

# Study of the Distribution and Modification of Writing in Proto-Chinese Language Communities: Correct Identification of Hieroglyphic Signs in the Jiaguwen Inscriptions

Maia Egorova\*, Aleksandr Egorov

Peoples' Friendship University of Russia (RUDN University), RU-117198, Moscow, Russia

**Abstract.** The Chinese language uses the ideographic writing. The outer form of Chinese characters in 3rd century BC and later underwent strong changes, which were caused mainly by change in the technique of writing. In the first century AD the writing style that has survived to the present time. We recall some of main results of the spread and change of linguistic information in certain Chinese community obtained with the help of a dynamic nonlinear model. Undoubtedly, Oracle script writings Jiaguwen (甲骨文) played an important role in this process. We consider it quite legitimate to discuss rather about the written language of the Chinese language, as main link between ancient Chinese, Middle Chinese, and modern Chinese languages Putonghua. In the study of the oldest examples of art, culture and science, for example, such as Jiaguwen, there is a need to study ancient hieroglyphs depicted on the surface of a tortoise shell, copper vessel or stone. New method for correct solving of an inverse problem for pattern recognition is presented. Our method is based on photometry of the studied surface. The obtained data are digitally processed in order to determine the characteristics and parameters characterizing the properties of the investigated surface. The possibilities and limitations of the method, as well as the mathematical model of the groove in the optical range, are given. The main attention is paid to the development of methods for more accurate calculation of the number of solitary grooves and structural elements such as hieroglyphs or ancient Chinese radicals (部首 / bùshǒu). This task is part of the complex inverse problem of pattern recognition, namely: identification of various structural elements, for example, ancient hieroglyphs depicted on the surface of a tortoise shell. To solve this important problem, two main methods were proposed, which are based on the study of both the initial photometric data and the study of their Fourier spectra. The first method allows you to visually identify the location of the characters along the photometric line and determine their approximate number. The second method allows us to get a stable and unambiguous solution to this problem: to determine accurately the number of characters on the photometric line.

## 1 Introduction

The modern Chinese language i.e. Putonghua (Pǔtōnghuà / 普通话) was created artificially in the middle of the 20th century [1-4]. It is based on the vocabulary and grammar of Mandarin (Mandarin Chinese or North Chinese) and the Beijing dialect for pronunciation, i.e., the phonetics and vocabulary of Putonghua is based on the pronunciation standard of the Beijing dialect, which belongs to the northern group of dialects of the Chinese language.

We consider it quite legitimate to discuss rather about the written language of the Chinese language, as the main connecting link between the ancient Chinese, Middle Chinese, and modern Chinese languages [4]. Three periods are usually distinguished in the development of the Chinese language: Ancient Chinese, Middle Chinese and Modern Chinese. In the history of the living spoken language, the ancient Chinese period

ends, most likely, around the sixth century AD, ancient Chinese language from the V-th century BC and to the II-nd century AD can be called classical, the language of earlier monuments is pre-classical, and the language of the third and fourth centuries AD is late ancient Chinese.

The Chinese language uses the ideographic writing; the signs of Chinese writing are called hieroglyphs. In ideographic writing, in contrast to phonetic writing, each sign does not correspond to a sound unit, that is, a sound or syllable, but a significant, a word or morpheme, which are written as a whole, and not divided into their constituent sounds. Different words are spelled differently, even if they sound the same. From this it follows that the same word has been spelled the same way throughout the history of the Chinese language, no matter how its pronunciation changes. Chinese writing does not give us any direct evidence of phonetic changes that have occurred in the language, or information about the phonetic differences between Chinese dialects; the

\* Corresponding author: [MeY1@list.ru](mailto:MeY1@list.ru)

reverse is true, a text written in hieroglyphs can be read aloud in any of the modern Chinese dialects, as well as in Japanese, Korean, Vietnamese and in any language using Chinese writing.

The outer form of Chinese characters in 3rd century BC and later underwent strong changes, which were caused mainly by change in the technique of writing. In the first century AD the writing style that has survived to the present time first appeared (the model or charter letter *kāishū* / 楷书). It differs from previously existing ones in that all the characters in this style are built from a small number of the same basic graphic elements (horizontal line, vertical line, hook, dot, etc.), which completely lose all resemblance to the drawings from which they came.

Oracle bone script writings Jiaguwen (甲骨文, XIV-XI centuries BC) on turtle shells, oracle bones and bronze vessels are among the most ancient Chinese art and culture samples [1-8]. Jiaguwen writings on turtle shells are hieroglyphic inscriptions that record the results of fortune telling or predictions. They are unique in their own way. These objects are often poorly preserved, but are objects of great historical, cultural and scientific value. Basically, their research is devoted to the study of hieroglyphic inscriptions applied to them with the aim of interpreting their content, as well as identifying ancient keys (Chinese radicals) [5-8]. We should note that the structure of the inscriptions practically did not change throughout all periods; they included the date, the name of the fortuneteller, the question, answer and mark of execution; however, the calligraphic style underwent a significant change, from large rough characters of an early period to the smallest ones that are hardly distinguishable by eye. The number of different characters on shells and bones is approximately 5000 characters, of which approximately 1500 are identified with modern characters [8]. In this regard, the most important task arises: to develop a method for recognizing a certain number of important characters depicted on the oracle bones that can be found in the specified list [8].

The present work is devoted to the development of a new method for solving some inverse problem for pattern recognition. The main goal of the work is to demonstrate the possibility of correct identification of groove-type elements that are an integral part of various complex structural elements, according to the obtained digital surface profiles of the sample under study. In particular, in technical applications, these can be various (rectangular, triangular, trapezoidal) solitary grooves, which are the constituent elements of such widely used structures as diffraction gratings (see e.g. [9-11]). In the study of the oldest examples of art, culture and science, for example, such as Jiaguwen, there is a need to study ancient hieroglyphs (keys) depicted on the surface of a tortoise shell, copper vessel or stone. As our study has shown, structural elements such as grooves are part of ancient hieroglyphs, therefore, their more detailed further study, including using rigorous mathematical methods, will provide new results in the study of Jiaguwen (Jiǎgǔwén) [12].

In this article, a new research method is used [12] as a basis, which can be a good addition to the currently used traditional methods (mainly visual) of research [1, 4-8]. This method allows, using digital data of surface photometry, to determine various characteristics and parameters that characterize the properties of the surface under study, including statistical ones [12]. The main advantages of our research method are its contactlessness, informational content, and potentially sufficiently high resolution, for example, by statistical parameters of the surface profile, which allows us to speak about its promising potential in the study, for example, of such ancient examples of art and culture as Jiaguwen.

The results obtained are based on the correct physical and mathematical approach, use rigorous mathematical models and, as a consequence, can be useful when developing other mathematical models, allowing to adequately describe structural elements such as grooves, which are included, for example, in the composition of the most ancient Chinese radicals (部首 / bùshǒu), as well as the hieroglyphs themselves. The developed method can be used in interdisciplinary research, which utilizes methods of mathematical linguistics, lexicostatistics and sociolinguistics, especially in the statistical analysis of hieroglyphic inscriptions.

## 2 Formulation of the problem. Numerical results and their analysis

### 2.1. Dynamic nonlinear model of distribution and changes of linguistic information

In our previous works [1, 4] we used a dynamic model of the spread and change of linguistic information in a certain Chinese community. Let us recall some of the main results we obtained earlier. This dynamic model can be described by a nonlinear equation:

$$I_{m+1} = [a_1 I_m (M - I_m) + a_2 (M - I_m)^2] \lambda, \quad (1)$$

where  $I$  is the value of the analyzed linguistic information,  $m = 12, \dots$  ( $m = 1$  corresponds to the first "measurement", i.e.,  $I_1$  is the initial value, for example, at some initial moment in time  $t_1$ );  $a_1$  is a coefficient that characterizes the distribution of linguistic information in the contacts of "unknowing" with "knowing" this information  $I_n$ ;  $a_2$  is coefficient that characterizes the impact only on "unknowing";  $M$  is the maximum value of linguistic information; and  $\lambda$  is a task parameter (control parameter). The nonlinear equation (1) allows one to explore the process of changing the spreading linguistic information depending on time and other parameters.

After replacing the variables, equation (1) can be written as follows:

$$y = \lambda_1 x(1-x) + \lambda_2 x(1-x)^2, \quad (2)$$

where  $y = I_{m+1}/M$ ,  $x = I_m/M$ ; and  $\lambda_1 = a_1 M \lambda$ ,  $\lambda_2 = a_2 M \lambda$  are new control parameters of the system and  $0 \leq x \leq 1$ .

We use the data obtained in [1, 4] in a numerical study of the distribution and change of linguistic information in the proto-Indo-European language community to analyze the distribution of language information in the proto-Chinese language community, given that there are no riders in it. As a result, in computer modeling, the ratio of the coefficients  $\lambda_1$  and  $\lambda_2$  was selected based on the ratio of the average speeds  $v_1$  and  $v_2$  (in km/h) of different types of pedestrians:  $\lambda_1 : \lambda_2 \propto v_1 : v_2 \approx 3 : 3 \approx 1$ . We consider the language policy of China as another important condition for the task that is aimed at linguistic unification, i.e., the creation of one national language in the future [1, 4].

We have conducted an analysis of the spread of language information in the proto-Chinese linguistic community (see e.g. [1, 4]). The main results of a computer simulation were presented in our papers [1, 4]. A brief analysis is given of a trend that originated in antiquity and developed with varying intensity depending on historical eras, to suppress dialects and reduce oral linguistic diversity to a single unified norm, which resulted in the creation in the middle of the 20th century of the national language of Putonghua (Pǔtōnghuà / 普通话). It is noted that final linguistic unification can be achieved no earlier than the year 3500. An estimate of the number of possible elementary “symbols” (“signs”, “radicals” or “base units”) of a hypothetical proto-Chinese script was obtained of 500–1000 characters for the first time. By basic units we mean the number of primary radicals. To verify our calculations, we examined modern typical SMS messages and made an original comparison of typical hieroglyphs and radicals from these messages with hieroglyphs and radicals on tortoise shells. As a result, matching hieroglyphs and radicals were discovered, which indicates the validity of our approach and the validity of the calculations [4].

We should note that mathematical model of the process of wave propagation and change of linguistics information in this system is described by a system of integral-differential equations and it discussed in sufficient detail in our previous papers (see e.g. [1]), therefore we do not consider it here.

## 2.2. Photometric analysis of the Jiaguwen example

At this stage of the study, there was an urgent need to develop a method for correct identification of groove-type elements that are an integral part of various complex structural surface elements, according to the obtained digital surface profiles of the sample under study. This problem was identified in our previous research when working with a number of ancient samples. It is especially important in light of the fact that the studied samples were the most important element in

the distribution of proto-Chinese script, proto-Chinese and then Ancient Chinese and Middle Chinese languages [1, 4, 12].

Structural elements such as grooves, which are part of the oldest hieroglyphs depicted on the surface of the tortoise shell (XIV-XI centuries BC) are the object under study. The subject of research is the profile of a dielectric surface with hieroglyphs depicted on it. The main task is to determine the number of elementary structural elements and the number of hieroglyphs on the scanning line according to the photometry data of the surface of the sample under study.

In this work, the main attention is paid to the development of methods for a more accurate calculation of the number of grooves and structural elements such as hieroglyphs (ancient keys or bǔshǒu). This task is part of a complex inverse problem for pattern recognition, namely: identification of various structural elements, for example, ancient hieroglyphs applied to the surface of a turtle's shell. The solution of this important problem will make it possible to avoid errors at the initial stage, for example, when approximating (fitting) autocorrelation functions (ACF) are found from the obtained digital profiles of the surface of the sample under study [9-12]. As a result, the correct data will be analyzed at all subsequent stages. The importance and urgency of solving this problem is due to the fact that this problem arises in the study of many structural elements of Jiaguwen. In the case of automation of the process of statistical processing of ancient hieroglyphic inscriptions, the solution of this problem will undoubtedly lead to a significant reduction in the time spent on conducting the entire complex of studies of ancient Jiaguwen samples.

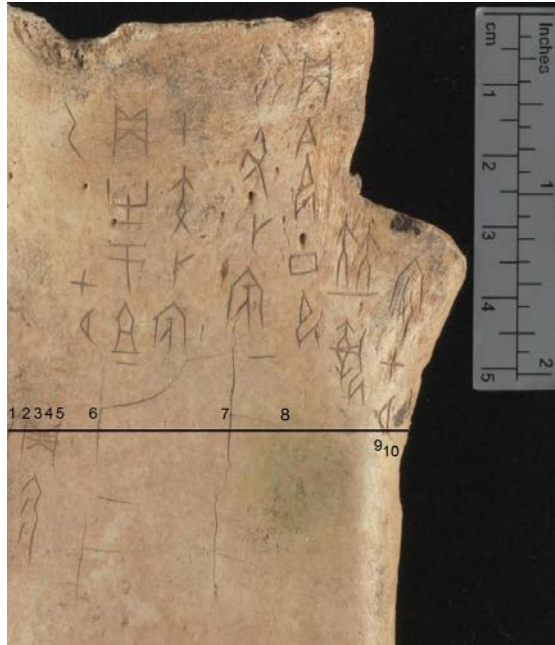
Recall that our method is based on photometry of surface profiles, in which a sampling of brightness levels in a discrete set of points is carried out, followed by conversion into digital form [12, 13]. The obtained data (optical response) are digitally processed in order to determine the characteristics and parameters (including statistical ones) characterizing the properties of the investigated surface [12-15]. As a result, from the digitized distribution of the intensity of the light scattered by the surface, one can find, for example, the form of the approximating ACFs, and also determine the corresponding statistical parameters of the roughness of the profiles of the studied surface: the standard deviation and the correlation radius (see details in [9-13]).

In the present work, using the numerical analysis of the digitized photometric data of the profile of the dielectric surface, we got the results that made it possible to obtain a more accurate value of the number of keys (ancient hieroglyphs, bǔshǒu). In the future, a previously developed statistical research method can also be applied, which makes it possible to find the statistical characteristics and parameters of the profiles of the studied sample. The shell of a turtle (XIV-XI centuries BC), on the surface of which hieroglyphic inscriptions were written, is a typical example of Jiaguwen (see photo in Fig. 1).

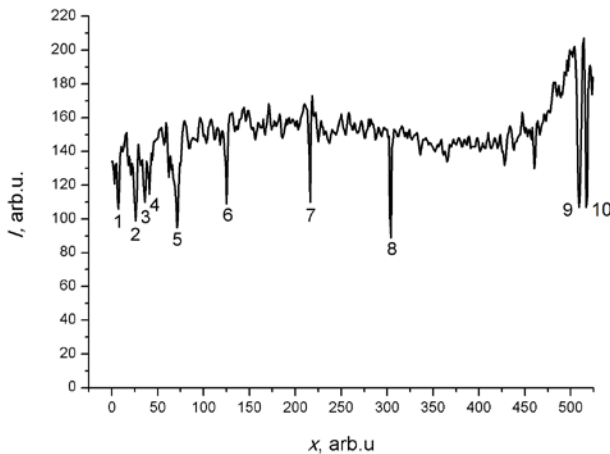
In the photograph given in Fig. 1 a horizontal line (along the axis  $x$ ) along which the surface was

photometric is shown. On the horizontal line, there are 10 groove-type structural elements (numbers from 1 to 10), most of which are undoubtedly part of the hieroglyphs.

Photometric data of the surface profile along the horizontal line shown in Fig. 1 are shown in Fig. 2 and Fig. 3. The signal-to-noise ratio  $S/N$  was greater than 7 in average.



**Fig. 1.** Jiaguwen example (甲骨文 – “writing on turtle shells”): hieroglyphic inscriptions fixing the results of fortune telling or predictions [7].



**Fig. 2.** Photometric data of the surface profile along the horizontal line shown in Fig. 1. 1 arb.u. (along the  $x$  axis)  $\approx$  400  $\mu\text{m}$ .

In Fig. 2 the numbers from “1” to “10” mark the dips in brightness  $I$  corresponding to 10 structural elements of the groove type (see Fig. 1), which are part of the oldest hieroglyphs written in the investigated surface.

The level of the ratio  $S/N$  could change from realization to realization, but on average it was at least 7.

### 2.3. The mathematical model of the groove. Scalar scattering

The mathematical model of the groove in the optical range of wavelengths can be described by the following expression (see e.g. [14]):

$$f(x) = \exp[-i2kh(x)], \quad (3)$$

where  $k = 2\pi/\lambda$  is wave number,  $\lambda$  is wavelength of radiation incident on an object ( $\lambda \approx 0.6 \mu\text{m}$ ),  $h(x) = -h$ , at  $|x| \leq \Delta x/2$  and  $h(x) = 0$  at  $|x| > \Delta x/2$ ,  $\Delta x$  is a groove width. In fact, the quantity  $f(x)$  in (3) is a function of the groove shape or the so-called optical profile.

It is important to note that the relationship in the general case of the nonlocal reflection coefficient  $r$  and the scattering matrix is determined on the basis of the scattering theory. However, under certain lighting conditions and recording the response, it becomes possible to describe the object by the local reflection coefficient. In particular, if the illumination wave is plane and the projection of the modulus of the wave vector on the plane of the object  $q = 0$ , then the reflected wave is determined by one column of the scattering matrix. Then the object can be characterized by the local reflection coefficient  $r$ , which depends only on the object shape function, which is, in fact, the Fourier transform of the scattering amplitude [14]. It is this method of the signal registration that was used in this paper.

When illuminated by a focused wave, an object can also be described by a local reflection coefficient if such a registration scheme is chosen, in which only one (for example, normally reflected with a projection  $q' = 0$ ) plane wave is recorded. Then the object is described by one row of the scattering matrix, and the corresponding local reflection coefficient is determined by a similar dependence (for more details, see, for example, in [14]).

Such a relatively simple mathematical model of the groove can be used in the case of a scalar consideration of the problem of forming an optical response. Indeed, in the analytical and numerical solutions we use only the distribution of the real function  $I(x)$ , obtained from photometric data of the profile of the investigated dielectric surface illuminated by a plane wave.

This mathematical model is also valid from a physical point of view, since structural elements of the groove type were investigated, which were homogeneous along the vertical axis  $y$  and represented an (almost) rectangular groove in cross-section with a width  $\Delta x$  and depth  $h(x)$  much greater  $\lambda$  [14].

If the profile of any groove contains parts that slightly deform a rectangular profile, then parts with dimensions smaller than the wavelength will be smoothed or lost with this approach. Note that with appropriate processing of the response, including in the presence of noise, some of these details can be restored when solving the inverse ill-posed problem. Moreover,

in the absence of noise, the image of a localized object (of finite length) can be reconstructed exactly and with superresolution. A detailed discussion of this problem is beyond the scope of this article (see e.g. [10, 14]).

In the scalar approximation, the following expression can be written, connecting the average value of the height  $h(x)$  of the profile irregularities and the intensity of light scattered by the surface [10, 15]:

$$h(x) \approx \left( \frac{\lambda}{2\pi} \right) B \ln[I_i/I_r(x)], \quad (4)$$

where  $B$  is dimensionless constant that is close to 1 in many practically important cases;  $I_{i,r}$  are light intensity incident on a surface and reflected (scattered) from it. Comparison of expressions (3) and (4) shows that the real part  $f(x) \propto |\ln[I_i/I_r(x)]|$ . In the case of a solitary groove-type object, the profiles  $h(x)$  in expressions (3) and (4) are almost identical. In the case of studying an extended surface profile, where there is a set of similar grooves, expression (4) will describe the profile of some groove averaged for a given section of the profile.

Note that the profile  $h(x)$  can be found more accurately using expression (3) than using expression (4). However, in the first case, to solve the inverse problem of reconstructing the groove profile, it is necessary to use complex methods of recording the real and imaginary parts of the nonlinear response. For example, register the amplitude and phase of a complex response or a correlation signal that requires the development of complex algorithms for its processing (see, for example, [14]). As a rule, this problem arises when solving various problems of super-resolution, i.e. exceeding the Abbe-Rayleigh diffraction limit [9, 10, 14]. In our case, there is no such need, since the width and depth of the investigated grooves are many times greater than the radiation wavelength (see below). And, therefore, it is sufficient to use the scalar approximation and register only the intensity distribution when photometry the profile of the investigated dielectric surface  $I(x)$ . From expression (4) it follows that to determine the surface profile, it is sufficient to know the value  $\lambda$  and the ratio of  $I_i/I_r$ .

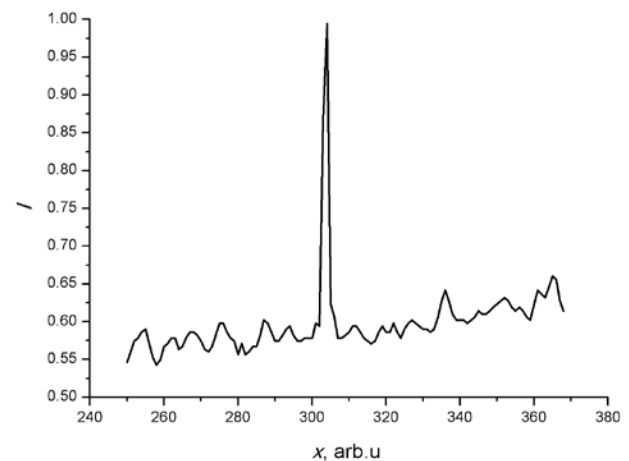
In Fig. 3 shows the inverted normalized brightness distribution  $I$ , shown in Fig. 2, for one structural element of the groove type (marked as “8” in Fig. 2). Using the graph shown in Fig. 4, found the groove width to be reliably determined from the response at 0.5 of the maximum:  $\Delta x \approx 0.75$  mm ( $\gg \lambda$ ). Estimation of its depth  $\Delta h$ : no more than 2–4 mm ( $\gg \lambda$ ).

If the width  $\Delta x$  (and/or depth  $\Delta h$ ) of the groove is comparable to or less than the radiation wavelength  $\lambda$ , then it will be necessary to use additional experimental and numerical possibilities to increase the resolution of the research method described in this work.

As you can see in the photo (see Fig. 1), the width of the vertical line of the cruciform element inside the hieroglyph located above this groove is about 2 times wider:  $\approx 1.5$  mm. Groove “8” was chosen as an example

because it is the least visible visually, especially at the intersection with the horizontal line along which the surface profile was photometric.

Similar data were obtained for grooves “6” and “7”, which are better visually distinguishable. As a result, the data obtained allowed not only to demonstrate the resolution of the method, but also to estimate the approximate width of the cutter used by the scribe - most likely it was in the range from 1 to 1.5 mm. In this case, the depth of the structural elements of the ancient hieroglyphs was probably approximately at first 0.5–2 mm. Note that over time, the size of the structural elements of hieroglyphs could probably increase slightly due to natural reasons.



**Fig. 3.** Normalized inverted profile photometric data of one structural element of the groove type (1 arb.u.  $\approx$  400  $\mu$ m).

Analysis of the numerical data of photometry of the surface profile shown in Fig. 2 and Fig. 3, allowed receiving results which will be useful in the following subsection of work for more exact definition of number of Chinese radicals (部首 / bùshǒu). To solve this problem, we will use a statistical approach, in which the useful signal (data containing information about the grooves) is separated from the total signal, which also contains the noise component. To do this, digital data are pre-processed (smoothed) to reduce the impact of different interference.

### 3 Estimation of the number of ancient Chinese hieroglyphs (radicals / bùshǒu / 部首)

#### 3.1. Estimation of the number of structural elements of hieroglyphic inscriptions such as grooves by the integral method

As an initial step, we define the number of grooves on the scanning line (see Fig. 2). First, indeed, as can be seen from the photometric data in Fig. 2, grooves are the most important structural elements of hieroglyphic inscriptions that record the results of divination or predictions of Jiaguwen (甲骨文). Therefore, the study

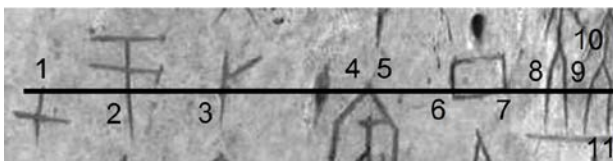
of grooves is an important stage in solving the complex inverse problem of pattern recognition, namely, the correct identification of ancient hieroglyphs applied to the surface of the tortoise shell. Secondly, it will allow us to demonstrate more widely the possibilities and limitations of our research method. Indeed, as the tortoise shell itself [4-7], and selected for analysis the grooves too can be considered as some test samples.

To achieve this goal, when calculating the number of grooves, an integral method is applied, the essence of which is reduced to: 1) finding the area  $S_0$  under the entire smoothed brightness curve (see Fig. 2); 2) finding the area  $S_1$  under the smoothed brightness curve for one structural element of the groove type (see Fig. 3); 3) determining the number of grooves by dividing  $S_0$  by  $S_1$ . After performing the integration procedure, we got:  $S_0 \approx 78690$ ,  $S_1 \approx 8112$ . As a result, we got:  $S_0/S_1 \approx 9.7$ . Thus, it was determined that the number of grooves is 10, with an error 3.1%. Note that to find the corresponding areas, we used numerical integration (in arbitrary units). To increase the accuracy, you can use several of the same type of photometry profiles with subsequent averaging over the sample and additional smoothing. Undoubtedly, this will require additional costs of computing resources and time.

### 3.2. Estimating the number of ancient hieroglyphs using the Fourier transform

In Fig. 4 shows a photograph of a fragment of the sample under study (see the photograph in Fig. 1 in the middle part), which shows a horizontal line (along the axis  $x$ ) along which the surface photometry was made.

On the horizontal line there are 11 structural elements of the groove type (numbers from "1" to "11"), which undoubtedly are part of 7 hieroglyphs.



**Fig. 4.** Photo of the test sample, which shows 11 structural elements of the groove type (numbers from "1" to "11").

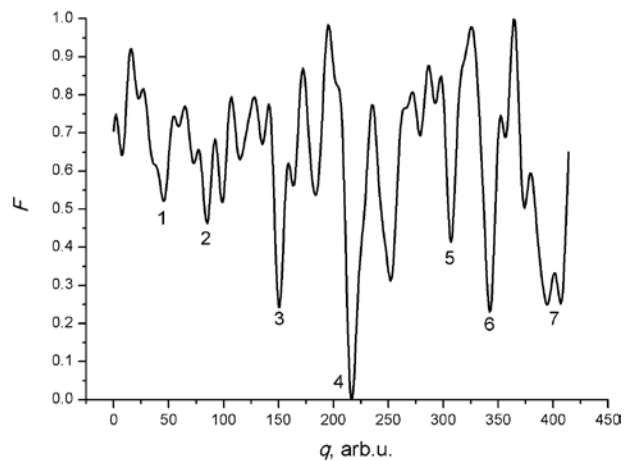
Let us show that to estimate the number of Chinese radicals  $\mathfrak{S}$  (bùshǒu), in contrast to the previously used method [4, 12] and the integral method described above, one can also use the Fourier transform. As a result, it becomes possible to visually identify the location of hieroglyphs along the scanning line and the ability to determine their approximate number. This method can be used in addition to our previously proposed methods for estimating the number of ancient keys in order to more accurately calculate the value  $\mathfrak{S}$ .

Let us apply the numerical Fourier transform to the smoothed data of photometry of the surface profile. Fig. 5 shows one of the obtained normalized Fourier spectra  $F(q)$  of photometric data of the surface profile (see

photo in Fig. 4). The coordinate  $q$  characterizes the position of spectral components of the Fourier spectrum  $F(q)$  in the Fourier plane [10, 11, 14],  $q$  is measured in inverse units of spatial coordinates  $x$  and  $y$ .

A relatively simple method for determining the number of keys  $\mathfrak{S}$  (Chinese radicals) in this case is as follows. The numerically obtained Fourier spectrum  $F(q)$  is normalized, by dividing all values of  $F(q)$  by the main maximum at  $q \approx 364$  arb.u. Then one can find the maximum dips for the function  $F(q)$ . Then, to count the keys in the function  $F(q)$  characteristic dips are found, from which less deep adjacent dips are excluded. In Fig. 5 these dips are in the vicinity of deeper dips, marked with numbers "1-7".

The accuracy of this method can be improved, for example, by setting a certain threshold in terms of the value level of  $F(q)$  (see Fig. 5). In this case, one can more accurately determine the number of hieroglyphic characters (keys). For example, setting the threshold at 0.5, we get exactly 7 characters i.e. Chinese radicals (side lobes of lower intensity are neglected). With this approach, obviously, there is no error in the determination, but the question remains of the exact identification of hieroglyphs location.



**Fig. 5.** Normalized Fourier spectrum  $F(q)$ .  $S/N \geq 10$  (1 arb.u.  $\approx 0.05 \mu\text{m}^{-1}$ ).

It should be noted that the above (integral) method for determining the number of hieroglyphs  $\mathfrak{S}$ , despite certain simplicity is less accurate than the method given in this section due to some understandable subjectivity.

## 4 Discussion

By the beginning of the XXI century in the course of the implementation by the PRC government of the program for the dissemination of the official Chinese language Putonghua throughout China, the Chinese language is gradually being introduced into those areas where it was historically not widespread [1, 2, 16]. Definitely, Oracle script writings Jiaguwen played an important role in this process, especially during this period: V-th century BC -

II-nd century AD. This is why the research of Oracle bone script writings Jiaguwen (甲骨文) is so essential.

It is important to emphasize that when solving the problem of finding the number of hieroglyphs  $\mathfrak{S}$  using the numerical Fourier transform, we have shown the possibility of not only a stable, but also an unambiguous solution to the problem of determining the number of hieroglyphic signs (Chinese radicals / bùshǒu / 部首), as well as their rather precise location.

For a more accurate determination of the number of keys  $\mathfrak{S}$  (ancient hieroglyphs), one can perform an effective numerical smoothing of Fourier spectrum. We should note that the smoothing procedure, on the one hand, suppresses the influence of noise, and on the other hand, it can lead to the loss of useful information.

Method for determining the number of ancient keys (bùshǒu)  $\mathfrak{S}$  and their precise location requires some improvement and enhancement, at least to exclude false triggering of the numerical Fourier transform when processing the original data with noise. For this purpose, it is possible to improve, for example, the method of counting the number of grooves by the integral method. This will at least improve the accuracy of determining the location of the hieroglyphs.

## 5 Conclusion

In this paper we primarily recall some of the main results of the spread and change of linguistic information in Chinese social-linguistic community obtained with the help of a dynamic nonlinear model.

A new method of researching writing on turtle shells (Jiaguwen / Jiǎgǔwén / 甲骨文), which dates back to the XIV-XI centuries BC, has been developed. The method is based on photometry of the investigated surface of the sample and subsequent digital processing in order to determine characteristics and parameters characterizing the properties of the investigated surface.

The possibilities and limitations of the method are described. We demonstrated the possibility of accurate counting of the number of solitary grooves and structural elements such as hieroglyphs (ancient keys / Chinese radicals / bùshǒu). The main advantages of the method are its contactlessness and high resolution in terms of the number of hieroglyphs. The method allows, in principle, to automate the process of statistical processing of ancient texts of the Jiaguwen type.

Therefore, the described method is undoubtedly promising in the study of the oldest examples of art, culture and science, like Jiaguwen, especially in cases where visual identification of hieroglyphic inscriptions is difficult, and the elements of the inscriptions are poorly visible or poorly distinguishable. A certain simplicity and clarity of implementation make it possible to apply this method in interdisciplinary research, which involve specialists from different subject areas of knowledge, for example, humanitarian (cultural studies, linguistics, history, etc.).

Considering that the oldest samples of Jiaguwen under study often have poor preservation, as well as high

historical, cultural and scientific value, our investigate method can be a good and promising addition to the traditional (mainly visual) methods used to research these antique objects.

## Acknowledgements

The publication was supported in part by the “RUDN University Program 5-100”. Authors are grateful to the colleagues for kindly advices, which contributed to the best completion of this paper.

## References

1. M. Egorova, A. Egorov, T. Solovieva, EPJ WoC, **224**, 06009 (2019)
2. *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics* / Eds. Blench R., Sagart L., Sanchez-Mazas A. (Routledge Curzon, London, 2005) ISBN: 978-0415322423
3. L.S. Vasil'ev, Historical questions, **12**, 86 (1974)
4. M.A. Egorova, A.A. Egorov, T.V. Solovieva, Automat. Doc. and Math. Ling., **54**(2), 92 (2020)
5. D.N. Keightley, *Sources of Shang history: the oracle-bone inscriptions of Bronze Age China* (London, Berkeley, 1985) ISBN: 978-0520054554
6. Oracle bone. URL: [https://en.wikipedia.org/wiki/Oracle\\_bone](https://en.wikipedia.org/wiki/Oracle_bone)
7. High-resolution digital images of oracle bones, Cambridge Digital Library. URL: <http://cudl.lib.cam.ac.uk/view/MS-CUL-00001-00155/1>
8. Request for comment on encoding Oracle Bone Script, L2/15-280. Working Group Document, ISO/IEC JTC1/SC2/WG2 and UTC. 2015-10-21. Retrieved 2016-01-23. URL: <https://www.unicode.org/L2/L2015/15280-n4687-oracle-bone.pdf>
9. A. Papoulis, *Systems and transforms with applications in optics* (NY, McGraw-Hill, 1981) ISBN: 0898743583
10. A.A. Egorov, Proc. of the 2nd Sci.-Tech. Conf. “Electronics, micro- and nanoelectronics”, 57 (2000) ISBN: 5-7262-0344-5
11. J.W. Goodman, *Introduction to Fourier optics* (NY, McGraw-Hill, 2005) ISBN: 9780974707723
12. A.A. Egorov, M.A. Egorova, T.V. Demidova, T.G. Orlova, Scientif. and Tech. Inform. Ser. 2, **8**, 24 (2020)
13. W.K. Pratt, *Digital image processing* (NY, Wiley, 2006) ISBN: 9780471767770
14. D.V. Baranov, A.A. Yegorov, E.M. Zolotov, Las. Phys., **13**(11), 1381 (2003)

15. A.S. Toporets, *Optics of the rough surface* (Moscow, Mashinostroenie, 1988)  
ISBN: 5-217-00015-5
16. B. Comrie, S. Matthews, M. Polinsky, *The atlas of languages: The origin and development of languages throughout the world* (NY, Facts on File, 2003).  
ISBN: 978-0816051236