

Statistics and a mathematical model of the phenomenon of oddness in palindromes of European and oriental languages

Dr. Yasir Z. Abo Rass
E- mail: abo.rass@gmail.com
Sakhnin Academic College
Dr .Saleh Y. Abo Romi
E- mail: salh.aboromi@gmail.com
Elzahraa Elementary School - Tamra

In 2004 - in "Science and Life"- they spoke about a strange linguistic phenomenon: 95% of the palindromic words in the Russian language have an odd number of letters. The following investigation showed a similar phenomenon also in other languages which use letters of the alphabet. In English, French, German, Spanish and Polish, the percentage of odd palindromic words is between 73% and 98%. The phenomenon of oddness is also characteristic of palindromic expressions in these languages. 70%-90% of the palindromic expressions have a central letter by contrast to 10%-30% with a central space.

They succeeded in giving a mathematical explanation to this phenomenon with the help of the theory of discrete groups and the "Antia" function.

As a basis to the data, a list of palindromic words was taken from Russian which includes 196 words and derived words. Derived words included inflections of nouns, adjectives, verbs, participles, gerunds in various times. First names were included in the list (Anna, Ella) and also abbreviations like M"M (deputy).

A list which is almost full of these words is found in "Science and Life" no. 11.

Below are the most common words from the list:

- 1.And-, for-, in-, from-,...
- 2.M"M, ...
- 3.Alla, Abba (father), Ima (mother), (Aya) Was, David, (Ana) Anna, Bob,...
- 4.Madam, (Shabash) disrupted,....

The average number of letters in the words is 4.1 with a standard deviation of 1.6.

An imbalance stands out in this list between words of an even length and words of an odd length.

Two research problems were formulated due to this:

1.To try to construct a mathematical model for palindromic words; to find an analytical function that describes them.

2.To test what happens in other languages using an alphabet of letters, chiefly in European languages.

Mathematical model

Distribution in this problem is not similar to any other distribution (Normal, Poisson, exponential,...).

We started in the simplest situation- words composed of one letter and afterwards gradually we added symmetric forms from both sides.

We will formulate the problem.

Combination with n letters ($n > 0$) is a chain of n "seating places" for letters comprising a possible word. Each letter can "sit" in any place (of course in language this is not so, but here only an estimate)

There are 33^n combinations (there are 33 letters in Russian).

A form with n letters is a palindromic combination with n letters. We will pay attention, that meanwhile a logical word is not concerned.

A form with one place or two can seat only one letter.

A form with 3 or 4 places can seat a maximum of two letters.

Etc.

Therefore, only $33^{\left\lceil \frac{n+1}{2} \right\rceil}$ forms are possible. A ratio of forms with n letters to combinations with n letters is $33^{\left\lceil \frac{n+1}{2} \right\rceil} = 33^n : 33^{\left\lceil \frac{n+1}{2} \right\rceil - n}$

Hereafter are the initial values of the function $y = \left\lceil \frac{n+1}{2} \right\rceil - n$

In a mathematical result as it is customary it is important to emphasize the existence of "horizontal steps". They delay reduction of ratio of forms with n letters when this ration decreases inasmuch as the number of letters in the form increases.

These steps are a factor for the increase of the number of odd palindromes with 2n-1 letters in relation to even palindromes with 2n letters.

In continuation, we will adopt the following conjecture:

The ratio of palindromes to normal words is approximately like the ratio of forms with n letters to combinations with n letters. Meanwhile we can test this conjecture only for Russian.

Table 1: Distribution of palindromes in Russian according to number of letters: True- row 3 and hypothetical- row 4.

Number of letters, n	1	2	3	4	5	6	7	8
Number of words in dictionary (Russian)	11	159	826	2144	4465	7077	10778	14645
Number of palindromic words in dictionary	11	3...1	33	1	23	1	1	0
Number of palindromic words in dictionary (conjecture)*	11	4,8	25	2	4	(0,2)	(0,3)	(0,01)

Dictionary with 130 thousand letters does not contain all the derivatives about which we spoke at the start.

As we see from Table 1, a mathematical model gives local extremums when n is odd.

And this is a main result. This result deals with palindromes in every language serving in a letter alphabet.

It is possible to see convergence in initial values n=1,2,3,4.

When n=1, there is overlap (according to definition).

When n=2, only one word appears in the dictionary.

A single palindromic word exists with 4 letters in the dictionary- ummu.

The dispersal (5 times more) when n=5 depends on the fact that there are many palindromic words in the language with 5 letters (absolute maximum) while the ratio of normal words with 5 letters is smaller.

When n=5 modal points tend to 0 like the number of logical palindromic words.

The law is kept: there are more odds than even.

What happens in other languages- Meet, see, esteem!

For a comparison, palindromic words were taken from the internet from English (75) and French (81). Also palindromic expressions were selected: 500 from English, 19 from French, 200 from German and 134 from Spanish.

Apart from this we alone (from memory and with the help of normal dictionaries, as we did not find palindromic words on the internet) we recorded palindromic words from German (40), Spanish (43) and Polish (58) which is approximately half of their quantity in these languages.

English:

Two lists of palindromic words appear on the internet (they appear in Table 2) as (a) and (b).

A first list contains 75 useful words including abbreviations:

a, aha, BB, Bab, bib, Bob, boob, bub, civic, dad, deed, deified, denned, dewed, did, dud, DVD, eke, ere, eve, ewe, eye, gag, gig, hah, huh, huh-uh, i, kayak, kook, level, ma'am, madam, mom, mum, noon, nun, o, oho, pap, PCP, peep, pep, Pip, poop, pop, pull-up, pup, put-up, racecar, radar, redder, refer, repaper, reviver, rotator, rotor, sagas, sees, sexes, shahs, sis, solos, SOS, stats, tat, tenet, tit, TNT, tot, toot, tut, tut-tut, wow,ZZZ.

A second list contains 207 words. We will present several examples. A lost contains old, dialectic words: desesed= disseized, devoved=deuoued=devoted, evitative, gnipping, goog (Australian)=egg, semitimes= half times; words which were taken from other languages like Geographic names:

Anahanahana, Kanakanak, Wassamassaw, Glenelg (Шотландия), Akasaka, Kрк, etc.;

There are innovative words and "words at the appropriate moment" (nonce words):

beeb (BBC), deleveled, finnif (5\$), redivider, tattarratta;

Scientific names of animals:

aidemedia, allenella, arara, biib, sunnus, ugotu, wow-wow, peeweep, tottot, etc.;

Chemical terms: xitix, detartrated, detannated.

At the end we will stress the word **aibohphobia**= fear from palindromes (this word appears also in Spanish). Apparently the word is created thus: they took the word **phobia** which means fear and attached the opposite to it from it.

According to a main list (a) the ratio of words with an odd number of letters equals **81%** = 0.1= 60/74

And according to a secondary list (b) 64% = 0/4= 132/207.

Equal average values (a) 3.9 and (b) 6.2. The standard deviation: (a) 1.4, (b) 1.8.

Table 2: distribution of words- palindromes according to the number of letters (n) in Russian, English, French and Polish.

Language	n=1	2	3	4	5	6	7	8	9	10	11	12
Russian	11	3	57	5	67	2	5	0	5	0	0	0
English	3 (8 0 (2	1 0	37 1	9 25	16 67	4 36	5 35	0 10	0 24	0 3	0 5	0 1
French	4	0	24	5	32	8	6	0	3			
German	5	0	15	5	7	5	2	1	1			
Spanish	5	0	20	2	12	2			1		1	
Polish	7	0	28	1	20	0	2					
Arabic	14	1	115	20	196	1	10	0	0	0	0	0
Hebrew	3	37	116	8	20	1	0	0	0	0	0	0

(Inflections of nouns, verbs, verbs at different times are included).
 We will bring a short list of palindromic expressions from English:
 A dog! A panic in a pagoda!
 A man, a plan, a canal: Panama!
 Beware era web!
 Bosnia ... pain ... sob.
 Bosnia gasps again--sob!
 Bombard a drab mob.
 Borrow or rob?
 Cain, am I maniac?
 Debate with girl last; if it's all right, I wet a bed.
 Devil never even lived.
 Diaper sin is repaid.
 Do geese see god?
 Dogma: I am God.
 Don't nod.
 Eros saw I was sore.
 God damn! Mad dog.
 God lived as a devil dog.
 No Garden, One Dragon.
 Now, Sir, a war is won!
 On a clover, if alive, erupts a vast, pure evil; a fire volcano
 Oozy rat in a sanitary zoo
 Rise to vote, sir.
 Sex at noon taxes.
 Star sees rats.
 Too hot to hoot.
 Was it a bar or a bat I saw?
 Was it a car or a cat I saw?
 We panic in a pew.
 Won't lovers revolt now?
 Zeus was deified, saw Suez.

At the end of the list we will bring a magnificent palindrome Never odd or even! As if especially for this article.

French:

(1) a, à, ô, y; (3) Ada, aga, ana, ara, asa, axa, Bob, ère, été (היה), été (קיי), eue, gag, ici, non, oxo, pep, pop, sas (מסנן), sas, ses, sis, sus, tôt, tût; (4) elle, erre, erré, esse, réer; (5) anona, aviva, étête, kanak, kayak, lebel, nagan, nanan, Laval, radar, rotor, rêver, sagas, salas, sanas, sapes, semés, senés, seres, sexes, seves, shahs, sidis, solos, sonos, stats, stots, talat, tapat, tarat, tatat, taxat, (6) sassas, selles, sennes, serrés, taggat, tallat, tannat, tassat, (7) Nauruan, réifier, retâter, sémames, sememes, snobons, (9) essayasse, malayalam, ressasser.

The list contains 82 words (not including abbreviations which are very widespread in this language). By contrast to Russian in French the number of words with double consonants in the center is large. The ratio of words with an odd number of letters is equal to

$84\% = 0.84 = 69/82$. The average number of letters in words is 4.5, the standard deviation is 1.7. The mark of oddness in French: $70\% = 0.07 = 13/19$.

Examples of palindromes in French:

Esopè rèste ici et se repose

Un roc cornu.

L'ami naturel? – Le rut animal.

Ce satrape repart à sec.

Eluparcettecrapule.

German:

(1) a, o, u, v, s. (аббрев.см.); (3) aha, Bub, Ede, Ehe, nun, oho, Ovo, Pop, rar, Tat, tat, tät, tot, tut, Uhu; (4) Anna, Esse, esse, Otto; (5) Kajak, Nagan, neben, neuen, Radar, Rotor, stets; (6) Neffen, nennen, netten, Renner; (7) Rentner, Rotator; (8) Reittier; (9) Geistsieg.

Oddness: $73\% = 0.73 = 29/40$. Average number of letters in word 4.2, the standard deviation is 2 letters.

Several examples of palindromic expressions in German:

Ein Neger mit Gazelle sagt im Regen nie.

Gras mit Niere vereint im Sarg.

DieliebeTote! Beileid!

Minimum - An, am, um, in, im...?

Nie reime, da kann Akademie rein!

O Vogel, leg ovo!

Spanish:

(1) a, e, o, u, y; (3) аса, аја (опозорил), аја (межд.), ala, Ala (Аллах), ама, ара, аса (жарил), аса (ручка), аса (видсока), аса (так...), ата, еје, есе, ојо, оро, осо, ото, ойо, susu; (4) alla; (5) ajaja, alala, anona, radar, rajar, rapar, rasar, rayar, rever, somos, sosos, sotos; (6) sallas, solos; (9) reconocer, (11) aibohphobia.

Ratio of words with odd number of letters equals $90\% \approx 0.9 \approx 37/41$. The average number of letters in words is 3.9, the standard deviation is 1.9. The ratio of the number of palindromes with an odd number of letters $90\% \approx 122/134$.

Several examples of palindromic expressions in Spanish:

La ruta nos aportó otro paso natural.

Somosonosomos?

Nodeseoesedon.

Lamoral? claro, mal.

Etna da luz azul a Dante.

Polish:

Polish is an interesting Slavic language for comparison with Russian. The list of words:

(1) a, à, i, o, u, w, z; (3) aha, asa, bab, bib, bóbb, cyc (כותנה), cyc (דש), dud, gig, jaj, jej (her), jej, kok (תסרוקת), kok (טבה), lal, mam, mim, nań, oho, oko, ono, oto, owo, pip, pop, pup, rur, udu; (4) Anna; (5) kamak, lamal, lapal, madam, magam, makam, maram, masam, matam, mazam, molom, mytym, nagan, nizin, radar, reper, rotor, ususu, zakaz, zaraz; (7) rotator, tartrat.

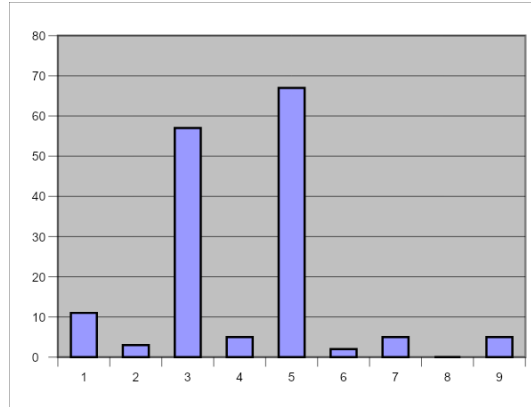
The average number of letters in a word is 3.6, the standard deviation 1.4. Almost all the palindromic words have an odd number of letters. A ratio of oddness $98\% = 0.98 = 57/58$. If you also took into account abbreviations like **SS**, **mm**, then the ratio will be equal to that in Russian. Palindromic expressions are not familiar to us. Therefore we will only bring two combinations composed of two words: kobyly, bok; Nysy syn. We think that to write palindromes in Polish is harder and their number is smaller by contrast to Russian. The reason for this is that many consonants in the language are expressed by two or three letters, and their reflection gives unfamiliar combinations in the language.

Observation of the phenomenon of oddness in two Slavic languages, two German languages and two Roman languages enables us to claim, that the phenomenon of oddness is universal to every language using a letter alphabet.

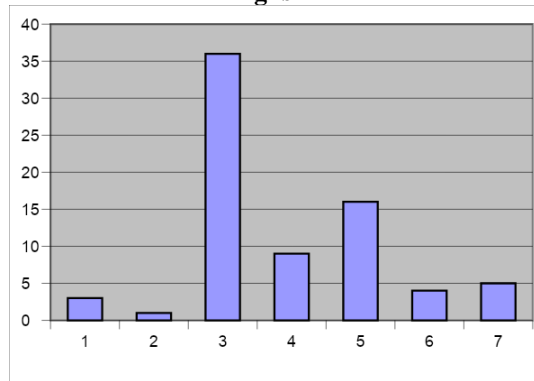
Drawings:

Histograms of distribution of the number of words on the number of letters in Russian, English (A- main list, B- secondary list), French, Spanish, German and Polish. For comparison also a histogram of the distribution of normal words is given according to a dictionary from a Russian dictionary containing 130 thousand words.

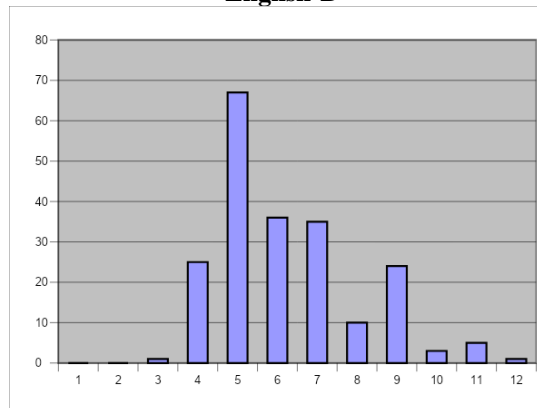
Russian



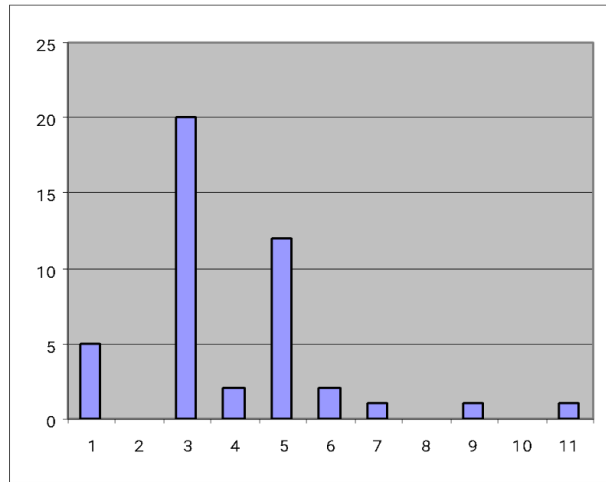
English-A



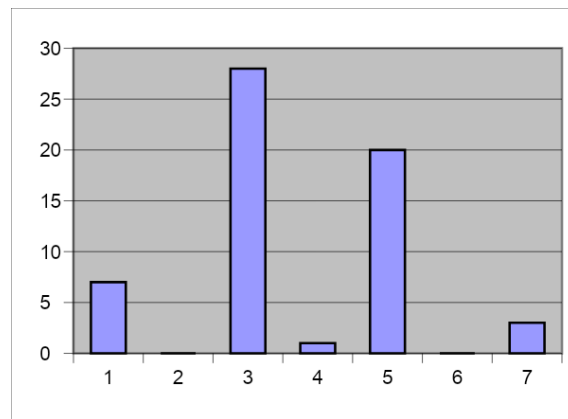
English-B



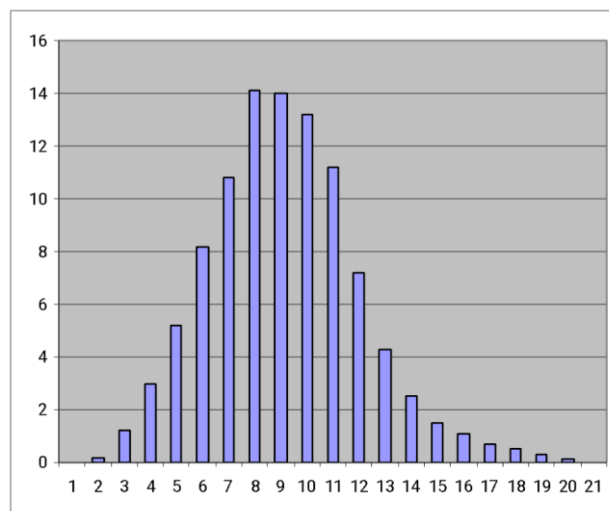
Spanish



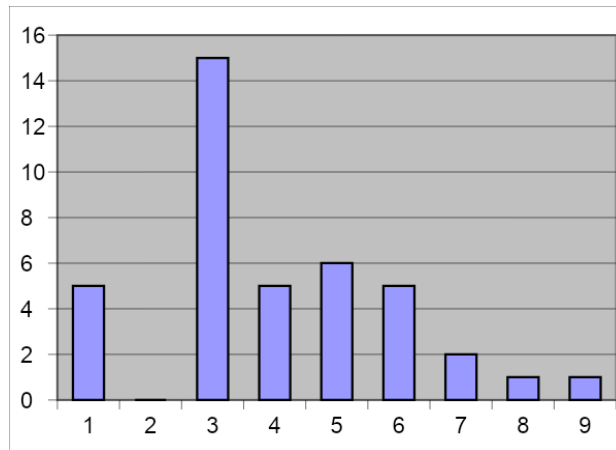
Polish



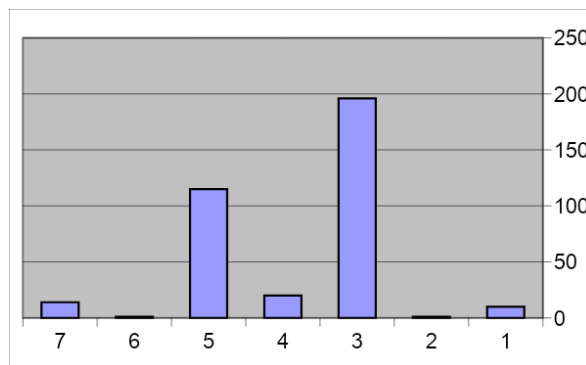
Russian- normal words



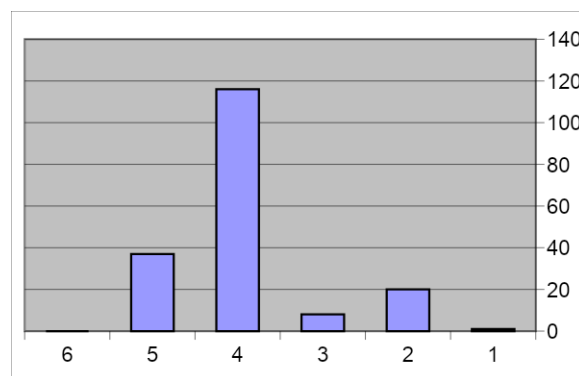
German



Arabic



Hebrew



References

Word Ways: The Journal of Recreational Linguistics. Greenwood Periodicals et al., 1968. ISSN 0043-7980.
The Palindromist. Palindromist Press, 1996.
Howard W. Bergerson. Palindromes and Anagrams. Dover Publications, 1973. ISBN 978-0486206646.

Stephen J. Chism. *From A to Zotamorff: The Dictionary of Palindromes*. Word Ways Press, 1992. ISBN 978-0963515209.

Michael Donner. *I Love Me, Vol. I: S. Wordrow's Palindrome Encyclopedia*. Algonquin Books, 1996. ISBN 978-1565121096