

DIGITAL SCRIPT AND METHOD FOR COMMUNICATION ON MOTHER TONGUE

PATENT BG 63704 B1



Georgi **Mitev**
inventor



Michaela **ŠIMKOVIČ**
inventor



Koycho **Mitev**
leader of inventors



The digital script is applicable in all the world languages and dialects. Thus, the ten numbers can be turned into a universal tool for communication. The unique sequence of operations integrated in the software design to allow real time translation of a finished thought, sentence by sentence, in the respective word order. The operations are common for the phonetic speech and the speech in-writing.

It will allow communication in the mother tongue: voice – to voice; text – to text; voice – to text and vice-versa.

The invention may turn into a tool to communicate in the mother tongue having all the languages and dialects fully equal. And then might come such individuals like Martin Luther King, saying “I have a dream...” all the nations, races and religions to be equal.

The more incredible an idea seems, the more possible it is, the idea to prove a truth to be.

Murphy's Law

“...The father of the first computer John Atanasov was really interested in the most important means of transferring information – the living human language and its written equivalent. The written languages, with regard to the spoken ones, are some kind of a symbolic expression of the latter, which is not synonymously defined and therefore could have a different quality. In connection with his unexpected award with the order of Cyril and Metodiy – Ist class, Jonh Atanasov showed a rare for an American knowledge of the deed of the brothers Cyril and Methodiy. That was due to his great interest in the different kinds of writing. He was often complaining of the high percentage of illiteracy in the USA, which he explained with the imperfection of the written English language, and declared the Cyrillic alphabet much more felicitous.

This motivated his wish to create a new writing, which should be completely phonetic and suitable for people, as well as for machines. This dream of his remained unachieved, although he was persistently striving for it till the end of his life.”

Blagovest Sendov, Academician –

“Centenarian Anniversary of John Atanasov”

<http://www.aba.government.bg/bg/BGpoSveta/Kariera/021203.html>

Aim of the Invention

To propose a new, unknown yet instrument for communication in mother tongue, through computer devices.

Task of the invention

- To compile a digital writing for the phonetic and the written speech of all languages, which people around the world use;
- The digital writing should comprise information for all phonetic and grammatical characteristics of the speech, which would be enough for a precise communication in mother tongue;
- The digital codes of the speech should be connected with a data base of inter-linguistic dictionaries, in order to enable communicating at the same time.
- The digital writing should enable recording on electronic device - electronic chip or processor of PC - of information for the dictionary of many languages with same signs (the digits from 0 to 9), which after their transformation in binary codes should be regenerative;
- The digital writing should enable performing of morphological and syntactical analysis of the sentences of the text to be translated in any language, as well as the reproduction of the translation in sound or text with the grammar and the word-order of the given language;
- The digital writing should enable communication in mother tongue of the type: sound-to-sound (on phone or microphone), text-to-text, sound-to-text and vice versa for all languages and dialects.

This invention is based upon well-known scientific facts.

First scientific fact:

As a result of their historic development almost all nations have created or accepted written signs, with which human speech is written. The different writings are incompatible. They are part of the world cultural heritage and they develop in their own individual way. Today exist about 6800 languages throughout the world, and the prognosis of the scientists is that at the end of this century half of them would disappear. The effect of this process is perhaps comparable with the effect of the disappearance of animal and plant species.

Second scientific fact:

The only common written signs in all languages are the digits (0-9) and the punctuation marks.

Third scientific fact:

The keyboards of the computers in the world are practically incompatible, because the only common keys on them are those for the digits and the punctuation marks.

Forth scientific fact:

In all languages around the world the grammar consists of one and the same components: phonetics, morphology, syntax, lexis, semantics.

Fifth scientific fact:

The mathematics is a universal instrument for analysis of the mechanisms, according to which the nature acts. It could be also an instrument for transferring language information (for example binary codes for recording and transfer of information units, which uses the inventor of the first computer John Atanasov in ABC computer).

Sixth scientific fact:

The component units of any human speech are: sound, phone, allophone, phoneme, morpheme, word, word group, sentence, text. (It should not be mixed up with the term lexical unit.)

Seventh scientific fact:

The organs of speech of Homo sapiens are one and the same for all races and do not depend on the ethnical origin.

Eighth scientific fact:

The number of the phonemes, which Homo sapiens uses in his phonetic speech, is a double figure. There exist technologies, which allow distinguishing among the phonemes, through applying a Differential indication (articulating or acoustic indication of the phonemes) and other means of phonetics.

Ninth scientific fact:

Like the finger prints, the human voice is unique for each person. If there would be compiled a list of limited number of words, which are previously selected in such a way, so that they comprise all possible positions of the phonemes of the speech, then, after a single record of sound from a specific human speech, could be created a technology, which could recognize these phonemes.

PHONETIC DIGITAL CODES

PH code (Phonetic code)

The purpose of this code is to identify the phonemes of a recognized voice with one and the same digital codes, which should be applicable for all languages. Regardless of the signs with which the words in the different writings are written, the sounds of the speech have common characteristics. The structure of the digital codes is, as follows:

- The first two digits show the number of the phonemes.
- The third digit shows length, duration of the sound (0 – normal duration of the sound, 1 – long vowel or geminated consonant).
- The fourth digit shows the presence of the indication “softness” in the consonants (0 – normal sound, 1 – soft consonant), and a stress in the vowels (0 – unstressed vowel, 1 – vowel in a stressed syllable).
- The fifth and the next digits could differentiate, if necessary, others characteristics of the phonemes.

Example:

Vowels (**Cyrillic** & **Latin** alphabet)

A **A** – **0100** **Á** – **0110** (**CZ** – nemáme)

E **E** – **0200** **Ě** – **0211** (**SK** – pekné)

И **I, Y** – **0300** **Í, Ý** – **0310** (**EN** – easy)

У **U** – **0500** **Ú** – **0510**

Ъ - **0600** (**EN** - camera, open)

- **Ä** – **0700** (**SK** – mäso)

- **Λ** "caret" – **0800** (**EN** – puff, blood, but [bʌt],

..... others

Diphthongs

Ей **ei** – **0900** (**EN** – day, come, made, examination)

Ай **Ai** – **1000** (**EN** – time, high, design)

Ой **Oi** – **1100** (**EN** – boy, noise, joy, employ)

- **eu** – **1200** (EN – hypothetic) **1201** (EN – shoulder)
- Ау** **Au** – **1301** (EN – house, town, sound)
- Иъ** - **1400** (EN – frontier, overseer)
- **1401** (EN – here, engineer)
- Еъ** - **1501** (EN – pair, there, chair, wear)
- Уъ** - **1600** (EN – contour, manicure)
- 1601** (EN – sure, tour, moor)
- Уо** **Uo**, **1700** (SK – pôvabný)
- Ю** **Iu** **1800**
- Я** **Ia** **1900**
- Йе** **Ye** **2000** (CZ – medvěd)
- Others till number **2900**.....

Consonants

- Б** **B** **3000**
- В** **V, W** **3100**
- Г** **G** **3200**
- Д** **D** **3300**
- Ж** **J,ž** **3400**
- З** **Z** **3500**
- К** **K** **3600**
- Л** **L** **3700**
- М** **M** **3800**
- Н** **N** **3900**

| | | | |
|-----------|---------------|-------------|-------------------------------|
| П | P | 4000 | |
| Р | R | 4100 | |
| С | S | 4200 | |
| Т | T | 4300 | |
| Ф | F | 4400 | |
| Х | H | 4500 | |
| Ц | C | 4600 | |
| Ч | CH | 4700 | |
| Ш | SH | 4800 | |
| Щ | SHT | 4900 | |
| РЖ | Ř,ř | 5000 | (CZ – například) |
| | H | 5100 | (SK, CZ – hmota) |
| ДЗ | DZ | 5200 | |
| ДЖ | Dj, Dž | 5300 | |
| - | | 5400 | (EN – thank) |
| - | Ó | 5500 | (EN – that) |
| Нь | Ń | 5600 | (EN – sing) |
| | | 5700 | |
| | | 5800 | |
| | | | Others till 9900 |

This system for digital encoding of the sounds of the human speech enables writing with digits of the sound equivalent of each word from a random language. For languages outside the Indo-European group, which possibly have sounds, essentially

10. Interjections

Eleventh scientific fact:

The parts of speech in all languages are variable and invariable (morphological classification).

Code B - Variable parts of speech

- Nouns
- Adjectives
- Numerals
- Pronouns
- Verbs

GRAMMATICAL CATEGORIES AT THE NOUNS

Gender, number, definiteness (with or without article), case, type: with one theme (simple nouns), with two and more themes (complex nouns), countable or uncountable (specific for the English language), others (if there are any).

First number of code B GENDER

Value and meaning of the code

0 – There is no gender

1 – Masculine gender

2 – Feminine gender

3 – Neuter gender

4 – Common gender

5 – Possess two genders

Second number in code B NUMBER

Value and meaning of the code

0 – without number

1 – singular

2 – plural

3 – only in singular

4 – only in plural

Third and fourth numbers of code B Definiteness

Value and meaning of the code

00 – Indefiniteness

11 – Definiteness masculine gender (BG- а, DE - der, EN
- a)

12 – Definiteness masculine gender (BG - ът)

13 - Definiteness masculine gender (BG - я)

14 - Definiteness masculine gender (BG - ят)

.....

21 – Definiteness feminine gender (DE - die, BG - та)

.....

31 – Definiteness neuter gender (BG - то)

.....

41 – Definiteness plural, masculine gender (BG - те)

42 – Definiteness plural, feminine gender (BG - та)

43 – Definiteness plural, neuter gender

.....

50 – Common definiteness (EN the, a, an)

Note: The forms for definiteness, which in some languages stand before the nouns: (EN the, a, an, DE der, die, FR le, la, s, un, une, des, les) and so on, cannot be encoded with the codes for the parts of speech (code A). They are not parts of speech.

Fifth and sixth numbers of code B CASE

Value and meaning of the code

00 – Without case

01 – Nominative case

02 – Accusative case

03 – Dative case

04 – Genitive case

05 – Instrumental case

06 – Prepositional case

07 – Vocative case

08 –

09 –

10 –

.....

21 to 99 Other case forms

Note: Languages, which do not have case forms (Bulgarian, English and so on) use code "00".

Seventh number of code B TYPE WITH ONE THEME (SIMPLE NOUNS)

Value and meaning of the code

0 – without

1 - Proper nouns

2 – Common nouns

3 – Nouns with material meaning

4 – Abstract nouns

5 – Collective nouns

6 – Nouns for actions

7 –

8 –

9 – others

Eighth number of code B TYPE WITH TWO OR MORE THEMES (COMPLEX NOUNS)

Value and meaning of the code

0 – without

1 – Proper nouns

2 – Common nouns

3 – Nouns with material meaning

4 – Abstract nouns

5 – Collective nouns

6 – Nouns for actions

7 –

8 –

9 – Others

**Ninth and tenth numbers of code B
COUNTABLE/UNCOUNTABLE (Characteristic of the
English language)**

Value and meaning of the code

00 – without

01 – Uncountable names of materials (material)

02 – Uncountable abstract nouns (abstract)

10 – Countable individual objects (individual)

11 – Countable, group of similar objects, persons

Example of digital encoding of grammatical categories of
the noun: Code A + Code B:

| | | |
|------------|----------------------|--------------------|
| EN: | Newspaper | 1000000310 |
| | The newspaper | 10050000310 |
| BG: | Инженер | 11100002011 |
| | Инженерът | 1111002011 |

GRAMMATICAL CATEGORIES AT THE ADJECTIVES

Gender, singular or plural, definiteness (with or without
article), case, type: with one theme (simple adjectives), with
two or more themes (compound adjectives), countable or
uncountable (specific for the English language), other (if there
are any).

First number of code B GENDER

Value and meaning of the code

0 – There is no gender

1 – Masculine gender

2 – Feminine gender

3 – Neuter gender

Second number in code B NUMBER

Value and meaning of the code

0 – without

1 – singular

2 – plural

Third and fourth numbers of code B Definiteness

Value and meaning of the code

00 – without

11 – Definiteness masculine gender (BG я)

12 – Definiteness masculine gender (BG ят)

13 – Definiteness masculine gender (BG й)

14 –

.....

21 – Definiteness feminine gender (BG та)

22 –

.....

31 – Definiteness neuter gender (BG то)

32 –

.....

41 – Definiteness plural, masculine gender

42 – Definiteness plural, feminine gender

43 – Definiteness plural, neuter gender

.....

50 – Common definiteness

Fifth and sixth numbers of code B CASE

Value and meaning of the code

00 – Without case

01 – Nominative case

02 – Accusative case

03 – Dative case

04 – Genitive case

05 – Instrumental case

06 – Prepositional case

07 – Vocative case

08 –

09 –

10 –

.....

21 to 99 Other case forms

Note: Languages, which do not have case forms (Bulgarian, English and so on) use code "00".

Seventh number of code B TYPE WITH ONE THEME (SIMPLE ADJECTIVES)

Value and meaning of the code

- 0 – without
- 1 – Qualitative adjectives
- 2 – Relative adjectives
- 3 – Adjectives from participles

Eighth number of code B TYPE WITH TWO OR MORE THEMES (COMPLEX ADJECTIVES)

Value and meaning of the code

- 0 – without
- 1 – Qualitative adjectives
- 2 – Relative adjectives
- 3 – Adjectives from participles

Ninth and tenth numbers of code B COUNTABLE/UNCOUNTABLE (Characteristic of the English language)

Value and meaning of the code

- 00 – without
- 01 – uncountable names of materials (material)
- 02 – uncountable abstract (abstract)
- 10 – countable individual objects (individual)
- 11 – countable, group of similar objects, persons

Eleventh and twelfth number of code B COMPARISON of the adjectives

Value and meaning of the code

- 10 – positive degree

20 – comparative degree

30 – superlative degree

Note: Words, which if separately written could be nouns, as well as adjectives, get digital codes for the two categories parts of speech.

Example of encoding of adjectives **Code A + Code B**:

EN: RESPONSIBLE = 200000100210

GRAMMATICAL CATEGORIES AT THE PRONOUNS

Pronouns nouns, pronouns adjectives, pronouns numerals, pronouns adverbs.

Personal, possessive, reflexive and emphatic, demonstrative, interrogative, relative, indefinite, negative, collective.

First number of code B

Value and meaning of the code

1 - Pronouns nouns

2 - Pronouns adjectives

3 - Pronouns numerals

4 - Pronouns adverbs

Second number of code B Type

Value and meaning of the code

1 – Personal

2 – Possessive

3 – Reflexive and emphatic

4 – Demonstrative

5 – Interrogative

6 – Relative

7 – Indefinite

8 – Negative

9 – Collective

Third number of code B Person

Value and meaning of the code

1 – first person

2 – second person

3 – third person

Forth number of code B NUMBER

Value and meaning of the code

1 – singular

2 – plural

Fifth number of code B GENDER

Value and meaning of the code

0 – without

1 – masculine gender

2 – feminine gender

3 – neuter gender

Sixth and seventh number of code B CASE

Value and meaning of the code

- 00 – without case
- 01 – Nominative case
- 02 – Accusative case
- 03 – Dative case
- 04 –
- 05 –

.....

and so on.

Eighth number of code B FORM

Value and meaning of the code

- 1 – Personal pronouns subjects
- 2 – Personal pronouns direct object
- 3 – Personal pronouns indirect object

Ninth number of code B FORM

Value and meaning of the code

- 0 – full form
- 1 – short form

and so on.

GRAMMATICAL CATEGORIES AT THE NUMERALS

First number TYPE

Value and meaning of the code

- 0 – cardinal
- 1 – ordinal (quantitative adjectives)

Second number - GENDER

Value and meaning of the code

0 – without gender

1 – masculine gender

2 – feminine gender

3 – neuter gender

Third number – NUMBER

Value and meaning of the code

0 – singular

1 – plural

Forth number – USE OF ARTICLE

Value and meaning of the code

0 – without article

1 – with article

and so on.

GRAMMATICAL CATEGORIES AT THE VERBS

Person, number, tense, type, voice, mood, conjugation.

First number PERSON

Value and meaning of the code

1 – first person

2 – second person

3 – third person

Second number NUMBER

Value and meaning of the code

0 – singular

1 – plural

Third number GENDER

Value and meaning of the code

0 – without gender

1 – masculine gender

2 – feminine gender

3 – neuter gender

Forth and fifth number TENSE

(In some languages the number of verbal tenses is a double figure.)

Value and meaning of the code

01 – Present simple tense

02 – Present continuous

10 – Past simple

11 – Past continuous

12 – Present perfect

13 – Past perfect

14 -

15 -

20 – Future simple

21 – Future simple in the past

22 – Future perfect

23 – Future perfect in the past

24 -

Fifth number TYPE

Value and meaning of the code

0 – without

1 – progressive

2 – simple

Sixth number VOICE (active and passive)

Value and meaning of the code

0 – without (impersonal verbs)

1 – active voice

2 – passive voice

3 – participial-passive forms

Seventh number MOOD

Value and meaning of the code

0 – without mood

1 – indicative

2 – paraphrased

3 – imperative

4 – conditional

Eighth number Conjugation

Value and meaning of the code

0 – without conjugation

1 – “e” conjugation

2 – “i” conjugation

3 – “a/ya” conjugation

The infinite verbal forms – participles, verbal adverbs and verbal nouns are encoded as verbs, as well as respectively adjective, adverb and noun.

Invariable parts of speech:

ADVERBS

First number ORIGIN

Value and meaning of the code

0 – Non-derivative

1 – derivative from names

2 – derivatives from verbs

3 – derivatives from adjectives

4 – from case forms

5 – from numerals

6 – pronominal adverbs

Second number TYPE OF THE PRONOMINAL ADVERBS

Value and meaning of the code TYPE

0 – demonstrative

1 – interrogative

2 – relative

3 – indefinite

4 – negative

5 – summary

Third number PURPOSE

Value and meaning of the code

0 – of place

1 – of time

2 – of manner

Forth number FORM

Value and meaning of the code

0 – full form

1 – short form

GRAMMATICAL CATEGORIES AT THE PREPOSITIONS

First number ORIGIN

Value and meaning of the code

0 – primary (simple)

1 – derivative (complex)

Second number DOUBLE AND GEMINATED

Value and meaning of the code

0 – single

1 – double

2 – geminated

Third and forth number MEANING

Value and meaning of the code

01 –

14 – surpassing, exceeding a limit

15 – place

.....

and so on.

Fifth number CONNECTION

Value and meaning of the code

0 – verb with a noun

1 – name with a noun

2 – noun with an adverb

3 – adjective with a noun

Categories at the CONJUNCTIONS

First number TYPE (according to their morphological construction)

Value and meaning of the code

0 – simple

1 – complex

2 – correlative

Second number FUNCTION

Value and meaning of the code

0 – definite

1 – objective

2 – explanatory and adverbial

Third number

Value and meaning of the code

0 – coordinating

1 – subordinating

Forth number TYPE OF THE COORDINATING CONJUNCTIONS

Value and meaning of the code

0 – without meaning

1 – copulative

2 – correlative (disjunctive)

3 – opposite

and so on.

Categories at the particles

Besides the categories of participles here will be encoded also the letters (symbols) of the relevant alphabet.

First number FUNCTION

Value and meaning of the code

0 – without meaning

1 – with a linking role in the sentences

2 – word-forming particles

3 – form-building particles

CATEGORIES AT THE INTERJECTIONS, PREFIXES AND SUFFIXES

First number NUMBERING ACCORDING TO THE TYPE

Value and meaning of the code

001-199: interjections

200 – 299: prefixes

300 – 399: suffixes

In addition: There exist languages, in which individually written words could be defined as different parts of speech (nouns, adjectives, verbs and so on) at one and the same time. These words receive digital codes for each of the parts of speech.

SYNTAX

Code C

It defines the role of the words in a given sentence (after a syntactical analysis).

We can use a code with four or more digits.

Structure of Code C

First three numbers

Value and meaning of the code

100 – Subject

200 – predicate

210 – simple predicate

220 – compound predicate

221 – compound verbal predicate

222 – compound nominal predicate

- 230 – complex compound predicate
- 231 – complex compound nominal predicate
- 232 – complex compound verbal predicate
- 300 – object
- 310 – direct object
-
- 400 – adverbial modifier
- 410 – adverbial modifier of time
-
- 500 – attribute
-

and so on until finishing off with all parts of speech.

Forth and fifth numbers

Value and meaning of the code

- 01 – first in order subject, predicate, attribute and so on.
- 02 – second in order
- 03 – third in order

and so on until 99.

Code D - Type of sentence

Value and meaning of the code

- 1 – declarative sentence (“.”)
- 2 – question (“?”)
- 3 – imperative sentence (“!”)

Code E - Type of sentence according to the number of predicative centers.

Value and meaning of the code

1000 – simple sentence

2000 – complex sentence

2100 – compound sentence

2110 – compound connective sentence

2120 – compound contrary sentence

.....

2200 – complex sentence

2210 – complex with subject subordinate clause

2220 – complex with predicate subordinate clause

.....

and so on until all types are mentioned.

Code AC – Area of Cognition

With this code we separate the cognition for the world into areas, in which the idioms have different meaning. Each word, which is a bearer of the quality "idiom", would receive an individual code for the different areas of cognition.

The digital code C is temporary and considers a given sentence from the text to be translated.

Value and meaning of the code

01 – Astronomy

02 – Biology and Medicine

.....

14 – general knowledge

.....

38 – sport

and so on.

KL – Kind of language

With this code we write down the different languages and the significant dialects in them.

Value and meaning of the code

00100 – Albanian language

00101 – First significant dialect in Albanian

00102 – Second dialect

00103 – Third dialect.....

.....

05000 – English language

05001 – American English

05002 – Texas dialect

.....

10000 – Bulgarian language

10001 – Rhodope dialect

10002 – Dialect of the Sofia district

.....

60000 – Slovakian language

60001 – Zahoratski dialect

60002 – Sharishski dialect

60003 – Spishski dialect

..... and so on.

Twelfth scientific fact:

Regardless of what the language origin is, the sentence is a basic grammatically formed unit of the coherent speech, which is distinguished for its unity in sense and intonation, with communicative significance.

The invention offers a method for translating complete sentences from a text.

Before beginning a translation, the sentences from the text are put under a syntactical analysis. Thus are defined the codes C, D, E.

Thirteenth scientific fact:

When doing a syntactical analysis in all languages are valid rules for GOVERNMENT, CONCORD, APPLYING, PREPOSITIONAL CONNECTION and so on, in accordance with the grammar of the relevant language.

By the syntactical analysis of the sentence the information from the codes A and B is necessary and enough. This information is enough also for defining the word order of the text to be translated.

The syntactical analysis begins with the so called **“Predicative center”** – the connection **subject – predicate**. For the purpose of the syntactical analysis the program would first search which word is the verb in the sentence (first number 5). After that - which noun (first number 1) or personal pronoun and so on is agreed in gender and number with the verb, and is a subject in this sentence. This information is hidden in the codes “B”. According to this logic is made a full syntactical analysis and each word from the given sentence receives its temporary code.

The digital codes from one language are transformed in digital codes of the second language. Again a syntactical analysis is made on the basis of the temporary codes "C", the rules of the syntax of the second language are applied (government, concord and so on) and the sentence is reproduced with the word order of the second language in sound or written text.

PHRASEOLOGY

Encoding of semantically synonymous phrases of two or more words (phraseological expressions)

In all languages the sentences consist of semantically synonymous separate words and/or semantically synonymous phraseological expressions (a semantic set of two or more words), phraseological coalescence, phraseological unities, phraseological word groups. The semantic meaning of these elements of the phraseology is encoded as a set of the separate words with a connecting symbol, for example: "_".

Phraseological coalescence: A word group, the common meaning of which does not have any relation with the meaning of the composing words.

Phraseological unities: A word group, which is used, as the idioms, in specific meaning, but the relation between the common meaning and the meaning of the lexical components is not lost.

Phraseological word groups: Word groups, which do not have a different meaning from that of the lexical components, but one of the components could be replaced with other, synonymous, without changing the common meaning.

Intermediate categories: Phraseological word groups at expressions, which have already separated from the free word groups, but on the other side are not yet completely stable expressions.

Verbal tenses with the participation of auxiliary verbs.

The digital equivalent of the last two examples will correspond with the digital codes of the verb "work", which define respectively first person, singular, present progressive tense and first person, singular, present progressive tense with negation.

All these phraseological expressions could be remembered as semantically synonymous structures, which would be translated with one or more words or relevant phraseological combinations in the other language. For the purposes of the syntactical analysis though, the software would treat the separate words as independent units in the sentence structure.

(It would ignore the sign "_").

EN: It _ is _ raining _ cat _ and _ dogs.

BG: Вали _ като _ из _ ведро.

PD Code - Place in dictionary code

Digital encoding of the place of the words in the language

The most significant achievements of the human race for its millenary existence are the languages of the individual nations and their amazing function to serve for communication. Each known, newly-coined or accepted by another language word is a treasure, which should be placed in the safe of the language. The great Russian scientist Mendeleev ordered the chemical elements, which compose the nature, in a special periodical table, giving them a number and leaving space for those not yet discovered.

Since, for the purposes of technology, we pursue the possibility the whole human speech to be presented only with digits, we could propose a preliminary system in which each word would receive its own and single number in the dictionary of the language. This could be also a subject of a separate

project. The digits of Mathematics give us unlimited quantitative possibilities; therefore we accept two basic categories of words:

- known words;
- words, which could appear later on.

Besides, it should be taken into consideration the possibility by personal names, geographical designations and so on, words which are not from the given language to be identified exactly by the translation from one language to another.

It is also important to keep in mind that, the choice of a system for ordering the words of the dictionary of each language, could give us the possibility to group the parts of speech according to other indications as well, for example those connected with the participation of the prepositions – duration, belonging, distribution, purpose and so on. Or, for example indications as: quality, movement, mechanism, means, plan, interrogative word, request, invitation and so on.

Such a system would enable easier finding of corresponding semantically-synonymous words and phraseological word-groups from the dictionary of the language, in which is translated.

Having in mind these and other considerations, for the needs of technology, preliminary and conditionally we accept following system for numbering the words in basic form – first person, singular, present tense for the verbs (we choose a system with six-digit numbering, but it is not a problem, if it should have more digits – seven, eight and so on.):

Common noun 000001-099999

Proper noun 100000-399999

Geographical designations 400000-499999

Adjectives 500000-599999

| | |
|-----------------------------------|---------------|
| Verbs | 600000-699999 |
| Numerals | 700000-799999 |
| Adverbs | 800000-899999 |
| Pronouns | 900000-924999 |
| Prepositions | 925000-925999 |
| Conjunctions | 926000-926999 |
| Particles (including the letters) | 927000-989999 |
| Interjections | 990000-999999 |

In this system for numbering the words we assume that there is a space for words which would appear later on in the language. In necessary, the system could be enlarged with more digits (instead of the maximum of 99999 adjectives which are foreseen in one class, it could be 999999, 9999999 or more).

The digital codes PD are the connection of the digital writing with a data base. It is known that most of the dictionaries show words in their basic mode (first person, singular, present tense). The single writing of the digital codes PD in the system's memory would enable by a translation of a text after each writing of the words to identify their PD code, and then, after a morphological and syntactical analysis, the other codes as well.

All spelling rules from the grammar of the given language could be entered as an algorithm by means of the digital writing.

Example:

The English verb WORK could receive PD code for a basic form (first person, singular, present tense): WORK=623125.

But as a variable part of speech, according to grammar, this verb can have forms for person, number, tense and so on.

And if, according to English grammar, The Present Participle is formed by adding “-ing” to the infinitive, then this grammatical rule can be written and the system will recognize, that the word WORKING is a participle, which is a component of a Progressive tense. The algorithm will consist of the digital code of the infinitive of the verb (Code A, with first number 5), plus the digital code of the suffix (Code A with first number 0 and the number of the suffix).

Digital form of a random sentence

Code for type of the language → Code for area of cognition → Σ || Code for the place of the words in the language → Code for the relevant part of speech → Code for grammatical categories → Phonetic code → Code for the role of the word in the sentence || → Code for the type of the sentence > Code for the structure of the sentence.

KL → AC → Σ || PD, A, B, PH, C || → D → E

Attention: The phonetic codes PH are presented at the digital-encoding system with possible variants from the conversational speech.

Example: BG - Literary speech - **отива на работа** → **(goes to work)**

Dialectical form → **Утива нь рабутъ**

The changed phonemes or whole words are written in advance together with the correct digital codes A and B and the program for phonetic translation should perceive and reproduce them in the translation as literary pronounced. This would enable a sound communication of the type: dialect-to-literary speech, dialect-to-dialect, literary speech-to-dialect and literary speech-to-literary speech.

Preparations for creating software for communication in mother tongue

- Defining the structure and the content of the codes PD (place in dictionary). Encoding with

the PD codes of the words in basic form. The software should enable automatic identification of the PD codes when writing the words.

- Creating auxiliary software products for defining the codes A, B, C, D, E. All spelling rules are written as algorithms, presented by the digital codes.
- Creating auxiliary software products for distinguishing the phonemes from a preliminarily recorded voice. A limited number of words is chosen, the phonemes of which are preliminarily encoded with the phonetic codes PH. The range of words comprises all possible positions of the phonemes. The software should enable a future random combination of the recorded phonemes, according to the phonetic images of the words from other languages. The missing phonemes from a language are supplemented from another voice or synthesizer.
- Creating or presenting of the parsers for syntactical analysis by means of the digital codes – finding of the so called “predicative center”, finding of the verb-predicate through code A = 5, finding of the subject, which is agreed by gender, number, word-order and so on.

Order of operations for translation text-to-text

(The program can be activated after the marks full-stop, question mark, exclamation mark and so on.)

Writing the sentence in the PC memory



Connection with the data base



Defining the PD codes



Syntactical analysis



Defining the codes A, B and the temporary codes C



Defining the codes D, E



Translation of the semantically synonymous units



Morphological and syntactical analysis



Defining the exact codes A, B and the temporary codes C



Defining the codes D, E



Written reproduction of the sentence in the other language

Order of operation for translation sound-to-sound

Recording of the sentence in the PC memory



Defining the combinations of component phonemes for the words



Defining the digital codes PD



Syntactical analysis



Defining the codes A, B and the temporary codes C



Defining the codes D, E



Translation of the semantically synonymous units



Morphological and syntactical analysis



Defining the exact codes A, B and the temporary codes C



Defining the codes D, E



Sound reproduction of the sentence in the other language

An exemplary digital image of a sentence with a translation in several languages:

A note for explanation:

- The digits are chosen at random and show only a general type of the digital codes;
- Each word is written also with the relevant phonetic codes, which enables a communication text-to-text, voice-to-voice, voice-to-text and vice versa;
- The word-order in each language is chosen in accordance with the presence of the temporary codes C.

Example:

EN: This invention will change our ideas of communication!

KL code = 05000; AC code = 14

KL→AC→| | PD, A, B, PH, C | | D→E

This (pronoun)

Σ938244→3→1431300→540003003500→500

Invention (noun)

**038339→1→1100000010→030039003100020139004800
06003900→100**

will _ change (verb)

**638256→5→313200100→050003003700→470009003900
5300→200**

Our (pronoun)

904321→3→223220020→13010600→310

Ideas (noun)

087666→1→2200003410→10003300030001004200→300

Of (preposition)

925110→7→0001→06003100→0000

Communication (noon)

055543→1→2100002011→36000100380038000500390
0030036000900480006003900→500]

BG translation:

Това изобретение ще промени нашите представи за комуникация!

KL code = 10000 AC code = 14

KL→AC→| | PD, A, B, PH, C | | D→E

Това (местоимение)

Σ014238→3→1431300→4300040031000101→500

Изобретение (съществително име)

023456→1→1100000010→030035000400000410002004
3000201390003000200→100

Ще_промени (глагол)

618763→5→31200100→49000200_40004100040038000
20039000301→200

Нашите (местоимение)

904321→3→223220020→390001014800030043000200→
319

Представи (съществително име)

087645→1→2200003410→400041000200330042004300
010131000300→300

За (предлог)

925114→7→0001→35000101→0000

Комуникация (съществително име)

**054593→1→2100002011→360004003800050039000300
36000101460003001900→500| |→2→100**

The translation on other languages is similar:

FR: Cette invention changera nos idées de communication!

DE: Diese Erfindung wird unsere Ideen von der Kommunikation ändern!

E: Esta invención cambiará nuestras ideas de la comunicación!

RU: Это изобретение изменит наши идеи относительно коммуникации!

PT: Esta invenção modificará as nossas idéias da comunicação!

Advantages

- The technology enables, after specifying the quantity of grammatical categories for the parts of speech in the languages around the world, to be created a common matrix for the codes A and B. In this common matrix each cell could have data for the relevant grammatical category of the given language.
- The phonetic codes could be finally specified from specialists-phoneticians, so that for all languages the phonemes should have a common digital record phonetic code PH.
- All grammatical rules could be entered by means of the digital writing. For example the rules for "reduction" of the sounds could be entered in the program so that the phonetic image of the speech to be perceived synonymously.
- The use of a common for all languages digital writing for the phonetic and written speech,

which is connected with a data base (inter-lingual dictionaries) would enable on an electronic bearer (chip, PC processor and so on) to be written the information for the lexical content of plenty of languages.

- By absence of sufficiently good inter-lingual dictionaries (for example if there is not yet a Bulgarian-Maltese dictionary in an electronic variant), the translation in sound and text could be accomplished also through a third language, for example the English one.
- This invention permit to arrange an international digital script for voice and written speech of all human languages – international standard for communication on mother tongue. It is like international standard for meaning of the chemical elements – H, Ca, Li, Na, Fe, Mg, Al....

-End-

THE OTHERS FOR US

Experts from Massachusetts, **USA**:

Bulgarian Invents Unique Translation Method

http://www.commonseadvisory.com/news/global_watchtower_one.php?wat_id=164

IZVESTIA, **Russia**

<http://www.izvestia.com/tech/article39443>

<http://kuda.ru/news.asp?mode=4&ID=755>

India (page 24 from the text, or page 19 from 132 of PDF file):

[http://www.saneinetwork.net/pdf/SANEI_VI/SANEI-VI-\(EcommerceandEconomicDevelopment_FPEPR\).pdf](http://www.saneinetwork.net/pdf/SANEI_VI/SANEI-VI-(EcommerceandEconomicDevelopment_FPEPR).pdf)

Bulgarian radio:

http://www.bnr.bg/radiobulgaria/emission_english/theme_science_and_nature/material/inventions.htm

Agency France Press

<http://www.bulgaria-france.net/kmitev.html>

Brazil

http://www2.correioweb.com.br/cw/EDICAO_20021016/vid_mat_161002_125.htm

One invention against terrorism:

http://www.democrit.com/category.php?n=330&cat=27&br=12&wh_n=news17

Presentation of invention "Digital script and method for communication on mother tongue" (BG text):

<http://www.democrit.com/Presentation-patent.pdf>

China radio international

<http://bg.chinabroadcast.cn/64/2005/09/29/1@28758.htm>

BablePort

<http://www.babelport.com/news/1106>

Велико българско изобретение чака своя час

<http://www.novavizia.com/?81&03043&97790>

.....
For more information - <http://www.google.com> – "koycho mitev" or "койчо МИТЕВ"

**This invention will transform our ideas for communication.
Invest in this project!**

Contacts:

Eng. Koycho MITEV

4, Madarski konnik Street, Vhod "V"

7020 Ruse, Bulgaria

Phone: +359 899103367

E-mail: info@patentservice.info