decode the telegram messages. Another copy is available on microfilm in the U.S. National Archives and on-line[14].

5.2 The Blue Cipher of 1899

After more than 20 years of use, Haswell thought it was time to update the Red Cipher of 1876. He wrote that "(...) since European Black Chambers[15] operated freely, and probably possessed copies of the U.S. Department of State 1876 Cipher, it would seem good policy, after a service of twenty-five years to establish a new and improved system of telegraphic communication between the Department and its diplomatic Agents" (Weber, 2013).

This new version of codebook saw the light in 1899 and became known as the Blue Book[16]. The design and format was very much the same as the previous version, although the new version was expanded from 1,200 to 1,500 pages (Weber, 2013).

5.2.1 Dutch solution: Legation Number Code

The Dutch name for this system can only be deduced from the title page of the Dutch archive material of "notes" on American codes as seen in Figure 1. There are two number codes, the consulary

[8] The codenumbers and codewords are in `typewriter` font and uppercase, the inline plaintext is *Italic*.

[9] Frank Lyon Polk was counsellor of the U.S. Department of State. `https://en.wikipedia.org/wiki/Frank_Polk`

[10] Wartraboard = abbreviation for War Trade Board.

[11] enz is the Dutch version of etc (*enzovoort*)

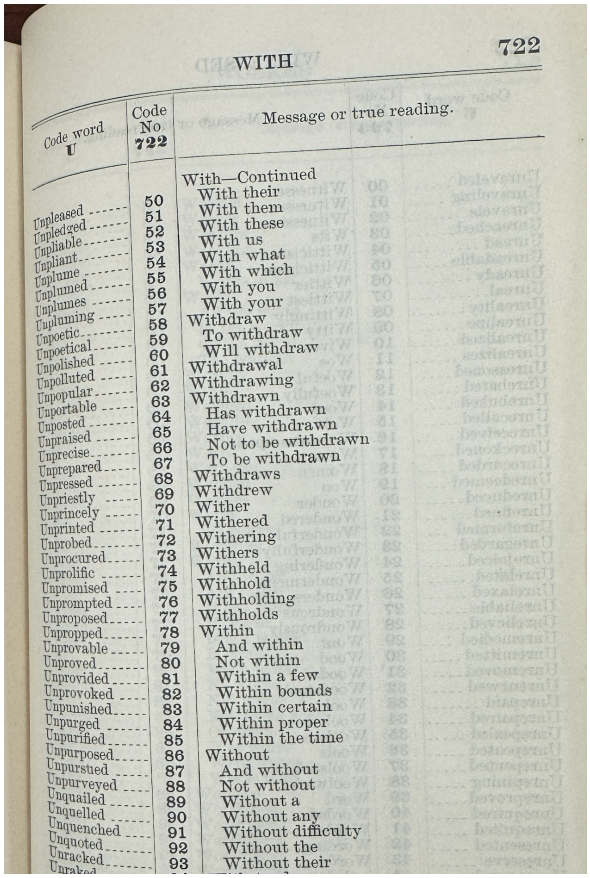
[12] Robert Lansing was the Secretary of State `https://en.wikipedia.org/wiki/Robert_Lansing`

[13] Philips is a well known Dutch electronics company that used to produce Lightbulbs

[14] For this article, a scan of the microfilm was requested. As a result, this scan is now available on-line for everybody to use: `https://catalog.archives.gov/id/417267710`

[15] See (de Leeuw, 2014) on the history of European codebreaking efforts organized in so called *Black Chambers*

[16] Although this codebook is often referred to as "The Blue Book", this paragraph was called "Blue Cipher" for the sake of consistency



WITH 722

Code word U	Code No. 722	Message or true reading.
		With—Continued
Unpleased	50	With their
Unpledged	51	With them
Unpliable	52	With these
Unpliant	53	With us
Unplume	54	With what
Unplumed	55	With which
Unplumes	56	With you
Unpluming	57	With your
Unpoetic	58	Withdraw
Unpoetical	59	To withdraw
Unpolished	60	Will withdraw
Unpolluted	61	Withdrawal
Unpopular	62	Withdrawing
Unportable	63	Withdrawn
Unposted	64	Has withdrawn
Unpraised	65	Have withdrawn
Unprecise	66	Not to be withdrawn
Unprepared	67	To be withdrawn
Unpressed	68	Withdraws
Unpriestly	69	Withdrew
Unprincely	70	Wither
Unprinted	71	Withered
Unprobed	72	Withering
Unprocured	73	Withers
Unprolific	74	Withheld
Unpromised	75	Withhold
Unprompted	76	Withholding
Unproposed	77	Withholds
Unpropped	78	Within
Unprovable	79	And within
Unproved	80	Not within
Unprovided	81	Within a few
Unprovoked	82	Within bounds
Unpunished	83	Within certain
Unpurged	84	Within proper
Unpurified	85	Within the time
Unpurposed	86	Without
Unpursued	87	And without
Unpurveyed	88	Not without
Unquailed	89	Without a
Unquelled	90	Without any
Unquenched	91	Without difficulty
Unquoted	92	Without the
Unracked	93	Without their

Figure 4: Original Blue Cipher codebook, page 722. Entry 86 is "without". Source: the National Cryptologic Museum.

number code and the legation number code. Since the consulary number code is the Red Cipher of 1876 and since the Blue Cipher of 1899 is the only other number code and since this code is in use by the American legation in The Hague, the Dutch probably referred to the Blue Cipher as the "Legation Number Code".

The proof that the Dutch codebreakers solved the Blue Cipher can be deduced from Lieutenant Vis's note: "How the code was found". His text takes us back to June 12, 1918. He explains how a sharp observation about two separate transmissions initiated the solution of this codebook. The first transmission, encoded in the Blue Cipher, is from Robert Lansing, the U.S. Secretary of State, to the American Legation in the Hague. The second transmission, in plain text, is from the Dutch representative in Washington, sent to the Dutch ministry of Foreign Affairs (a channel the Dutch codebreakers apparently also monitored!).

> On Wednesday afternoon, June 12, around 2 p.m., a long coded telegram addressed to "Amlegation[17] the Hague" from Lansing passed through the telegraph office in Amsterdam. In it, a few ship names appeared in plain language. From this it was possible to deduce the traditional word "Dampfer", this time "steamship".
>
> Two hours later, an approximately equally long telegram to "Celer[18] Haag" from the Dutch representative in Washington passed through. The same names also appeared in this telegram. Upon comparison after transmission, both telegrams turned out to be almost identical[19], allowing more than three hundred words to be recorded.
>
> The code found is in five-digit form and appears to be purely alphabetical: between 10425 "A" and 72286 "without" are the alphabetically lowest and highest groups found. Geographical names lie within this range. Nothing can yet be reported regarding numbers, letter groups, and personal names.
>
> Work on deciphering earlier telegrams can begin immediately, and due to the regularity and the experience gained with irregular codes, results can be obtained relatively quickly.

The notes of Lieutenant Vis show that this new code started with the number 10425 for the letter "A" and went all the way up to 72286 for "Without". Both these observation can be confirmed with the original codebook from 1899. The latter one can be seen in Figure 4. It shows page 722 of the Blue Cipher codebook with entry 86 for "without". This leads to a codenumber of 72286 in accordance with the analysis of the Dutch codebreakers.

5.2.2 Sources

There is one copy of the original Blue Cipher codebook of 1899 in the collection of the National Cryptologic Museum. That version was used as

[17] Short for "American Legation"

[18] The word "Celer" in the signature of telegrams represents the postal address of the Dutch Ministry of Foreign Affairs.

[19] This somewhat unusual situation could have been caused by American and Dutch officials each quoting the same newspaper.

the source for the image in Figure 4 and to verify the results obtained from the Dutch archive.

5.3 The Green Cipher of 1910

The U.S. Department of State published its next code, the Green Cipher, in 1910. This 1418 page volume introduced a completely new code design for enhanced security. In the absence of original codebook scans, this example relies on the information provided by (Weber, 1979, page 246). The pages of the Green Cipher codebook are divided into 5 columns:

Bab 102	Bab	102	Bab	A
Aa	ba	00	ab	a
Aachen	ca	01	ac	–bad
Aal	da	02	ad	–badge
Aalborg	fa	03	af	–band
Aam	ga	04	ag	– – of

In this example the page number is 102 and the header of this page shows the stem "Bab". That means that all the codewords for this page start with "Bab". The coding system makes a distinction between plaintext words and plaintext phrases. The codeword for *Aalborg* would be "Babfa" while the codeword for *A band* would be "Babaf". Careful observation shows that plaintext words are encoded with *stem* + *consonant* + *vowel* and plaintext phrases are encoded with *stem* + *vowel* + *consonant*. The notes left in the archive show that the Dutch codebreakers took particular professional pride to have noticed this structure.

5.3.1 Dutch solution: Legation Letter Code A

Lieutenant Berenschot published his analysis of the Green Cipher on July 19, 1918. In his note from the Rotterdam bureau he identified this code as "The American Word Code" as can been seen in Figure 5. However, the December 4, 1918, note discussing Letter Code B (examined in the next section on the Gray Cipher) refers to this code as "Letter Code A". This shows how the Dutch sometimes struggled with their own nomenclature of intercepted codes. With hindsight, the choice of "Legation Letter Code A", in combination with the subsequent "Letter Code B", actually makes much more sense because of the similarity between the Green code (Letter Code A) and the Gray code (Letter Code B).

Berenschot writes that the code consists of pronounceable words of five letters where every code-

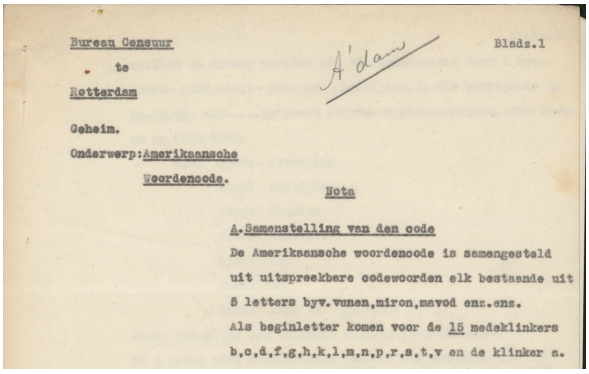

Bureau Censuur Bladz.1
te
Rotterdam
A'dam
Geheim.
Onderwerp:Amerikaansche
Woordencode.
Nota
A.Samenstelling van den code
De Amerikaansche woordencode is samengesteld
uit uitspreekbare codewoorden elk bestaande uit
5 letters byv.vunen,miron,mavod enz.enz.
Als beginletter komen voor de 15 medeklinkers
b,c,d,f,g,h,k,l,m,n,p,r,s,t,v en de klinker a.

Figure 5: Heading of the note *"Concering the American Word Code*. This copy was written in Rotterdam and was probably intended to inform and instruct the Amsterdam ("A'dam") bureau. Source: Dutch National Archives.

word is made of a "stem" of three letters and a suffix of two letters. Here we clearly see the basic structure of the codebook as described above. Berenschot gives a description of the methods and techniques he used to solve this code. The Dutch archival materials hold particular value because they preserve the codebreakers' actual thoughts and analytical processes. Therefore, the analysis is presented below as much as possible in Berenschot's words.

Analysis of message start

All telegrams sent on the same day begin with codewords that share the same root but have varying endings. These roots progress alphabetically from one day to the next. It could therefore be safely assumed that the first codeword represents the date of the telegram, particularly because American plaintext telegrams always start this way.

In American telegrams, the second codeword immediately following the date often includes the codewords `miron` and `vunen`, which are always followed by two new words starting with the letter "a". Since the "a" only appears as an initial letter in this situation, and since the construction of these codewords starting with the "a" deviates from the rest, it was deduced that these codewords represent numbers, especially because telegrams often begin with:

"Your / My (telegram) `number number`".

Assuming the simplest interpretation, the code is alphabetical, where `miron` means "My" and `vunen` means "Your" and the two codewords start-

ing with "a" represents "number number".

Telegrams Between The Hague and London

Many telegrams between The Hague and the American Embassy in London had similar message headers. The start of these messages, the message headers, were very similar in structure:

From	To	Header
The Hague	London	`fakaw rospo fibof fucku`
The Hague	London	`fakaw godmy fibof fucku`
London	The Hague	`fakaw fucku fibof rospo`

Analysis of known American number codes revealed that Sheldon and Frothingham were American envoys in London and that Edwards served in The Hague. This led to the following conclusion:

`fakaw`	for
`fibof`	from
`fucku`	Edwards
`godmy`	Frothingham
`rospo`	Sheldon

It was concluded that codewords ending with a vowel (like the "o", the "y" and the "u" in this example), represent personal names, while those ending with consonants represent common words (e.g., `miron` and `vunen`). Later, it was also discovered that these codes represent place names and abbreviations, for example: `pystu = Rotterdam`. (See table 2 for more examples).

Washington Messages

Plaintext telegrams were sent almost daily by Garrett, the American envoy in The Hague, to Washington to request visas for specific individuals. These telegrams always began with "Asks visa", followed by a name, and often include the phrase "nationality parentage Dutch".

The opening lines of the coded messages sent to Washington were now carefully compared with these plaintext telegrams in the hope that such visa application matters were also submitted in code. Having similar messages in both plaintext and code would greatly facilitate the solution of codewords. This hope was indeed fulfilled. There were quite a few telegrams intended for Washington that began with `cudam tisuk` followed by some words from the supplementary code (personal names) and in which the expression `mitel nipaf gigep` appeared. So that led to:

`cudam`	asks
`gigep`	Dutch
`mitel`	nationality
`nipaf`	parentage
`tisuk`	visa

This analysis revealed that the code for common words was indeed alphabetical.

Frequently Used Words

Common words like "for, from, to, and" had multiple codewords that were spread throughout the codebook:

`hugys, fakaw, filuc`	for
`hekyd, fibof, fadin`	from
`degyl, nuvog, somet`	to
`bykef, myses`	and

And with this, First Lieutenant Berenschot ended his report on the analysis of the Green Cipher. An entry from August 9, 1918 in the report indicates that the Dutch codebreakers were able to decode all Green Cipher messages.

5.3.2 Sources

The information on the Green Cipher of 1910 used in this sections comes from a special source. The U.S. National Archives have in their collection a file of recaptured material from the Germans after World War II. Apparently, the Germans had in their possession a copy of the Green Cipher as reconstructed by English codebreakers during WWI[20].

Table 2 shows a comparison of letter codes A (Green Cipher) and B (Gray Cipher). We can see in this table that for the Green Cipher the word "bank" corresponds to the codeword "cylub". This particular example can be seen in figure 6 that shows the page for codewords starting with "cyl".

5.4 The Gray Cipher of 1918

The U.S. Department of State introduced the Gray Cipher on March 22, 1918. At 1,173 pages, this volume was more compact than its 1910 predecessor. The Department designated this code for all consular telegrams, with specific instructions to exclude verbatim quotations from newspapers or government sources. This last instruction makes

[20] NR 3922 ZEMA121 36655A 19410000 Cryptographic Codes and Ciphers: United Kingdom Code Books for B3, Green Cipher, T and Sonder List B 1 Systems Recaptured from Germans in WWII. `https://catalog.archives.gov/id/2810671`

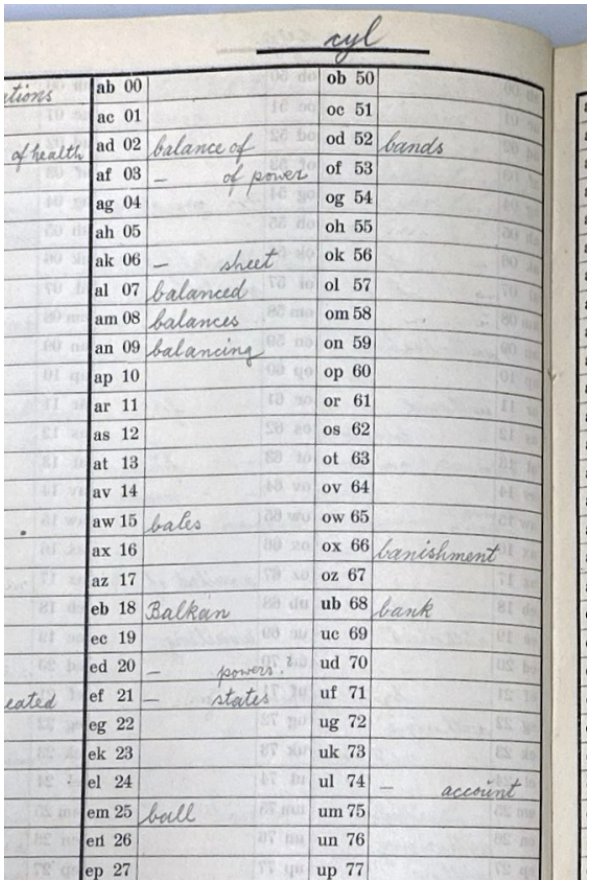


Figure 6: Inside page of the reconstructed version of the Green Cipher by the British codebreakers with “cylub” as “bank”. Source: U.S. National Archives.

a lot of sense because if the interceptor would know the verbatim underlying plaintext, for example from the latest newspaper, the solution of the codewords would be trivial. The design of the Gray Cipher was very similar to that of the Green Cipher (Weber, 1979). The well known U.S. Cryptographer William F. Friedman describes, somewhere between 1940 and 1946, the properties of the Gray Code as follows: ”This system has no confidential character whatever [sic] and is used for economy purposes” (Friedman, 1940).

5.4.1 Dutch solution: Legation Letter Code B

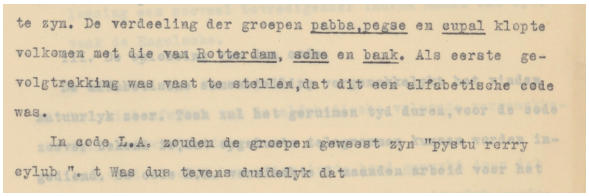
Reserve First Lieutenant de Vries from the Rotterdam bureau published his note on Letter Code B, the Gray Cipher, on December 4, 1918, only nine months after this new codebook was published. This date, falling several weeks after the November 11, 1918, armistice that ended WWI, demonstrates that Dutch intelligence continued monitoring and decoding American diplomatic traffic even after the war’s conclusion. This might very well be related to the fact that the Dutch were keen to follow the developments after the war in preparation for the peace negotiations in Versaille (Moeyes, 2014). Historical sources from the same time show that the Dutch codebreakers were breaking German codes to follow specifically the communications between the Americans and the Germans. (Jacobs and van Kampen, 2024a)

In the following section, First Lieutenant de Vries describes how he solved Letter Code B.

How the solution was found

A superficial examination of the telegrams written in this code already reveals that its structure matches that of Letter Code A (the Green Cipher). The identical formation of codeword stems and the similar alternation of endings — sometimes consonant / vowel, and at other times vowel / consonant — are immediately noticeable. The carelessness of the Americans was the reason the code was broken: the usage of the predictable sequence `pabba pegse cupal` was the breakthrough that would unravel the otherwise meaningless jumble of letters.

Captain Boomsma’s speculative remark that this sequence might mean “Rotterdamsche Bank” (Bank of Rotterdam) turned out to be spot on. The division of the groups `pabba`, `pegse`, and `cupal` perfectly aligned with the pattern of Rotterdam’s bank names. The first conclusion drawn was that this was an alphabetical code.


te zyn. De verdeeling der groepen pabba,pegse en cupal klopte volkomen met die van Rotterdam, sche en bank. Als eerste gevolgtrekking was vast te stellen,dat dit een alfabetische code was.

In code L.A. zouden de groepen geweest zyn "pystu rorry cylub ". t Was dus tevens duidelyk dat

Figure 7: Excerpt from the note about Legation Letter Code B showing the “De Rotterdamsche Bank” hypothesis in both Letter Code A and B. Source: Dutch National Archives.

Table 2 shows the Dutch phrase “Rotterdamsche Bank” or “Bank of Rotterdam” in both Letter Code A and B. This comparison shows that Letter Code B uses fewer letter groups than Letter Code A. The stem `pys` in Letter Code A corresponds with `pab` in Letter Code B and `cyl` corresponds to `cup`, and so on. The note contains a complete mapping between the stems of both let-

Word	L'code A	L'code B
Rotterdam	**pys**tu	**pab**ba[21]
sche	**rer**ry	**peg**se
bank	**cyl**ub[22]	**cup**al

Table 2: Words in Letter Code A and B

ter codes, part of which can be seen in Figure 8.

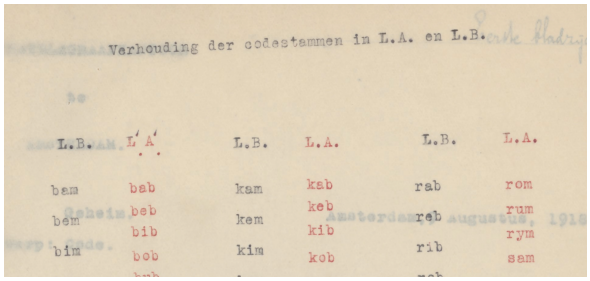


Figure 8: Mapping between stems for Letter Code A and B ("L.A." and "L.B."). Source: Dutch National Archive

5.4.2 Sources

Access to the original 1918 codebook that was used in this section was obtained through the Georgetown University Library. This library is home to the Russel J. Bowen collection with more than 16,000 titles including a large number of photocopied articles. The Bowen books cover a range of topics from military intelligence and spycraft all the way through spy fiction. Part of this collection is a unique original copy of the 1918 Gray Cipher codebook[23], as can be seen Figure 9.

6 Conclusions

Between June and December 1918, the Dutch codebreakers from Amsterdam and Rotterdam successfully broke all four diplomatic codes from the U.S. Department of State that were in use at that time: The Red Cipher of 1876, The Blue Cipher of 1899, The Green Cipher of 1910 and the Gray Cipher of 1918.

The fact that these different codes, with a similar design, were in use at the same time was actually an advantage for the Dutch. It was a deliberate choice by the U.S. Department of state as can be seen in introduction of the 1918 Gray Cipher

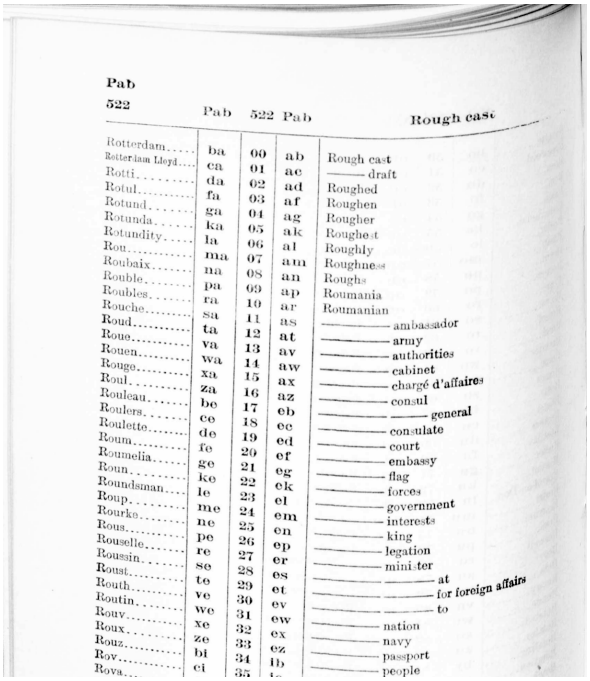


Figure 9: The original Gray Cipher codeook with the page for "Pabba" Rotterdam. Source: Georgetown University Library

codebook. Page XII reads: "This cipher (meaning the Gray Cipher) should never be used for the transmission of verbatim messages from individuals or from officials of a foreign Government. (...) Verbatim messages will be transmitted only in the Consular Red Code, The Diplomatic Blue, or such other codes as may be designated for this purpose by departmental instructions."

But the biggest advantage for the Dutch was the fact that the Americans did not use any advanced security measures on the channels the Dutch monitored, like super encipherment or similar techniques that were known at the time. The American officials were using straightforward, alphabetical one-part codebooks, greatly facilitating the Dutch in their codebreaking efforts.

In the note about the Blue Cipher, written on June 18, 1918, the author mentions a request for copies of coded telegrams dating all the way back to January 1, 1914. This is a striking historic example of the modern signals intelligence strategy of "store now, decrypt later".

As a result, by 1918, the Dutch had accumulated substantial knowledge of confidential and secret American diplomatic communications, both during and immediately after the war.

[22] See Figure 9 for source example

[22] See Figure 6 for source example

[23] https://wrlc-gu.primo.exlibrisgroup.com/permalink/01WRLC_GUNIV/1ok4lbs/alma991001599509704111

Acknowledgements

This research would not have been possible without access to original American historical codebooks. Special thanks are due to the following individuals who made extraordinary efforts to retrieve historical codebooks from special storage facilities and provide scans and images: Molly Kamph, Research Rooms & Augmented Processing Branch, National Archives at College Park; Jay Silvestre, Curator of Rare Books, Georgetown University Library and Robert J. Simpson, Museum Archivist and Librarian, National Cryptologic Museum, Maryland. Their efforts are highly appreciated. Finally, I would like to thank Prof. Dr. Bart Jacobs for his valuable ideas and comments on this article.

References

Karl de Leeuw. 2014. Books, Science, and the Rise of the Black Chambers in Early Modern Europe. *Geheime Post: Kryptologie und Steganographie der diplomatischen Korrespondenz europaïscher Höfe während der Frühen Neuzeit*, pages 87–89.

William F. Friedman. 1940. Communications Systems in Use by the Department of State. *NSA Friedman documents - Folder 467 - nr A67345*.

Mark T. Hove. 2011. *History of the Bureau of Diplomatic Security of the United States Department of State*. U.S. Dept. of State, Bureau of Diplomatic Security, Global Publishing Solutions.

Bart Jacobs and Florentijn van Kampen. 2024a. Geheimschrijverij bij de Meijer. In *Uiterst vertrouwelijk, Achter de schermen van de Nederlandse geheime diensten*, pages 127–135. Querido Facto.

Bart Jacobs and Florentijn van Kampen. 2024b. A new perspective on Dutch WWI codebreaking with its international ramifications. *Proceedings of the 7th International Conference on Historical Cryptology (HistoCrypt 2024)*.

David Kahn. 1996. *The Codebreakers : The Story of Secret Writing. Revised edition*. New York: Scribner.

Daniel Larsen. 2017. British codebreaking and American diplomatic telegrams, 1914–1915. *Intelligence and National Security*, 32(2):256–263.

Paul Moeyes. 2014. *Buiten Schot, Nederland Tijdens de Eerste Wereldoorlog*. de Arbeiderspers.

Ralph E. Weber. 1979. *United States Diplomatic Codes & Ciphers 1775-1938*. Transaction Publishers.

Ralph E. Weber. 2013. *Masked Dispatches: Cryptograms and Cryptology in American History, 1775–1900*. Center for Cryptologic History, National Security Agency.

Amy B. Zegart. 2022. *Spies, Lies, and Algorithms: The History and Future of American Intelligence*. Princeton University Press.