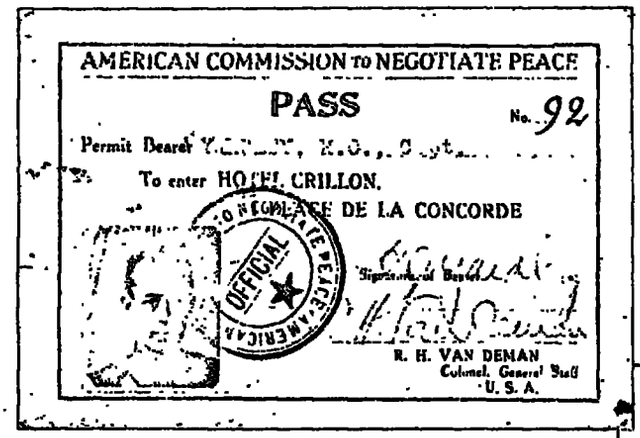
CODES



AMERICAN COMMISSION TO NEGOTIATE PEACE

PASS No. 92

Permit Bearer Y.C.M.C.A., U.S., Captain

To enter HOTEL CRILLON,

10 PLACE DE LA CONCORDE

OFFICIAL

Signature of Bearer

R. H. VAN DEMAN
Colonel, General Staff
U. S. A.

A Pass Used by Captain Yardley

CRYPTOGRAPHY demands a type of mind difficult to define. We call it "cipher brains," but can't describe it. To excel, a cryptographer should have not only great originality and imagination of a peculiar kind but years of experience, and no one in America had such experience in 1917. We never were able to formulate an intelligence test that would indicate the potentialities of a student. The most successful pupils, when put on their own responsibility, were, more often than not, useless except as clerks.

Later I was to have the opportunity of studying under the British, French and Italians, and learned that their experience had been the same. In the various Allied cryptography bureaus were thousands of devoted men and women, but among them all there were no more than a dozen with cipher brains.

When we began, I had, in the name of Colonel Van Deman, cabled to London, Paris and Rome, urging that pressure be brought on the Allies to send us expert cryptography instructors. I also asked for a few hundred examples of intercepted enemy code and cipher messages and all available exposition on their solutions. The latter were sent, but the former refused. Cryptographers could not be spared. While I was studying with the British, an English colonel told me that Captain Hitchings, their most brilliant cipher expert, was worth four divisions to Haig.

Military Intelligence Division, Section 8, or MI-8, had, by January, 1918, expanded into a great organization of five subsections: Code and cipher compilation, communications, shorthand—solution of intercepted shorthand documents—secret-ink laboratory, and code and cipher solution. In the last we had the double task of training students for our own use and others for the A. E. F. cryptographic bureau in France. This twin function handicapped us badly, for in justice to General Pershing, we felt honor bound to send our most promising graduates abroad. I regret to say that not more than two of all those dispatched to France distinguished themselves there, but this was not the fault of MI-8.

Judging from letters I found in the War College files, every other American was a dabbler in codes and ciphers. The writers either offered the Government their services or had new and indecipherable codes for sale. From among the former group, I hurriedly selected a few whose letters read most convincingly, and ordered the writers commissioned as captains.

Ciphers That Were Easily Deciphered

THE first of these captains to report was Dr. John M. Manly, a small, quiet-spoken scholar, who was head of the English department of the University of Chicago. To our great good luck, Captain Manly had the rare and precious gift of originality of mind, and developed into the most brilliant of all our cryptographers.

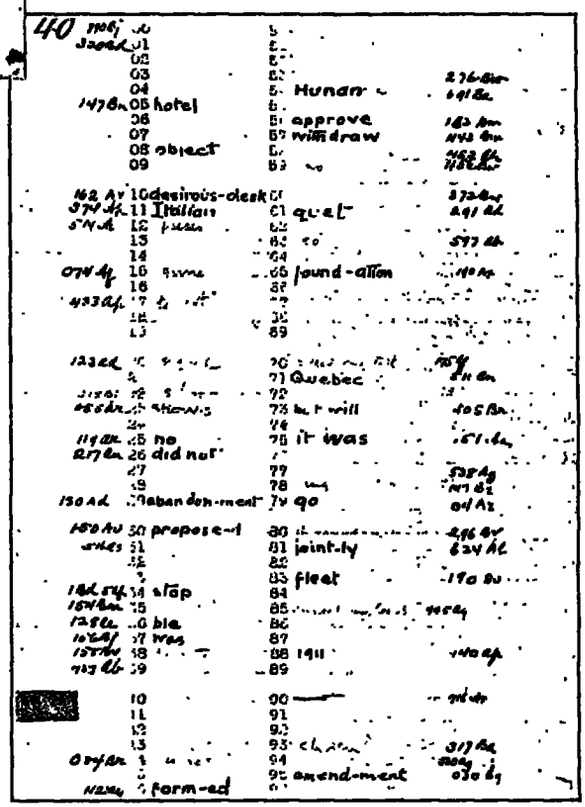
I had just begun to map out a course of schooling when my plans were upset by a memorandum from the State Department to the effect that the British considered our War Department's method of coding cablegrams seriously unsafe. The British reported, moreover, that the Germans were intercepting all Atlantic cable messages. The success or failure of the American Army depended upon the secrecy of its communications. The American people were allowed to know nothing of what passed between Washington and Pershing, but German submarines were lying alongside the Atlantic cables, copying the passing messages by induction and, in all likelihood, decoding them.

No wonder the memorandum startled the War Department. The Chief of Staff made a personal request for a prompt report. I learned that a copy of the War Department code book had been stolen in Mexico during the 1916 punitive expedition and that Germany presumably possessed a photographic copy. Whether the enemy had a copy or not, I soon demonstrated from actual tests that because of the technical construction of the code it could be solved easily without a key.

I doubt that my report was taken very seriously, but the opinion of London was held in respect and I was ordered to drop everything and revise the entire system of War Department codes and ciphers. I promptly chose one of the best men in the State Department code bureau and tempted him with a commission.

He took that detail off my hands. The compilation of codes and ciphers was, by General Orders, a Signal Corps function, but the war revealed the unpreparedness of this department. How much so is indicated by a talk I had with a higher officer of the Signal Corps who had just been appointed a military attaché to an Allied country. It was not intended that attachés should actually encode and decode their own telegrams, but as a part of an intelligence course they were required to have a superficial knowledge of both processes in order that they might appreciate the importance of certain precautions enforced in safeguarding our communications.

When the new attaché, a veteran of the old Army, appeared, I handed him a brochure and rapidly went over some of our methods of secret communication. To appreciate his attitude, the reader should



A Page From a "Skeleton Code" of 10,000 Words, Showing the British Foreign Office Code During Process of Decipherment by the Black Chamber. An Example of a Thoroughly Disarranged Code

understand that the so-called additive or subtractive method of garbling a code telegram, used by us during the Spanish-American War, is about as effective as the simple substitution cipher which we read as children in Poe's Gold-Bug.

He listened impatiently to my explanations; then growled: "That's a lot of nonsense. In the Spanish War we just added 1898 to all our figure code words, and the Spaniards never did find out about it."

He outranked me greatly or I might have added that we weren't fighting Spain this time. Amazing as it may seem, his attitude was characteristic, even at the Front. One of the young officers we graduated and sent to France brought the Signal Corps house of cards down with one puff. His first act was to induce his superiors to intercept our own Army's radio code and cipher messages used to transmit the most secret and vital orders. With no knowledge of the A. E. F. method of encipherment, the young officer solved these messages within a few hours.

The Germans, of course, were intercepting these messages constantly. If a student cryptographer could solve them, the German experts, with their long experience, without question had broken them down more quickly. Once the system was broken, the enemy could read every message as easily as the person to whom it was addressed.

By

Herbert O. Yardley

As it happened, the contents of this particular decipherment were so important and their secrecy so imperative that the General Staff was panic-stricken. For our student had read the disposition of our troops along the St.-Mihiel salient, the number and names of our divisions, the amount and location of the artillery support, even the precise moment at which the A. E. F.'s first great offensive would be launched. This, then, the enemy knew. It was too late to swoop down upon the German in a surprise attack. Either he would stand and was prepared for our every move, or he would retreat at once. The latter was true. The American offensive of September 12, 1918, was considered a triumph, but it represented only a fraction of what it might have been. Pershing pursued a retreating horde that fought only a rearguard delaying action. Of this whole episode we read but one sentence in a history of the

The Home Office of the British Censorship. A Sister Department to the Secret Ink Laboratory. In This Room Suspicious Letters and Communications Were Decoded

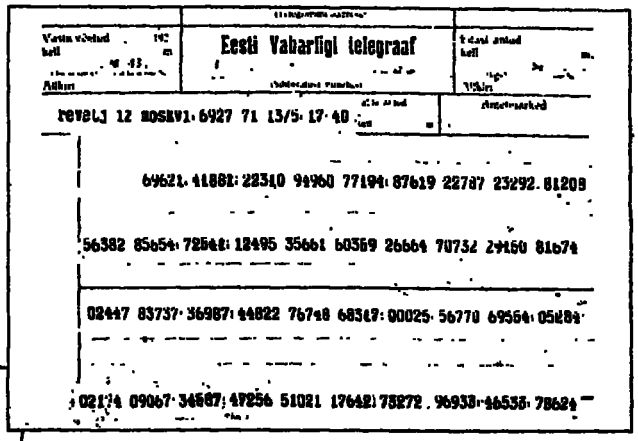
World War: "Despite all Pershing's precautions for secrecy in the St.-Mihiel sector, the Germans expected attack and began to withdraw."

Late in January, we received a communication from General Nolan, Chief, Intelligence Section, General Headquarters, France, inclosing two messages, made up of five-figure code groups, that were destined to make history. The letter read, in part:

These messages are sent at irregular intervals by POZ, a powerful station used for transmitting German official communications. So far as learned they are never addressed to any particular station and are never acknowledged. It is thought possible the messages are intended for German secret agents in some hostile or neutral country—possibly the United States or some other American nation—where the establishment of a high-power transmitting station is not permitted, but where the small equipment needed for interception can readily be concealed.

Due to lack of personnel and the necessity of putting most of our time on trench codes, we have been unable to give this particular cipher much attention. The French are much interested in its solution, but I am told neither they nor the British have made anything out of it up to now.

POZ was the powerful German station at Nauen. The same messages had been intercepted in America, for they were sent on an arc of 16,200 meters.



Eesti Vabariigi telegraaf

revsl 12 ROSEV1 6927 71 13/5 17 40

69621 41881 22310 94960 77194 87619 22787 23292 81209

56382 85654 72544 12495 35661 60369 26664 70732 24150 81674

02447 83737 36987 44822 76748 68347 00025 56770 69564 05284

02174 09067 34687 47256 51021 17642 73272 96933 46535 78624

An Original Soviet Code Message

Transmitted without address or signature and repeated as often as sixty times to our knowledge, it was clear that they were intended for a station at a great distance from Berlin.

We had good reason to believe that Mexico was the target; that they were answers to some curious messages emanating daily from some high-powered but unknown station in Mexico. The following is a typical example—note the repetition of each group:

HSI HSI HSI
DE DE DE
HSI HSI HSI
ATTENTION ATTENTION ATTENTION
WNCHL FTQBN CPDBQ TOCK IT
WNCHL FTQBN CPDBQ TOCK IT
WNCHL FTQBN CPDBQ TOCK IT
PERIOD PERIOD PERIOD
BREAK BREAK BREAK

The station call, HSI, which begins the message, was not assigned to any known or legal station. Listening in each night, we noted that the sender always gave a different call and never signed a call of his own. At first he sent every night at 10:30, 11, 11:30, and 12. Later he operated only three times a night, sending longer messages at irregular intervals. At first the messages were identical, with the exception of the number of times each code word was repeated. A little later there began to be slight variations in the code words. The signals were as strong at Dallas, Texas, as at San Antonio, and the back kick of his sending condensers could be heard at times at San Antonio, indicating great power and overloading of his condensers.

Tracing an Outlaw Transmitter

HIS slow and careful sending, his repetition of signals, were to make sure that every letter was correctly received—all implying that he was aiming at a far-away station. We heard him call combinations assigned to Russia, Great Britain, Germany, the United States, Brazil, and frequently to ship stations, convincing us that the call letters had no significance, but were used merely to enable the receiving operator to tune in his set to the incoming wave.

He received no answer, never allowed for the possibility of an answer, for at the end of each message, unless it was the last one of the night, he sent the character "wait," and on the last message the character "finish." All these facts made it evident that he operated on some agreed schedule.

Using a radio goniometer at different points along the Mexican border, we traced the strange station down. A goniometer is an instrument which shows the direction from which the wave emanates. Thus, if you take a reading at Laredo and draw a line on the map, then take a reading at Del Rio and draw another line on the map, the point on the map at which these two lines meet is the position of the sending station.

To the amazement of some, the lines converged at Chapultepec, the Mexican Government official station. Was Mexico in league with the Germans? Were they permitting Germany to use their official station for the transmission to Berlin of information German spies were running across the Rio Grande? Was Chapultepec receiving messages from Berlin and delivering them to the German Minister at Mexico City? There could be no other conclusion. The two code messages sent to us by General Nolan must have been destined for Mexico, and must have been secret indeed if Germany did not dare to send either address or signature.

The first thing we noted when we tried to solve them was that they were encoded in the same code. Such groups as 42036, 19707, 47230 occurred often in both. Our first task was to make a frequency table showing how often each group was repeated in both messages. This finished, we observed that the lowest number was 00308, the highest 55936, indicating that the code used had approximately 60,000 words and phrases. This is not unusually large; many contain 100,000 or more words and phrases.

(Continued on Page 141)

CODES

(Continued from Page 17)

We knew from experience that a code of only 10,000 words will express any language, except for unusual words, names of cities, of people, and so on, which may be spelled out either letter by letter in the code or, if the code so provides, by syllables. The code with which we were dealing, then, could contain only about 10,000 of the more common words, and the remaining 50,000 would be proper names, tables and common phrases and sentences. Probably it also had several different code groups for the first 1000 most common words. These we call variants. In English some of them would be "telegram," "you," "your." The variants for "you" might look something like the following in this code book:

You	40126	You	57754
You	09420	You	19227
You	13549	You	20648

Variants are used to confuse a hostile cryptographer. Repetitions are the weak spots of any code, and common words must be repeated often in a long message. In this case the hostile cryptographer must identify six code words for "you," instead of one. The code also may have several hundred nulls—meaningless code groups scattered at random through the message in order further to confuse the analyst.

The Problem of Frequency

Now then, if this code had approximately 50,000 code groups dealing with nulls, phrases, variants and entire sentences, there should be very few repetitions. If, however, the code consisted of single words only, and had neither nulls nor variants, phrases nor sentences, all words and phrases would be sent identically, piling up high frequencies and many repetitions.

Our frequency chart shows us that 42635 occurs sixteen times, three other groups eight times, and eleven other groups from four to six times in the POZ messages. The two messages contain a total of only 276 groups. Does it not seem unusual that in two messages of only 276 groups, one of the groups should occur sixteen times? Is it possible to write 276 words of text naturally and repeat one word sixteen times? Pick up a newspaper, select a story at random and count off 276 words, omitting "a," "of" and "the," for in telegraphic communication they seldom are required. The highest frequencies in the newspaper text I have selected are the words "be" and "in," each occurring only four times. "To" appears three times and no other word more than twice. Why this great discrepancy? What sort of code can we be dealing with that produces such high frequencies?

Seeking a Language Clew

The reader will recall the sensational Zimmerman-Carranza note which President Wilson read to Congress just before we entered the war—the note in which the German Minister for Foreign Affairs promised Mexico financial aid and the states of New Mexico, Texas and Arizona if she would declare war against the United States. This radiogram was deciphered by the British cryptographic bureau early in 1917, before our declaration of war. The Germans now were aware that their latest code between Germany and Mexico had been solved. They had to reëstablish secret communication. But how? Unless by submarine, it was virtually impossible for Berlin to deliver a new code book to her minister to Mexico. Yet we knew that we were dealing with a new code and were convinced that it was not dispatched to Mexico by submarine. What kind, then?

We made a careful analysis of the sixteen-time-repeated group, 42635, showing each recurrence and the prefix and suffix to each. We found immediately that, although frequently having the same prefix, 42635 nearly always had a different suffix. Obviously, then, it must be a termination, or ending, of some sort. It would do very nicely for "stop," but it scarcely seemed possible that there could be as many as sixteen "stops" in 276 words.

Before we went further we had to identify the language. The natural supposition was that it was either German or Spanish. Failing, we next tried English. To simplify this exposition, I do not go through the analysis. We shall assume that the message is in English. This will not be difficult to prove.

Scientifically constructed new codes such as Germany would use if free to do so give no such enormous repetition. This, it is fair to conclude, must be an alphabetical code—that is, the text words or meanings in the code are

FREQUENCY OF 5-FIGURE CODE WORDS IN NAUEN MESSAGES

1	00308	1	08732	1	19069	1	27918	2	34911	1	41137	1	49423
1	01156	5	08736	6	19707	6	28223	1	34004	1	41345	1	49513
1	01315	1	09415	1	19801	1	28530	1	34128	1	41363	1	49688
1	01426	1	09501	1	20344	8	28709	1	34642	2	42022	1	49642
1	01602	1	11230	1	20614	3	29540	1	35749	2	42235	1	49618
1	01930	2	11511	2	21109	2	29801	1	36060	16	42635	3	49632
3	02002	6	12201	2	21205	1	29913	1	36092	1	42740	1	49713
1	02206	2	13201	1	21209	1	30205	1	37013	2	43644	1	49853
1	02406	1	13503	1	21630	1	30334	2	37112	2	43827	6	50015
2	02502	1	13717	1	21737	1	30548	4	37950	1	44049	1	51316
2	02914	1	14210	1	23436	1	30949	2	37406	2	45021	4	51336
1	02915	2	14521	1	23827	1	31256	1	37803	1	45020	1	53025
1	02947	1	14913	1	23810	2	31259	1	38085	1	46027	1	54409
3	02957	1	15636	1	23145	1	31416	1	38303	3	46533	1	54628
2	03455	4	15726	1	23309	1	31511	1	38219	6	47329	1	55142
3	04186	1	15742	1	23634	1	32531	1	38606	1	48001	1	55320
1	04306	1	15929	2	24211	4	32045	1	38938	1	48234	2	55331
2	04628	1	16636	1	24235	3	33104	1	39408	1	49117	1	55841
2	05101	1	16733	1	24735	1	33186	1	39501	1	49126	2	56027
1	05732	1	17213	5	24909	1	33446	1	40252	1	49140	2	56034
2	06336	1	17256	1	26424	2	33556	1	40429	2	49223	1	56036
2	06420	1	18507	5	27141	6	33951	1	41111	2	49419		

arranged in dictionary order, and their figure equivalents arranged in numerical sequence.

When we turn back to our frequency chart we find several pairs of figure groups that are very close to one another, such as:

02914	21306	31258	49138	55934
02915	21309	31259	49140	55936

This is no coincidence. Each pair must imply words that are close together in the English language, such as:

21306 . . . Arrange   02914 . . . Arrive
21309 . . . Arranged   02915 . . . Arrived
31258. . . Name
31259. . . Named

If the book was alphabetically arranged—and it had to be—we should be able to place with reasonable accuracy in the alphabetical order of the English language the approximate position of each figure group. Any unabridged English dictionary would help us. We began with the assumption that 00308, the smallest figure group occurring in the two messages, equaled a word beginning with the letter A; and that 55936—the highest figure group—equaled a word beginning with Y or Z.

Discovering the Key

We found three groups beginning with 559; they must equal words occurring near the end of the dictionary and words that are close together. Words beginning with Z being infrequent, we could discard that letter. They must, therefore, begin with Y. Bear in mind that 55927 and 55934 each occurs twice. These must be common words beginning with Y.

The only dictionary I had before me was Appleton's New Spanish. I opened to the Y's in the English-Spanish section and searched the four columns of Y's. The word "you" seemed to be outstanding. I copied down the words that immediately followed "you" in the dictionary. They are: young, younger, youngish, youngling, youngster, younker, your, yours, yourself, youth, youthful.

If we assumed arbitrarily that 55927 equaled "you," we discovered that 55934, in this dictionary, would equal "your" and 55936, "yourself"—all very common words. We probably had correctly identified these three words and disclosed the encoding system used. Apparently these messages were encoded in an English dictionary: the first three figures indicating the page, the last two the line. Unlikely as this seems, our discovery could hardly be a coincidence.

We already had tabulated the range of the first three figures—003 to 559. Now we tabulated the last two figures to see whether this suggested any new ideas. The lowest number we found was 01, the highest 62. Why was it that with a total of 276 code words, the last two figures of none was higher than 62? What had happened to the endings from 63 to 99? There could be but one explanation. If the first three figures indicate the page of a dictionary and the last two the word or line number, we should not expect the last two figures to run very high, for there are seldom more than sixty words on a dictionary page. Then, too, why did 00 never appear? It would occur before 01 in a code book, normally. The reason is, of course, that the first word on each page of a dictionary must be indicated by 01; 00 would have no meaning. The use of a dictionary also would account for high frequencies and repetitions, for the same word always would be expressed by the same symbol.

We were, then, most certainly dealing with a dictionary. By spy or other method, the Germans had slipped word past the censors to their minister in Mexico to use a certain English dictionary for encoding and decoding dispatches.

We already had identified "you," "your" and "yourself" to our satisfaction without much effort. The complete solution of these famous messages was not quite so simple. But we had made an important discovery and had outwitted the German cipher experts after a few hours' labor. We continued. If we took Appleton's dictionary as a guide, we now could place the frequent group 42635 among the R's or S's. It might mean "stop." Yet "stop," again, could hardly occur so often in two short messages. We knew that it equaled some sort of termination. What?

How are the plural forms of words expressed in a code message encoded with a dictionary? By adding S, of course. S is a termination, the end of a word. Did 42635 mean S? It fitted perfectly into its alphabetical position and answered the requirements of high frequency and termination. So we assumed that S was correct and continued; though, eventually, 42635 proved to mean "stop," after all.

The beginning of the first message offered possibilities. It started off 19707, 21206, 31511, 31259. Now, 19707 occurred eight times in two messages and must be a common word. If we had the correct dictionary before us we would find the word on Page 197, Line or Word 7.

We lacked the right dictionary, but we could search for this word in any English dictionary of approximately 600 pages, by turning to Page 197 and looking in both directions at least ten pages. The most common word I could find between Pages 187 and 207 was "For," which occurred on Page 203, Line 11, or 20311. This was six pages farther advanced than the correct dictionary.

Success at Last!

This gave us a tentative formula, anyway. We added six pages to the next group—21206—giving us 21806. We found no clew on Page 218, but on Page 217, Line 20, was the word "German." Now we had two words—"For German." This began to look interesting. We should almost guess the next word—31511—but we stuck to the dictionary and our formula. Adding six pages to 31511, we had 32111. My dictionary told me that this word must begin with the letter M. What was it?

For German M—— That's right! For German Minister.

The next word—31259—occurred just three pages earlier in the dictionary than 31511. It must be some word between "let" and "mie." For German Minister Le-Mi? This message must be addressed to Mexico. What was the word? "Mexico," of course. For German Minister Mexico!

There were two methods of completing the solution of these messages. Several weeks of painstaking work would unravel their secret by this process, or we might go to the Library of Congress and search through every dictionary in print. If we were lucky, we should find the right dictionary.

It was the English-French half of Clifton's Nouveau Dictionnaire Français. Both of these historic documents were encoded in this volume. The first message reads, decoded:

FOR GERMAN MINISTER MEXICO BLEICHROEDER ANY TIME READY FOR LOAN NEGOTIATIONS. AT PRESENT REMITTANCE FROM GERMANY IMPOSSIBLE. MEANTIME FIRM PLACES TEN MILLION SPANISH PESETAS AT YOUR DISPOSAL GERMAN OVERSEA BANK MADRID. YOU ARE AUTHORIZED TO OFFER THIS PRELIMINARY AMOUNT TO MEXICAN GOVERNMENT IN NAME OF BLEICHROEDER FOR THREE YEARS, INTEREST SIX, COMMISSION HALF PER CENT, ON SUPPOSITION THAT MEXICO WILL REMAIN NEUTRAL DURING WAR. ALL GOOD ARRANGEMENTS LEFT TO BUSSHE. GENERAL STAFF POLITICAL SECTION BERLIN NUMBER HUNDRED.

A Spy in the Ranks

Mexico offered a bribe to remain neutral. America and the Allies had been straining every effort to induce Central and South American countries to declare war against the common enemy.

The other message read, decoded:

TELEGRAM JANUARY TWO AND TELEGRAPHIC REPORT S. ANTHONY DELMAR VIA SPAIN RECEIVED. PLEASE SUGGEST PRESIDENT [Mexico] TO SEND TO BERLIN AGENT WITH FULLER POWER FOR NEGOTIATION OF LOAN AND SALE OF RAW PRODUCT. DO NOT EMBROIL YOURSELF IN JAPANESE AFFAIR, BECAUSE COMMUNICATION THROUGH TOO DIFFICULT. IF JAPANESE ARE IN EARNEST THEY HAVE ENOUGH REPRESENTATIVES IN EUROPE FOR THAT PURPOSE. FOREIGN OFFICE BUSSHE. MACHINERY PLANS FOR RIFLE MANUFACTORY CAN BE PUT AT DISPOSAL. DETAILS OF *MACHINERY, TECHNICAL STAFF, AND ENGINEER* FOR AIRCRAFT COULD BE ARRANGED HERE WITH THE AUTHORIZED MEN OF PRESIDENT [Mexico] TO BE SENT BY HIM FOR NEGOTIATIONS ABOUT LOAN. WE AGREE PURCHASE ARRANGED BY CRAFT [Kraft?] IN JAPAN OF TEN THOUSAND RIFLES ETC. WISHED BY PRESIDENT. GENERAL STAFF POLITICAL SECTION NUMBER [?]

There was general excitement in Washington when these two messages were deciphered, for this would obviously open a new avenue to the United States for information not only of the intrigues of Germany but also of the true aims and intentions of Mexico and Japan. What would the decipherment of further messages reveal? A hundred instruments tuned in on the powerful Nauen wireless station to listen.

But Nauen suddenly became silent. Why? Because the news of our success had been flashed back to Berlin. There could be no other explanation, for Nauen resumed after a time, but with a new code. MI-8, with all its care in the selection of its personnel, had a German spy in its midst. A finger of suspicion now pointed to every cryptographer.