

# ORIGINS OF CRYPTOLOGY: THE ARAB CONTRIBUTIONS

Ibrahim A. Al-Kadit

To cite this article: Ibrahim A. Al-Kadit (1992) ORIGINS OF CRYPTOLOGY: THE ARAB CONTRIBUTIONS, Cryptologia, 16:2, 97-126, DOI: [10.1080/0161-119291866801](https://doi.org/10.1080/0161-119291866801)

To link to this article: <https://doi.org/10.1080/0161-119291866801>



Published online: 04 Jun 2010.



Submit your article to this journal [↗](#)



Article views: 349



Citing articles: 20 View citing articles [↗](#)

# ORIGINS OF CRYPTOLOGY: THE ARAB CONTRIBUTIONS\*

Ibrahim A. Al-Kadi<sup>†</sup>

ADDRESS: Advanced Electronics Company, Ltd., P. O. Box 90916, Riyadh 11623, SAUDI ARABIA.

ABSTRACT: Recently discovered old manuscripts show that the origin of cryptology, and the Arab contributions to it, are older and more extensive than previously thought. The word ‘cipher’ in European languages comes from the Arabic word *ṣifr*. The 9th-century Arab scientist al-Kindī is the author of the oldest known book on cryptology, antedating any other by more than 300 years. This paper highlights the specific contributions of some Arab cryptologists based on newly discovered documents that include books of al-Kindī, ibn Adlān and ibn ad-Durāhim. Factors behind the emergence and advancement of Arab cryptology are discussed. The discoveries reported in this paper push the frontiers of the history of cryptology back by about 500 years.

KEYWORDS: History of cryptology, Arab cryptology, al-Khalīl, al-Kindī, ibn Dunainīr, ibn Adlān, ibn ad-Durāhim, combinatorics, statistics, statistical cryptanalysis.

## 1. INTRODUCTION

In his celebrated history of cryptology, *The Codebreakers*, David Kahn asserts that “Cryptology was born among the Arabs” [9, p. 93]. Kahn attributes most of his account of Arab cryptology to the section on cryptology in the 1412 encyclopedia *Ṣubḥ al-A‘sha*, written by the Arab scholar al-Qalqashandī [2,9,13]. Nearly all of al-Qalqashandī’s writings on cryptology was taken from an earlier and more extensive book of ibn ad-Durāhim (1312-1361) entitled *Miftāḥ al-Kunūz fī Idāḥ al-Marmūz* (Treasured Key for Clarifying Ciphers), which was considered among the lost books of cryptology. This book was recently discovered and published [12].

---

\*This paper is based on two presentations given by the author at the Swedish Royal Institute of Technology in Stockholm and at the University of Linköping in Linköping SWEDEN during September 1990.

<sup>†</sup>The author is on temporary leave from his permanent position as an Associate Professor, Electrical Engineering Department, College of Engineering, King Saud University, P. O. Box 800, Riyadh 11421 SAUDI ARABIA.

Actually, Arab cryptology is much older than Kahn's account suggests, and Arab contributions to it are much more extensive than previously thought. Recently discovered old manuscripts show that the origin of cryptology is attributed to the 9th-century Arab scientist al-Kindī. His study of cryptology is the oldest available book on the subject, antedating any other by about three centuries, although his predecessor al-Khalīl (718-786) is reported to have written *Kitāb al-Mu'ammā* (The Book of Cryptographic Messages) about a century earlier. Unfortunately this book has not been found.

The aim of this paper is to report and discuss not-widely-known Arab contributions to cryptology. The study is largely based on the recent discovery of three old Arab manuscripts of cryptology including al-Kindī's book, a book by ibn Adlān, in addition to the long-reported but never-seen book by ibn ad-Duraihim. These books were edited and published (in Arabic) by a group of researchers at the Arab Academy of Damascus [12].

The next section briefly discusses the factors that led to the emergence and advancement of cryptology among the Arabs. The origin of the word 'cipher' is touched upon in Section 3. Section 4 introduces some known Arab cryptologists and their works in general. The following five sections (5 to 9) highlight the contributions of al-Khalīl, al-Kindī, ibn Adlān, ibn Dunainīr and ibn ad-Duraihim, respectively. The original contributions of al-Kindī, in particular, are discussed in some detail in Section 6.

## 2. FACTORS BEHIND THE ARAB ADVANCEMENT IN CRYPTOLOGY

Many factors contributed to the early emergence and advancement of cryptology in the Arab civilization. These can be grouped into five main related factors, each of which is discussed in the following.

### 2.1 Translation

Like many other rising civilizations, the Arabs considered knowledge as a shared human wealth. They started, shortly after their emergence from Arabia, to absorb the learning of older neighboring civilizations. One important means of their acquiring knowledge was through translation. In particular, the first three centuries of the Islamic civilization (700-1000 A.D.) witnessed intensive translation activities into Arabic of major works written in Greek, Indian, Farsi, Syriac, Armenian, Hebrew, Roman and other languages. The Caliph al-Ma'mūn (reigned 813-833) established the famous Bait al-Hikmah (The House of Wisdom)

in Baghdad as a national translation academy, state library and research center [1,3]. So much emphasis did the Arabs place on acquiring other cultures' knowledge that Greek manuscripts were obtained from the Byzantine Empire through peace treaties in return for maintaining the status quo. The main subjects of translation included philosophy, astronomy, mathematics, physics, chemistry and medicine. Some source books were written in dead languages forcing the Arab scholars to study them as ciphertexts and to classify their letters and symbols. They developed cryptanalytic techniques to try to understand those writings. In addition, some writings, were enciphered, especially in the fields of philosophy, theology<sup>1</sup>, witchcraft and chemistry<sup>2</sup>.

A number of Arab scholars have written on other languages and alphabets. Thobān al-Maṣry (? - 859) wrote a book entitled *Ḥall ar-Rumūz wa Brā'at al-Asqām fī Oṣūl al-Lughāt wal-Aqlām* (Solution of Symbols and Answers of Doubts on Principles of Languages and Scripts) [12]. The book of Aḥmad ibn Waḥshiyyah (? - 919), *Shawq al-Mustahām fī Ma'rifat Rumūz al-Aqlām* (Seekers' Joy in Learning about Other Languages' Written Symbols) contained 93 alphabets of different languages [12]. Nine centuries later, this book was translated from Arabic into English by the orientalist Joseph von Hammer and was published in both languages in London in 1806 under the title *Ancient Alphabets and Hieroglyphic Characters Explained* [2,9,7]. The French scholar Sylvestre de Sacy published a study about the book in Paris in 1810 that might have helped the famous archeologist J. F. Champollion in deciphering hieroglyphics [12].

## 2.2 Linguistic Studies

Being the language of the *Holy Qur'ān*, Arabic spread swiftly across vast areas of the world. It also served in different parts of the Islamic empire as a unifying official and scientific language. The first four centuries of the Arab civilization, in particular, witnessed intensive activities in studies of different linguistic aspects of Arabic, including phonetics, morphology, syntax, semantics, lexicography, grammar, prosody and computational linguistics. Such studies were prerequisites for the advancement of cryptography and cryptanalysis. Consequently many of the Arab cryptologists were also prominent linguists (see Sections 4-9).

---

<sup>1</sup>Some extremist sects of Islam cultivated cryptography to conceal their writings from the orthodox [9, p. 93].

<sup>2</sup>Chemistry was considered a precious secret because it was believed that through chemistry cheap metals could be converted to gold!

### 2.3 Administrative Studies

The need of the emerging Islamic state for administrative organization and services led the Arabs early on to concentrate on developing secretarial and management skills especially in the fields of composition and correspondence. The practical need to protect sensitive state affairs, particularly in the mails, was one of the main driving forces behind using, and therefore writing about, cryptography and cryptanalysis. The hands-on experience of practicing message concealment and codebreaking is evident in the available biographies of many Arab writers and secretaries who were considered among the most important personnel in government offices, especially in the Caliph's court [1,12]. Use of special codes by tax officials and accountants is well documented [9,10].

Many Arab administrative scientists wrote on cryptology either within their books of general administration or separately. Among those were Abū Bakr aṣ-Ṣūlī (? - 946) in his book *Adab al-Kuttāb* (The Secretaries' Manual), Is-ḥāq al-Kātib (~10 Cent.) in his book *al-Burhān fī Wōjōh al-Bayān* (The Authoritative Guide to the Art of Articulation), the 'father of sociology' Abdur-Raḥmān ibn Khaldūn (1322 - 1406) in his book *al-Muqaddimah* (The Introduction [to his history Encyclopedia]) which was called by the British historian Arnold Toynbee "undoubtedly the greatest work of its kind that has ever yet been created by any mind in any time or place", and al-Qalqashandī in *Ṣubḥ al-A'shā* [2,9,10,12,13].

### 2.4 Public Literacy

The Arab-Islamic civilization witnessed an expansion of reading and writing skills never approached by any previous civilization. This unprecedented level of literacy advanced cryptology in two ways:

- It increased the need to protect sensitive written messages. (In illiterate societies a mere writing of messages guaranteed their secrecy).
- The larger pool of educated people bred many qualified scientists in fields related to cryptology, such as translation, linguistics, mathematics, and administrative sciences.

### 2.5 Advanced Mathematics

The Arabs acquired the mathematical knowledge of their neighbors, especially the Greeks and Indians. During the first three centuries of Islam, in particular, they embarked on an extensive campaign of translation and absorption of major mathematical works of the time including the books of Euclid (*Elements*),

Ptolemy (*Tetrabiblos* and *Al-magest*), Archimedes, Aristotle, Diophantus (*Arithmetica*) and Brahmagupta (*Siddhānta* which the Arabs called *Sindhind*). By the early ninth century, Arab mathematicians started to produce major contributions of their own in the established fields of mathematics (arithmetic, trigonometry and geometry), but, more importantly, they pioneered in establishing two new fields, namely algebra and statistics, both of which are extremely important in modern cryptology. The glorious history of Arab mathematics cannot possibly be given a fair treatment here, but excellent references abound (see for example [1-6, 11, 14-16]). Here only aspects related to cryptology will be discussed.

Al-Khwārizmī (780 - 847?) is the most famous Arab mathematician. *Algorismus*, as he was known in Latin, was probably the best mathematician of his time [11]. He wrote two books on arithmetic and algebra, which played important roles in the history of mathematics. The original Arabic version of his first book (written around 820) on arithmetic is lost, but its contents have survived in different Latin translations. It was through those Latin versions of al-Khwārizmī's book that the (Hindu-)Arabic numerals were introduced into Europe (see Section 3). Al-Khwārizmī's second book, *al-Jabr wal-Muqābalah*, established the new field of algebra. It was translated in Spain into Latin by the Englishman Robert of Chester in the 12th century under the title *Algebra et Almucabala* [3, 11, 9]. From al-Khwārizmī and his works stem a number of common technical terms such as 'zero', 'cipher', 'algorithm', 'algebra' and 'Arabic numerals'. As Boyer wrote in his history of mathematics [3, p. 252], to al-Khwārizmī belongs the title "the father of algebra."

The Arabic foundation of algebra is widely acknowledged, but this is hardly the case with statistics. Historians of mathematics (see Boyer [3, p. 397] for example) attribute the first writings on probability and statistics to correspondence between Pascal and Fermat in 1654. In his recently discovered manuscript, al-Kindī gave the first description of statistical methods in cryptanalysis. He even explicitly required texts to be long enough to allow letter statistics to be meaningful (see Section 6). This is the world's first known writings in statistics, antedating those of Pascal and Fermat by about 800 years<sup>3</sup>. Statistical techniques were routinely used by Arab cryptologists after al-Kindī (see Sections 6 to 9).

---

<sup>3</sup>Historians of science and mathematics are urged to investigate the evidences of the new findings reported here.

### 3. ORIGIN OF “CIPHER”

Up to the thirteenth century (A.D.), the West was using Roman numerals (I, V, X, L, C, D, M), which were very cumbersome and made the simplest arithmetic operation (e.g. addition) extremely complicated. The reason is that the Roman number system is not a decimal system and has no symbol for an empty space (the zero). The lack of the concept of zero was the great stumbling block for the medieval mathematicians in Europe [11].

The decimal number system and the zero were originally developed in India. The Arabs became aware of the new numerals, in the early ninth century, through their translation of Brahmagupta's *Siddhānta* from Sanskrit into Arabic. The new numerals together with the zero were quickly adopted and put into widespread use throughout the Islamic empire from China in the east to Spain in the west. The Hindu-Arabic numerals were later introduced to Europeans through their contacts with the Arab civilization in Spain, where these numerals became known simply as the Arabic numerals. The spread of the new numerals throughout Europe started in the twelfth century with the Latin translations of al-Khwārizmī's book on arithmetic first by Robert of Chester and then by John of Seville [11]. In the thirteenth century, many authors helped to popularize the new Arabic numerals. Among those were the French Alexander de Villa Die (~1225), the English John of Halifax (~1200-1256), also known as Sacrobosco, through his book *Algorismus vulgaris*, and the Italian Leonardo of Pisa (~1180-1250), best known as Fibonacci, who studied under a Muslim teacher and traveled in Egypt, Syria and Greece. Fibonacci later wrote his book *Libre abaci*, in which he strongly advocated the use of Arabic numerals [3].

The new concept of zero confused people in the Middle Ages Europe [11]. In Sanskrit, the zero was called sunya or “empty”. The Arabs translated the Indian name into its Arabic equivalent *ṣifr*. When the Europeans became aware of the new digit, they no longer translated the name, but adopted the concept and the symbol as well as its Arabic name *ṣifr* which was transformed into the Latin words *cifra* and *cephirum* (as written by Fibonacci). The two Latin words worked their way into different European languages. In Italian *cephirum* was changed to *zefiro*, *zefro*, and *zevero* which was later shortened to *zero*.

The French formed the word *chiffre*, and also took over the Italian *zero*. The English used *zero* and *cipher* from which the word ciphering was some times used in the sense of ‘computation’. The German used *ziffer* and *chiffer* [11].

The concept of the *zero* or *ṣifr* or *cipher* was so confusing and ambiguous to common Europeans at first that a person used to say in arguments or discussions that he is “talking about something clear and comprehensible, and not about

some ambiguous and far-fetched thing like the cipher.” From such common confusion developed the use of the word “*cipher*” to mean concealment of clear meaning of messages or simply encryption [12]. In conclusion, the Arabic word *ṣifr* for the digit “zero” (0) developed into European technical terms that mean encryption.

	Scientist Name	Period	Comments	References
1	Al-Khalīl	718 - 786	see Section 5	[9,12]
2	Jābir ibn Ḥayyān (The Chemist) [known in Europe as Geber]	? - 815	<i>Ḥall ar-Rumūz wa Miftāḥ al-Kunūz</i> (Code Solving and Key to Treasures)	[1,4,12]
3	Thobān al-Maṣry	? - 859	See Subsection 2.1	[12]
4	Al-Kindī	801 - 973	See Section 6	[1,4,12]
5	Ibn Waḥshiyyah	? - >919	See Subsection 2.1	[2,9,12,15]
6	Moḥammad ibn Aḥmad ibn Kaisān	9th Cent.	Wrote a book on crypt- tography and prosody	[12]
7	Moḥammad ibn Aḥmad ibn Tabātabā	? - 934	Wrote a book on crypt- analysis preserved in the Ottoman Archive	[12]
8	Is-ḥāq al-Kātib	~10th Cent.	See Subsection 2.3	[12]
9	As’ad ibn Muhadhhab ibn Mammāti	1149 - 1209	<i>Khaṣā’iṣ al-Ma’rifah fī al-Mu’ammīyāt</i> (Guides of Knowledge on Cryptograms)	[12]
10	Ibn Adlān	1187 - 1268	See Section 7	[12]
11	Ibn Dunainīr	1187 - 1229	See Section 8	[12]
12	Ibn ad-Durāhim	1312 - 1361	See Section 9	[2,9,12,13]
13	Alī ibn Moḥammad ibn Aidamur al-Jaldakī	? - >1341	<i>Kanz al-Ikhtīṣāṣ fī Ma’rifat Asrār Ilm al-Khawāṣ</i> (Treasures of Specialization in Learning Secrets of Elites’ Science)	[12]
14	Al-Qalqashandī	1355 - 1418	See Sections 1 and 9	[2,9,13]

Table 1. Some known Arab Cryptologists.

#### 4. EARLY ARAB CRYPTOLOGISTS

Cryptography has been practiced to conceal messages since antiquity by different civilizations, including the ancient Egyptian, Chinese, Indian, Mesopotamian, Greek and Roman. But in none of them was there any cryptanalysis; only cryp-



tography existed [9]. Cryptology, the science of both making ciphers (cryptography) and breaking them (cryptanalysis), was born among the Arabs shortly after the rise of the Arab-Islamic empire. Many Arab scholars wrote on, and excelled in practicing, both branches of cryptology. Table 1 lists, in a chronological order, Arab cryptologists known to us who wrote on the subject. The space limitations, and the meager information available on some of them, preclude writing in detail on each of the individuals in the table, but we will highlight the contributions of some of them in the following five sections.

## 5. AL-KHALĪL

Abū Abdur-Raḥmān al-Khalīl ibn Aḥmad ibn ‘amr ibn Tammām al-Farāhīdī (718-786), who was among the greatest Arab linguists of all times, was also the world’s first great philologist. Among other achievements, his record of ‘firsts’ includes the following:

1. He was the first to conceive the idea of, and produce, a comprehensive dictionary. *Al-Ain*, a dictionary of Arabic vocabulary, is the oldest known Arabic book of its kind.
2. He was the first to study and classify the prosody of Arabic poetry.
3. He wrote the first known book on cryptology, *Kitāb al-Mu‘ammā* (The Book of Cryptographic Messages). This book is apparently lost [9,12].
4. He used combinatorics to count all possible Arabic words by considering all different combinations of letters with and without vowels. Appendix B presents a sample of al-Khalīl’s writings on this subject. His writings on the subject are eloquent and easy to understand.

Although considered the greatest linguist of his time, al-Khalīl was not known to be a mathematician. Yet his spontaneous use and apparent command of combinatorics (which may as well be the first such use in history) without much fanfare or elaboration raise a number of interesting questions that are worth consideration by historians of mathematics. Did the Arabs develop combinatorics and use it to such an extent that a nonmathematician linguist found it natural to employ it implicitly within a dictionary? Or was it the modesty of this great scholar that made him unaware of the importance of his ideas that could have contributed greatly to mathematics early in the 8th century?

[illegible]

الحمد لله - والحمد لله رب العالمين صلوات الله على من علم به محمد وآل بيته

[illegible]

Figure 1. The first page of al-Kindī's manuscript [12].

## 6. AL-KINDĪ

Abū Yūsūf Ya‘qūb ibn Is-hāq ibn aṣ-Ṣabbāh ibn ‘omrān ibn Ismaīl al-Kindī, the author of the oldest known book on cryptology, was born around the year 801 A.D. in al-Kufah, where his father was a governor. He grew up in Baghdad, where he got his education. He excelled in many fields including philosophy, medicine, astronomy, mathematics, linguistics and music. Al-Kindī, the first philosopher of Islam, was known as “The Philosopher of the Arabs”<sup>4</sup>. He became the caliph’s court physician, and was later named the first director of Bait al-Hikmah Library and Translation Academy, where he led the ongoing translation campaigns (see Subsection 2.1). Three of the caliphs al-Kindī served supported his research pursuits, but the unsympathetic attitude of the fourth caliph towards his philosophical thoughts brought about loss of prestige and personal fortune. His large famous personal library, the al-Kindiyyah, was confiscated and he was publicly beaten. He returned home sad and with a broken spirit; he later died in Baghdad in 873 at the age of seventy-two.

Al-Kindī left a wealth of knowledge in diversified fields. His writings were authoritative and encyclopedic, covering about 290 titles in different subjects. One of his books, *Risālah fī Istikhrāj al-Mu‘ammā* (A Manuscript on Deciphering Cryptographic Messages), is the oldest available book on cryptology, antedating any other by more than 300 years. A manuscript of it was recently discovered in the Sulaimaniyyah Ottoman Archive in Istanbul. It was edited and published by the Arab Academy of Damascus in 1987 [12]. Figure 1 shows the first page of al-Kindī’s manuscript.

The main contributions to cryptology of al-Kindī are briefly discussed in the following:

1. **Cryptanalysis Techniques:** Al-Kindī started by discussing principles of cryptanalysis. He divided cryptograms into two types—those of normal texts and those of poetry. For normal texts he specified four cryptanalysis methods: quantitative techniques (which he called ‘quantitative tricks’), qualitative techniques (‘tricks’), probable words (which he called ‘opening statements and praises’), and vowel-consonant combinations. For poetry he called for using prosody in addition to all the above methods of normal texts.
2. **Cipher Types:** Al-Kindī gave a tree-diagram classification of the major types of cipher systems, described each, and explained how to cryptanalyze

---

<sup>4</sup>The Arabs reserved the more prestigious title, “The Philosopher” for Aristotle, for whom they held much respect and admiration.

them. Figure 2a shows his tree diagram as it appeared in the manuscript, while a translation of it is given in Figure 2b. Al-Kindī also introduced super encipherment, but he did not discuss it in much detail. All he wrote about it is the following:

... [Super-encipherment] is composed of a number of simple [ones],  
by using two or more possible combinations of simple encipherments  
... Solving it is more difficult [12, pp. 224 and 234].

3. **Arabic Phonetics:** He classified Arabic phonetics into consonants (he used the term ‘silent’), long vowels (which he called “al-muṣṣawīyāt al-‘izām” = major louds) and short vowels (“al-muṣṣawīyāt aṣ-ṣighār” = minor louds).
4. **Arabic Syntax:** Al-Kindī gave a detailed study of Arabic syntax, in which he explained possible and impossible letter combinations. Figure 3 shows his study in the manuscript.
5. **Statistical Cryptanalysis:** The above contributions are remarkable, but the most important and original of al-Kindī’s contributions comprises the world’s first description and use of statistical techniques in cryptanalysis (or in any other application, for that matter). In his discussion of the quantitative techniques in cryptanalysis, al-Kindī clearly described how to use letter frequency statistics of the cryptogram to solve it. He also explained how to find these letter frequencies (by using a sample of the same language). Equally important and original is the condition he set on the length of the ciphertext under consideration; he required texts to be long enough to allow letter statistics to be meaningful. This concept, introduced by al-Kindī more than 1100 years ago, is extremely important in statistics today. The ideas of al-Kindī are remarkably clear in his manuscript. He wrote:

One way to solve an encrypted message, if we know its [original] language, is to find a [different clear] text of the same language long enough to fill one sheet or so and then we count [the occurrences of] each letter of it. We call the most frequently occurring letter the “first”, the next most occurring the “second”, the following most occurring the “third” and so on, until we finish all different letters in the cleartext [sample]. Then we look at the cryptogram we want to solve and we also classify its symbols. We find the most occurring symbol and change it to the form of the “first” letter [of the cleartext sample], the next most common symbol is changed to the form of the “second” letter, and the following most common



Figure 2a. Al-Kindī's tree diagram classification of cipher types as it appears in his manuscript [12].

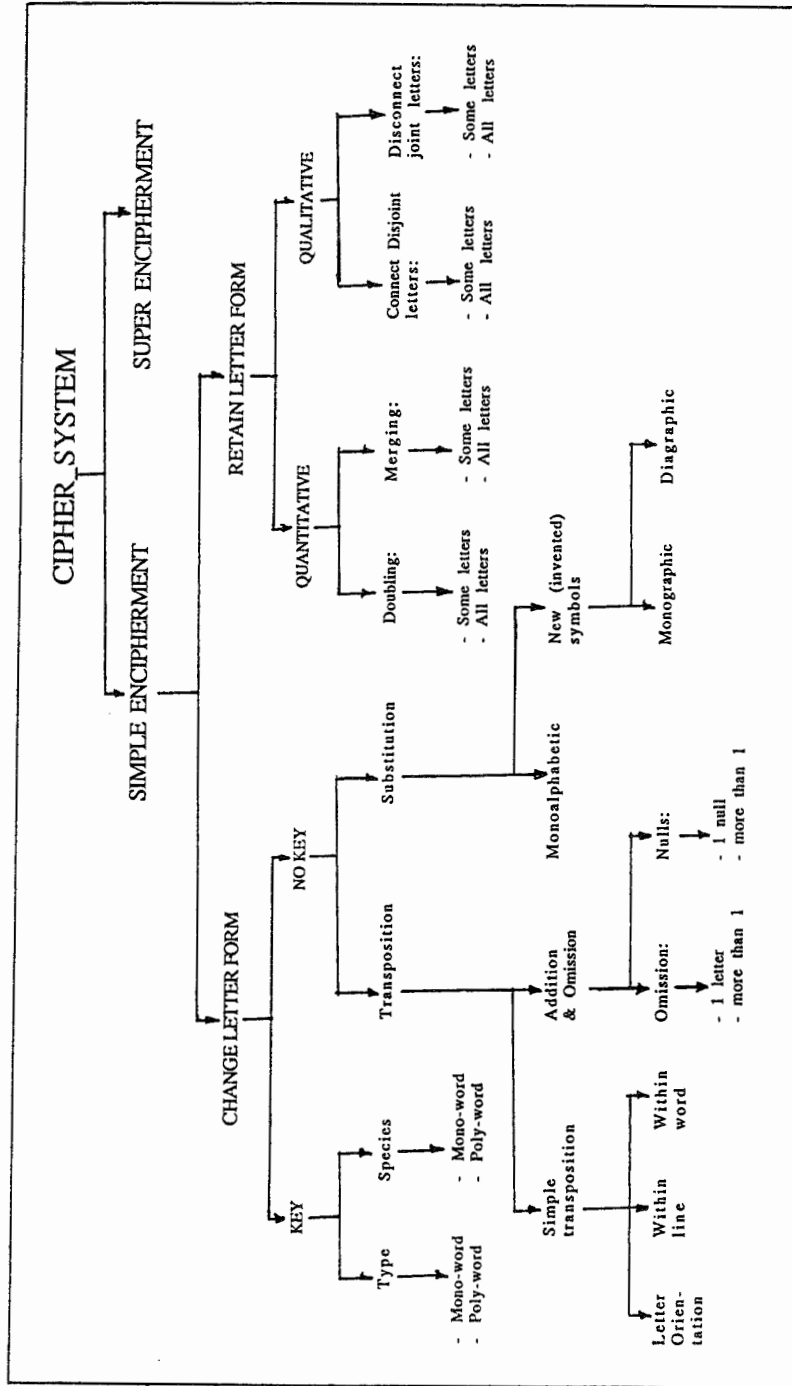


Figure 2b. A translation of al-Kindi's tree-diagram classification of cipher types.

symbol is changed to the form of the “third” letter and so on, until we account for all symbols of the cryptogram we want to solve.

It could happen sometimes that the cryptogram is too short to have all different letters. The high and low [frequency] counts will not be correct, for high and low counts are only correct in long enough messages to correspond to all places of frequent and rare occurrences so that if some letters are [too] few in one segment of the message, they will be [too] many in others. But if the cryptogram is too short, equivalence does not apply, letter ranks are not correct and [consequently] a second trick should be used to recover letters. Such a trick is qualitative which is ... [here al-Kindī wrote in detail on possible letter combinations in a language like ‘al’ in Arabic... etc.] [12, p. 216].

6. **Arabic Letter Frequency:** Later in his book, al-Kindī set out the first tally of Arabic letter frequencies. His sample consisted of “seven pages of Arabic” totaling 3667 letters. Al-Kindī’s results are shown in Figure 4, together with results of a much more recent study [8]. Not only did al-Kindī pioneer in computing letter frequencies, but he also analyzed his results and explained what might appear as a partial contradiction between the assertion at the beginning of his book, that vowels are the most common letters on all languages, and the fact that his results show that although the Arabic vowel ‘Alif’ is the most common letter, the other two vowels ‘Wāw’ and ‘Yā’ come only fifth and sixth after the consonants ‘Lām’ (l), ‘Mīm’ (m) and ‘Hā’ (h). He explained that vowels are underrepresented in written Arabic texts because short vowels are not usually written explicitly in Arabic scripts (see Appendix A). Al-Kindī wrote:

We said before that vowels are naturally the most common letters in all languages... Our results show that [the consonant] Lām (l) is more frequent in the Arabic language than [the vowels] Wāw and Yā. So are [the consonants Mīm (M) and] Hā (h). This is not a contradiction to what we previously said, because vowels appear in Arabic writing [only] if they are major [or long]. Minor [i.e., short] vowels do not appear in Arabic scripts... We have explained this in our [other] book [entitled] *fī Ṣinā‘at ash-Sh‘ir* [On the Art of Poetry]... [12, p. 236].

Al-Khalīl’s combinatorics and al-Kindī’s statistics are, most probably, the world’s first writings in the field of computational linguistics.

[illegible]

111



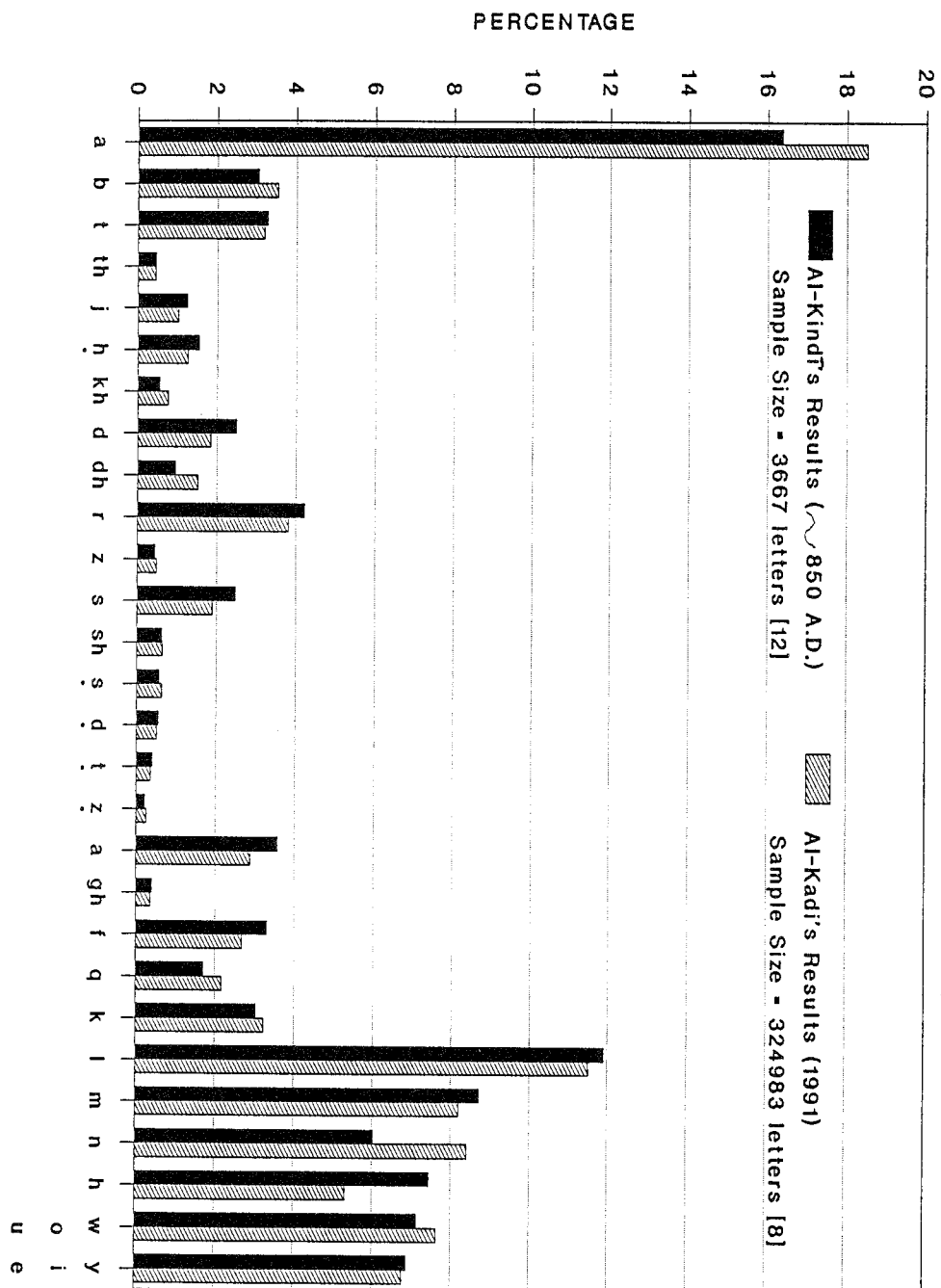


Figure 4. Al-Kindi's computed letter frequencies in Arabic, compared with recent results.

Thus it appears that al-Kindī is the father not only of cryptology but of statistics as well (see Subsection 2.5).

## 7. IBN ADLĀN

Afif ad-Dīn Alī ibn Adlān ibn Ḥammād ibn Alī al-Mouṣilī an-Naḥwi al-Mutarjim was born in Mouṣil (1187), educated in Baghdad, lived in Damascus for some time, and taught in Cairo, where he died in 1268 at the age of eighty-one. He was a poet who also excelled in grammar and linguistics, but he did not write many books. He left four, two of which were on cryptology. The first, *al-Mu‘lam* (The Told [Book]), is lost, while the second, *al-Mu‘allaf lil-Malik al-Ashraf* (The [Book] Written for King al-Ashraf), was recently discovered and published [12]. Figure 5 shows the title page and the first page of the document. The book is a manual or user-guide for cryptanalysis; unlike al-Kindī, ibn Adlān does not go into cryptography much. At the beginning, he discusses qualifications necessary for cryptanalysts, examples of simple substitution ciphers and Arabic letter combinations. The main body of the book is devoted to techniques of cryptanalysis where Ibn-Adlān gives 20 rules in 9 topics. In the conclusions, he works out a detailed example of solving a cipher as a training exercise for the reader. Ibn Adlān’s main contributions include the following:

- The concept of variable key for simple substitution. (He called it the ‘controller’, and used different verses of Arabic poetry as a variable key.)
- A detailed study of word spaces, identification of word starting points in a cryptogram, probable words and doubled letters.
- An explicit lower bound on the required ciphertext length. He wrote:

The text to be solved should be [at least] 90 letters long approximately because letters will have circulated about three times [ $3 \times 28 = 84$ ]. [12, p. 276]

- A classification of Arabic alphabet characters according to their frequencies into three groups—common (7), medium (11) and rare (10).
- An eloquent articulation of cryptanalysis methodology. In Rule 18, ibn Adlān clearly stated the consecutive steps of solving a cryptogram—moving from ciphertext to plaintext systematically from the unknown (ciphertext) through the set of all possible solutions, to suspected solutions, to probable ones and finally to the confirmed solution (the correctly solved plaintext).

الملك المليك المليك المليك  
في الدار الجديده  
التي بناها الملك المليك  
الذي هو الملك المليك  
على يد الملك المليك

مرسحروبن العرب المظلم  
الفلّاحم كاسم نوز وارباء راء كاسم اعنف  
فأثاف دال جاسم اعنف لاس زنا لاجم طالع ثا

114

He briefly but lucidly explained that one should consider all possible solutions, fix the sure parts of them, neglect the impossible combinations, and try all suspected solutions using elimination to arrive at the final solution. This is probably his most impressive contribution.

## 8. IBN DUNAINĪR

Ibrāhīm ibn Moḥammad ibn Dunainīr was born in Mouṣil (1187), lived in Cairo and Damascus, and died in Baniyas (1229) at the age of forty-two. He wrote at least one book on cryptology, entitled *Maqāṣid al-Fuṣūl al-Mutarjamah an Ḥall at-Tarjamah* (Clear Chapters' Goals on Solving Ciphers). The historical manuscript was recently discovered but has not yet been published [12]. In addition to his use of statistical technique in cryptanalysis pioneered earlier by al-Kindī (see Section 6), ibn Dunainīr was the first to describe an arithmetical cipher, or what he called "sentence reckoning", in which cleartext characters are converted into numbers and then some simple arithmetic operations are performed on them (see Figure 7 for some examples). He was also the first to describe some elementary cipher mechanisms, such as the chessboard [12].

## 9. IBN AD-DURAIHIM

Tāj ad-Dīn Alī ibn Moḥammad ibn Abdul'azīz ibn ad-Duraihīm was born in Mouṣil in 1312. A wealthy merchant who traveled between Damascus and Cairo, he held various teaching and official positions in both cities. Ibn ad-Duraihīm died in Qaus on his way to Eritrea in 1361 at the age of forty-nine. He excelled in theology, arithmetic, linguistics and cryptology. He wrote about 80 books on different subjects. Ibn ad-Duraihīm's most extensive work on cryptology is contained in his book *Miftāḥ al-Kunūz fī Idāḥ al-Marmūz* (Treasured Key for Clarifying Ciphers) to which most of al-Qalqashandī's writings on the subject in *Ṣubḥ al-A'shā* are attributed. The book of ibn ad-Duraihīm was considered lost [9, p. 95] until fairly recently when it was found, in Istanbul's Sulaimaniyyah Ottoman Archives, and published in a book by Mrayati et.al. [12]. Figure 6 shows two pages of the manuscript. Our discussion of ibn ad-Duraihīm's contributions will be brief and will concentrate on aspects not sufficiently covered by other authors [2,9,13].

Ibn ad-Duraihīm starts his book by briefly outlining the prerequisites of successful cryptanalysis. He then, also briefly, discusses the alphabets of 15 different languages including Greek, Farsi, Hindi and French. (Al-Qalqashandī included less than a third of this discussion in his book [13]). Like his predecessors, ibn

بسم الله الرحمن الرحيم  
 الحمد لله الذي ابتداء بخلق القلم . وصرفه في الكون فرقاه . وقسم<sup>لينة</sup>  
 واللغات المختلفة بين الأمم . العالم فلا يخفى عليه سر مكنتم . نحمد  
 على ما كشفه لنا من مكنون علمه وتوفيقا حمد من النعم . واشهد<sup>الله</sup> لا  
 إلا الله وحده لا شريك له شهادة من اليها التجا فرب اعظم . وشهد  
 ان محمدا عبده ورسوله الى العرب والعجم . ونجية المقرب حتى سمع<sup>صريف</sup>  
 نصريف الاقلام بما حكم وختم . بحله اولاً في الفضائل وبه<sup>الدين</sup>  
 ختم . فمدانا لا وضع الثم . وبين لنا مشكلات الحكم . صا<sup>وعا</sup> الله عليه  
 آله واصحابه الذين كل منهم في الهداية علم . سلاة دائمة مانتركا<sup>ت</sup>  
 ونظم . وبعد فانه كنت صنفت كتابا في وضع التراجم وحله<sup>ت</sup>  
 ايضا المبرم . في حل المترجم . ثم اختصرته ومرت عليه برهة  
 من الدهر ولم يكن الآن عند<sup>نفس</sup> نسخة . وسألني من يجب امتثال  
 ولا سبيل الى رده . فنظمت هذا القدر الكافي مما على ذهني من قواعد  
 هذا الفن وضوابطه وجعلت هذه الحاشية عليه موضحة لنظرة  
 مؤذنة ان شاء الله تعالى بفهمه . وتسمية مفتاح الكاوز في ايضا  
 المرموز والله تعالى اسأل الالاعانة والتوفيق وهو حبيبنا ونعيم  
 الوكيل . ان حل المترجم وايضا المعنى من اجل الفوائد  
 لا يستغنى عنه في اوقات تدعو الضرورة اليها وينتفع به في<sup>سخر</sup>

Figure 6a. The first page of ibn ad-Duraihim's manuscript [12].

[illegible]

Figure 6b. An example of cryptanalysis as it appears in ibn ad-Duraim's manuscript [12].

ad-Duraihim gives an extensive study of the linguistic characteristics of Arabic, which was preserved nearly completely in al-Qalqashandī's writings [2,13].

One of the most important contributions of ibn ad-Duraihim is his detailed and exquisite description of 8 systems of cipher with many variations and examples. His systems include the following:

1. Transposition with 24 variations and 31 examples including column transposition.
2. Substitution with 11 variations and 24 examples including the variable key substitution pioneered by ibn Adlān.
3. Letter addition (such as "ssifr" in "sifr") and omission (like "ifr" in "sifr").
4. Simple cipher mechanisms discussed earlier by ibn Dunainīr. An example of such mechanisms is a board with holes corresponding to letters of the alphabet and a string passed through holes of the intended message—a device described a millennium earlier by the Greek writer Aeneas the Tactician.
5. Arithmetical cipher. Ibn ad-Duraihim expanded on this method, pioneered by ibn Dunainīr, and gave 8 variations and 8 examples. Some of these examples are illustrated in Figure 7.
6. So-called letter-word substitution of four main types. One of these hides plaintext in the first (or second or last ... etc.) letter of words in innocent-looking but meaningful sentences. An example of which is hiding the plaintext Ali in the ciphertext sentence *pAint aLl pInk*.
7. Substituting for plaintext letters, under the control of a key, names from some species or families (e.g. animals, birds, cities, plants, etc.).
8. Use of some invented symbols for letters.

In his discussion of substitution cipher (No. 2 above), ibn ad-Duraihim analyzed the various possibilities and introduced the idea of what is now known as a 'Vigenère' table two centuries before Vigenère. Ibn ad-Duraihim's extensive and lucid study of cipher types was condensed beyond recognition (to about 10 percent) in al-Qalqashandī's book. Even though the presentation in *Ṣubḥ al-A'shā* was deformed by contraction, ibn ad-Duraihim's study of cipher types was commended by Kahn as remarkable and important [9, p. 98].

At the end, ibn ad-Duraihim gives a systematic methodology for cryptanalysis including the statistical techniques of al-Kindī and ibn Adlān. He falls short of them, however, in not giving actual frequency counts of Arabic letters but only

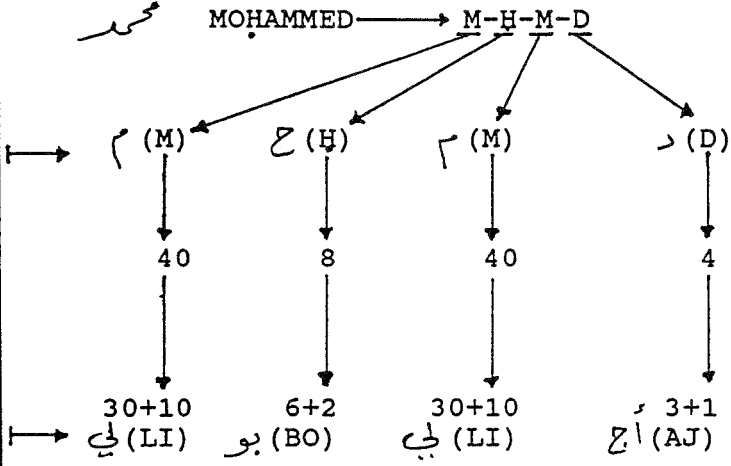
Plain Text:	 <p>MOHAMMED → M-H-M-D</p> <p>(M) → 40 → 30+10 → لي (LI)</p> <p>(H) → 8 → 6+2 → بو (BO)</p> <p>(M) → 40 → 30+10 → لي (LI)</p> <p>(D) → 4 → 3+1 → اذ (AJ)</p>				(Short vowels or diacritics do not appear in Arabic writings.)
Numbers:					
Sum Factors:					
Or:	20+20 كه (KK)	7+1 أز (AZ)	20+20 كه (KK)	2+2 بب (BB)	
Or Doubling x 2 =	80 ف (F)	16=10+6 يو (YO)	80 ف (F)	8 ح (H)	
Or Tripling x 3 =	120=100+20 قل (QK)	24=20+4 كر (KD)	120=100+20 قل (QK)	12=10+2 يب (YB)	
...					
and so on ...					Possible Cipher Texts

Figure 7. Some variations and examples of arithmetic cipher as given by ibn ad-Duraihim. (See Appendix A.)



their order of usage (rank). He also differs in that his ranking is based on the *Holy Qur'ān*. Next, ibn ad-Duraihim works out two detailed examples of solving sample cryptograms (see Figure 6b). Both the methodology and the two examples of ibn ad-Duraihim appear in their entirety in al-Qalqashandī's book, and were later described by Kahn as "the first exposition on cryptanalysis in history"<sup>5</sup>[9].

## 10. CONCLUSIONS

Cryptology originated, and developed to an advanced state, in the Arab-Islamic civilization (650 - 1400 A.D.). Among the factors that led to the early emergence and advancement of Arab cryptology were translation, linguistic studies, administrative sciences, public literacy and advanced mathematics. The origin of cryptology is much older, and Arab contributions to it are much more extensive, than previously reported. The new findings, reported in this paper, push back the frontiers of the known history of cryptology by more than 500 years. These findings are largely based on 8th century and later manuscripts written by the prominent Arab cryptologists al-Kindī, ibn Adlān and ibn ad-Duraihim. The contributions of these scientists, in addition to ibn Dunainīr, to cryptology are remarkable and important. Al-Kindī's contributions, in particular, are original and pioneering. He is the father of cryptology.

Among the new findings, reported in this paper, is the fact that the first writings in statistics are due to al-Kindī about 800 years before Pascal and Fermat. In addition, the great Arab linguist, al-Khalīl, used combinatorics more than 1200 years ago.

There are, however, some unanswered questions on Arab cryptology that cannot be cleared up yet. Arab cryptologists do not mention unbreakable ciphers, unlike many later inventors of cipher systems who claimed, and genuinely believed, that their systems were unbreakable.

It should be kept in mind that the sudden and quick decline of Arab civilization may have led to the loss of many books on the subject, and may have set back cryptology [9, pp. 98-99]. That decline seriously delayed the progress of human knowledge in many other fields as well.

---

<sup>5</sup> Al-Kindī's exposition was 500 years earlier!

## APPENDIX A THE ARABIC SCRIPT

The basic Arabic alphabet consists of a set of 28 letters. The Arabic script is extended to some 90 elements by additional shapes, marks and short vowels formally recognized in Arabic morphology. Arabic, written from right to left in elegant cursive form, does not differentiate between upper and lower case. Table A-1 shows Arabic letters, together with their transliteration as used in the etymologies and the numerical values of the different letters as adopted by early Arab linguists and cryptologists. In standardizing these values, early Arab scholars adopted the “abjadi” order of letters, which is different from the alphabetical order shown in the table. They gave the first nine “abjadi” letters consecutive numbers 1 to 9; the following nine letters were assigned numbers 10, 20, 30, . . . , 90; letters 19 through 27 were given numerical values 100, 200, 300, . . . , 900; the last letter in the “abjadi” order, which is ghayn, is assigned the numerical value of 1000.

Arabic letters are all primarily consonants; three of which (alif, wāw and yā) are also used to represent long vowels. Full indication of short vowels or phonetic effects are not usually provided in normal Arabic scripts except in important texts or where confusion may arise. When they are provided, it is by means of a number of strokes and diacritical signs which are similar to accent marks in some European languages. Those signs are written above or below characters to mark short vowels and emphasize or loosen a letter’s sound. They can also be mixed to produce composite phonetic effects.

## APPENDIX B AL-KHALĪL’S COMBINATORICS

The great Arab linguist al-Khalīl, the author of the oldest dictionary of Arabic words, used combinatorics in the eighth century to count all possible Arabic words by considering all different combinations of letters with or without vowels. In his dictionary, *Al-Ain*, al-Khalīl wrote:

If you want to exhaustively know all of the Arabic language double-letter words, either meaningful or not, which the Arabs either used or rejected, such as *qd*, *km*, *an* . . . etc., take the [Arabic] alphabet letters which are 28, then multiply them with each other to get 784 [=  $28^2$ ]. A single letter is not a word. If you take two letters [without reversal], you get 392 [=  $784/2$ ] such as *dm* and the like. If you reverse [the two letter positions] it comes back to 784, 28 of which have identical letters

Letter Name <sup>1</sup>	Transliteration	Arabic Form				Numerical Value
		Letter..... stand alone <sup>2</sup>	Joined..... from right <sup>3</sup>	Joined..... from left <sup>4</sup>	Joined..... from left and right <sup>5</sup>	
Alif	‘,a <sup>6</sup>	ا	آ			1
Bā	b	ب	با	با	با	2
Tā	t	ت	تا	تا	تا	400
Thā	th	ث	ثا	ثا	ثا	500
Jim	j	ج	جا	جا	جا	3
Hā	h	ه	ها	ها	ها	8
Khā	kh	خ	خا	خا	خا	600
Dāl	d	د	دا			4
Dhāl	dh	ذ	ذا			700
Rā	r	ر	را			200
Zā	z	ز	زا			7
Sin	s	س	سا	سا	سا	60
Shin	sh	ش	شا	شا	شا	300
Ṣād	ṣ	ص	صا	صا	صا	90
Ḍād	ḍ	ض	ضا	ضا	ضا	800
Ṭā	ṭ	ط	طا	طا	طا	9
Zā	z	ظ	ظا	ظا	ظا	900
‘ayn	‘	ع	عا	عا	عا	70
Ghayn	gh	غ	غا	غا	غا	1000
Fā	f	ف	فا	فا	فا	80
Qāf	q	ق	قا	قا	قا	100
Kāf	k	ك	كا	كا	كا	20
Lām	l	ل	لا	لا	لا	30
Mīm	m	م	ما	ما	ما	40
Nūn	n	ن	نا	نا	نا	50
Hā	h <sup>7</sup>	ه	ها	ها	ها	5
Wāw	w	و	وا	وا	وا	6
Yā	y	ي	يا	يا	يا	10

Table A-1. (Opposite) Arabic letters, their forms, transliteration, and “abjadi” values.

## Notes:

1.  $\bar{a}$ ,  $\bar{i}$ , and  $\bar{u}$  are pronounced like *a* in *father*, *i* in *machine* and *u* in *rude*, respectively.
2. The form of the letter when it stands alone.
3. The form of the letter when it is joined to the preceding letter only.
4. The form of the letter when it is joined to the following letter only.
5. The form of the letter when it is joined to both the preceding and following letters.
6. Alif represents no sound in itself, but is used principally as an indicator of the presence of a glottal stop (transliterated ' medially and finally; not transliterated when initial) and as the sign of a long *a* (transliterated  $\bar{a}$ ; see Note (1) above).
7. When 'o' has two dots above it ( $\ddot{o}$ ), it is called *ta marbūtah* and, if it precedes a vowel, is transliterated *t* instead of *h*.

like *hh* which do not change when reversed. 600 of these [ $784 - 28 = 28 \times 27 = 756$  words] are perfect words [i.e., consonants only] with no *Wāw*, *Yā* or *Hamzah* [these are the three basic vowels in Arabic], which come to 300 before reversal [ $(28 - 3)(27 - 3)/2 = 300$ ]. 150 words [of the 756] contain one of these [vowels]: *Yā*, *Wāw* and *Hamzah*, with 75 before reversal [ $25 \times 3$ ]. 6 words [of the 756] contain two [different] vowels [ $3 \times 2$ ], with three before reversal. 3 double-letter words [of the 784] contain the same vowel, 25 [double-letter words], contain identical consonants. You should understand what I just explained to you of the double-letter word counts which the Arabs spoke or rejected.

If you want to count all triplets [three-letter words], multiply the 3 vowels by the 9 double-letter words, they become 27 structures with three letters all of which are vowels [ $3^3 = 27$ ]. Then multiply the 3 vowels by the 150 double-letter words where one letter is a consonant and the other is a vowel; you get 450 forms of three letters, two of which are vowels and one is a consonant. Also multiply the 3 vowels by the 600 double-consonant words; you get 1800 forms of three letters-2 consonants and 1 vowel. Multiply the 25 [consonant] letters by the 600 double-consonant words; they become 15625 three-consonant words [al-Khalīl might have made a computational error here (it should be 15000), but we believe this to be a slip; see below]. This is all that can result of three letters.

If you want to count the quaternary [four-letter] words, use the same method. Multiply the 3 vowels by the 27 all-vowel triplets, then multiply by 450, then by 1800. Then multiply the 25 consonants by the 15000 three-consonant words [al-Khalil is right this time, which proves that the 15625 above was indeed a slip]. These are all the quaternary structures. The same method is used for five-letter words. Six-letter [root] words are not possible except by using superfluous letters [a property of Arabic morphology].

In another part of al-Ain, al-Khalil wrote:

A double letter word has two variations like  $qd$ ,  $dq$ , and  $\dot{s}d$ ,  $d\dot{s}$  [= 2!]. A triple-letter word has 6 variations [= 3!]; the collection of the six variations is called a hexa-word, an example of which is:  $\dot{q}rb$ ,  $dbr$ ,  $br\dot{d}$ ,  $b\dot{d}r$ ,  $r\dot{d}b$ ,  $rb\dot{d}$ . A four-letter word has 24 variations because its letters, which are 4 are multiplied by the three-letter word variations which are 6 and the result is 24 [= 4! = 4 × 3!]. A five-letter word has 120 variations because its letters which are 5, are multiplied by the four-letter word variations which are 24, and the result is 120 [= 5!]....<sup>6</sup>

### ACKNOWLEDGEMENTS

The author spent part of his sabbatical leave as a Visiting Researcher at the Data Transmission Division of the Electrical Engineering Department (Institutonen för Systemteknik) of Linköping University, Linköping SWEDEN, during which time this paper was written. I would like to express my gratitude to Prof. Thomas Ericson for hosting my stay in Sweden, and to Linköping University for using its services and facilities. I am grateful to Dr. Shakir Abdul-Jabbar for the helpful discussions and critical comments on the manuscript. I would also like to acknowledge the critical review and expert editing of Dr. David Kahn, which helped make the paper more focused and, hopefully, easier to read.

### REFERENCES

1. Badeau, J. S. *et al.* 1983. *The Genius of Arab Civilization: Source of Renaissance*. Second Edition. Cambridge: MIT Press.

<sup>6</sup>The exact Arabic text of these two quotations is given by Mrayati *et al* [12]. The translation given here is the attempt of the author [texts between square brackets are added to clarify the original writings]. Apology is extended to the great al-Kahlil for the inevitable weakness and loss of eloquence of this translation in conveying the actual meaning and beauty of the original words.

2. Bosworth, C. E. 1963. The Section on Codes and Their Decipherment in Qalqashandī's *Ṣubḥ al-A'shā*. *Journal of Semitic Studies*. VIII-i (Spring): 17-33.
3. Boyer, Carl B. 1985. *A History of Mathematics*. Princeton: Princeton University Press.
4. Bynum, W. F., E. J. Browne and R. Porter, eds. 1983. *Dictionary of History of Science*. London: Macmillan Press Ltd.
5. ad-Daffa', Ali. A 1978. *Nawabigh Ulama'i al-Arab wal-Muslimeen fi ar-riyadiyyat* (The Genius of Arab and Muslim Mathematicians). in Arabic. New York: John Wiley & Sons.
6. Fauvel, John and Jermy Gray, eds. 1987. *The History of Mathematics: A Reader*. London: Macmillan Press Ltd.
7. Hammer, Joseph von. 1806. *Ancient and Hieroglyphic Characters Explained*. London: W. Bulmer & Co.
8. al-Kadi, Ibrahim A. 1991. Cryptography and Data Security: Cryptographic Properties of Arabic. In Arabic. *Proceedings of the Third Saudi Engineering Conference*. Riyadh, Saudi Arabia: Nov 24-27. Vol. 2: 910-921.
9. Kahn, David. 1967. *The Codebreakers: The Story of Secret Writing*. New York: Macmillan.
10. Khaldūn, Abdur-Raḥmān ibn. 1958. *The Muqaddimah : An Introduction to History*. Trans. Franz Rosenthal. New York: Pantheon Books.
11. Menninger, Karl. 1969. *Number Words and Number Systems: A Cultural History*. Trans. Paul Broneer. Cambridge: MIT Press.
12. Mrayati, Mohammad, Yahya Meer Alam and Hassan al-Tayyan. 1987. *Ilm at-Ta'miyah wa Istikhrāj al-Mu'ammā Ind al-Arab*. (Origins of Arab Cryptography and Cryptanalysis), Volume I : *Analysis and Editing of Three Arabic Manuscripts of Al-Kindī Ibn-Adlān and Ibn ad-Duraihim*. In Arabic. Damascus: The Arab Academy of Damascus.
13. al-Qalqashandī, Aḥmad ibn Ali. 1412. *Ṣubḥ al-A'shāfi Ṣina'at al-Inshā*. (The Light of the Blind in the Profession of Writing). in Arabic. Section 8 of Chapter 2 in Part 4. Reprinted in Cairo (1957): Ministry of Culture and National Guidance.
14. Sarton, George. 1927-1948. *Introduction to the History of Science*. 3 vol. Baltimore: Carnegie Institution.
15. Waerden, B. L. van der. 1985. *A History of Algebra: From al-Khwārizmī to Emmy Noether*. Berlin: Springer-Verlag.
16. Youshekevitch, A. P. 1964. *Geschichte der Mathematik im Mittelalter*. Leipzig, Germany: Teubner.

### BIOGRAPHICAL SKETCH

Ibrahim Al-Kadi was born in Onaizah, SAUDI ARABIA on 19 January 1954. He received the BS from Riyadh University (currently King Saud University), Saudi Arabia in 1978, the MS from the University of Michigan, Ann Arbor MI, in 1980 and the PhD from Stanford University, Stanford CA, in 1984 all in electrical engineering. He worked as a graduate assistant at Riyadh University (1978-79), as a research assistant at the Communication Satellite Planning Center of Stanford University (1981-84), and as an assistant professor at King Saud University (1984-88) where he is now an associate professor. He teaches undergraduate and graduate courses in systems, electromagnetics and communications, and performs research and consultation. He is a senior member of the IEEE. His areas of interest include communication systems, coding and cryptography, spectrum management, radio wave propagation, remote sensing, technology transfer, satellite positioning, and education.

Starting September 1991, he is on a temporary leave from the University working as Vice President for Engineering at the Advanced Electronics Company (AEC).