Comparison of Languages

E. ADALI

Measure in Comparing Languages

- *Understandability*
- *Easy to speak, nice sound, energy conservation*
- *Vocabulary*
- *Efficiency*
- *Clarity*
- *Being regular*

Understandability

A method used to measure the communication fidelity of telephone lines can also be used to measure the *understandability* of a language.

$$AD = 1 - \frac{YS_1 + YS_2}{M_1 + M_2}$$

AD: degree of *understandability*

YS: Number of wrong

YS$_1$: The wrong number of the first listener

YS$_2$: The wrong number of the second listener

M: Number of words in the text read

M$_1$: Number of words in the text read by the first reader

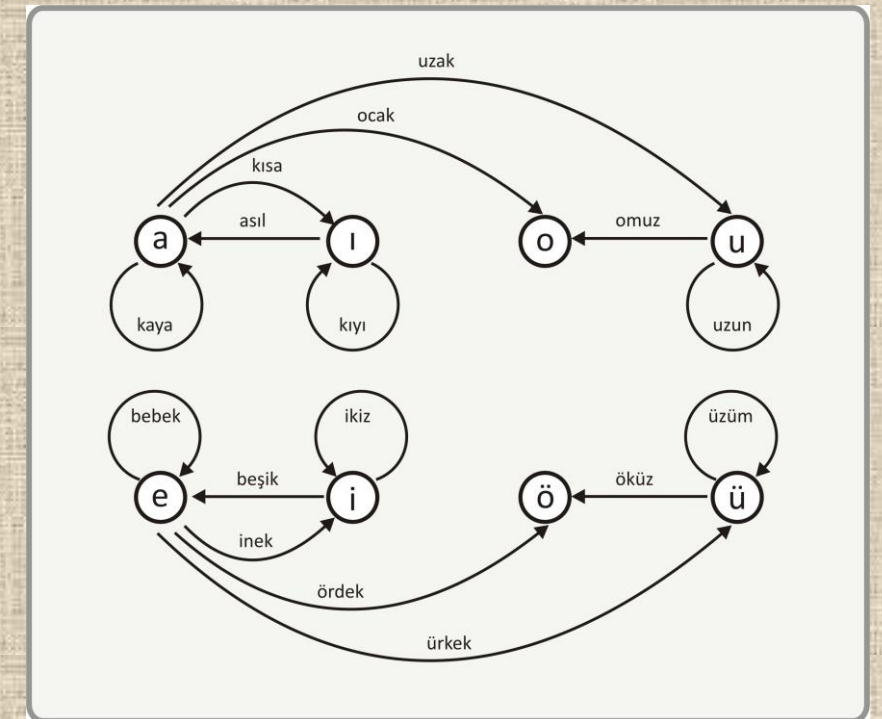M$_2$: the number of words in the text read by the second reader

These experiments show that the understandability of Turkish is high.

Easy of Saying, Nice Sounding, Energy Conservation-I

Vowel Harmony

- The sound we make from the back of our mouth when our lips are straight and open is the "a" sound. After the "a" vowel, we can say "a" without distorting the shape of our mouth and lips, and "ı" with a little change.

- The sound we make from the middle of our mouth when our lips are straight and closed is the "ı" sound. After the "ı" vowel, we can say "ı" without distorting the shape of our mouth and lips, and "a" by changing it very little.

- When our lips are round and open, we can make the "o" sound from the back of our mouths, after the "o" vowel we can make the "o" sound again. However, this situation is not encountered in Turkish words because it does not sound nice. The "u" and "a" sound can be easily produced with a small change in the lip structure.

- While our lips are round and closed, we can make the "u" sound from the back of our mouths, after the "u" vowel we can come back to the "u" vowel or we can make the "a" sound by changing our mouth shape a little.
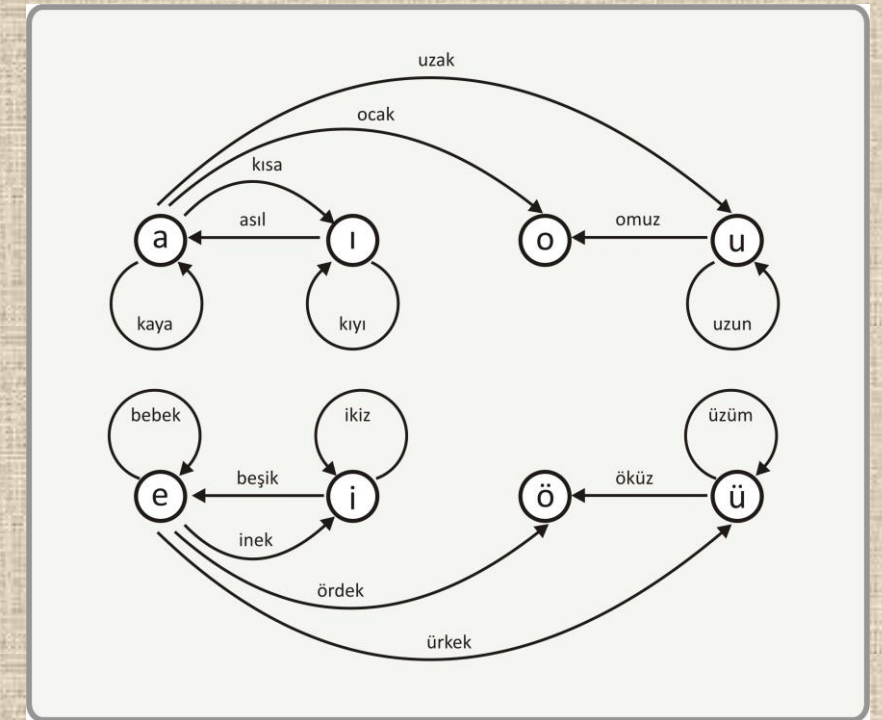

	Front		Middle		Back	
	Flat	Round	Flat	Round	Flat	Round
Close	i	ü	ı			u
Open	e	ö			a	o

Easy of Saying, Nice Sounding, Energy Conservation-II

Vowel Harmony

- One of the sounds we make from the front of our mouth when our lips are straight is "**e**" and the other is "**i**". There is no need to change our mouth structure in order to remove the "**e**" after the "**e**". After the vowel "**e**", we can make a small change in the shape of our mouth and lips, making one of the sounds "**i**", "**ö**" and "**ü**". After the "**i**" vowel, the vowel "**i**" and "**e**" can be easily removed without making any changes. However, it is difficult to say the vowels "**e**" or "**i**" after the vowels "**ö**" and "**ü**".

- When our lips are round and open, we can make the "**ö**" sound from the front of our mouths, after the "**ö**" vowel, we can make the "**ö**" sound again. However, this situation is not encountered in Turkish words because it does not sound nice.

- When our lips are round and closed, we can make the "**ü**" sound from the front part of our mouth, followed by "**ü**" again.

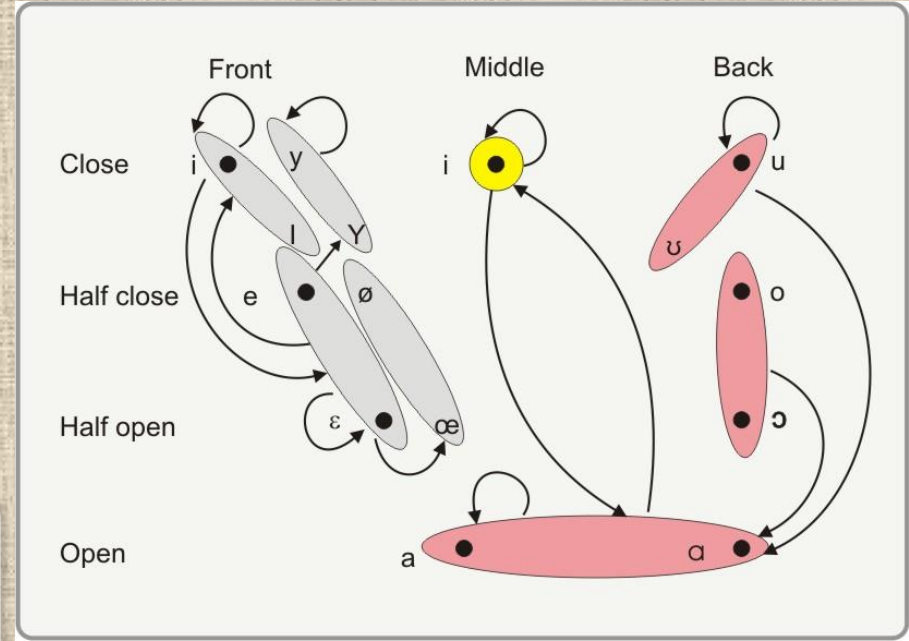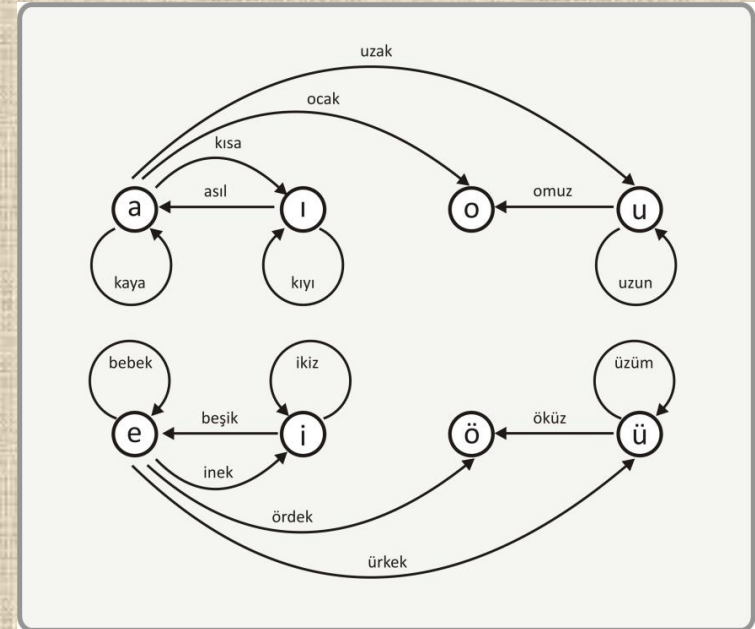|	Front		Middle		Back	
	Flat	Round	Flat	Round	Flat	Round
Close	i	ü	ı			u
Open	e	ö			a	o

Easy of Saying, Nice Sounding, Energy Conservation-III

Vowel Harmony

- The consecutive vowel, which we can remove without any change in our mouth structure, is the same as the first vowel.

- Whatever the condition of our lips when singing the first vowel (open or closed), keeping the same when singing the consecutive vowel makes it easier to say the consecutive vowel.

- Whatever the condition of our lips (flat or round) when singing the first vowel, keeping the same when singing the consecutive vowel makes it easier to say the consecutive vowel.

- Switching from round vowels to straight vowels "o" and "u" can only be done if you stay in the back region. In this case, the transition can only be to the vowel "a".

- Transition from straight "e" to round vowels ("ü" and "ö"), provided that they stay in the same region (front). When we show the transition between vowels in the vowel's quadrant, the comments we made above are clearly visible.
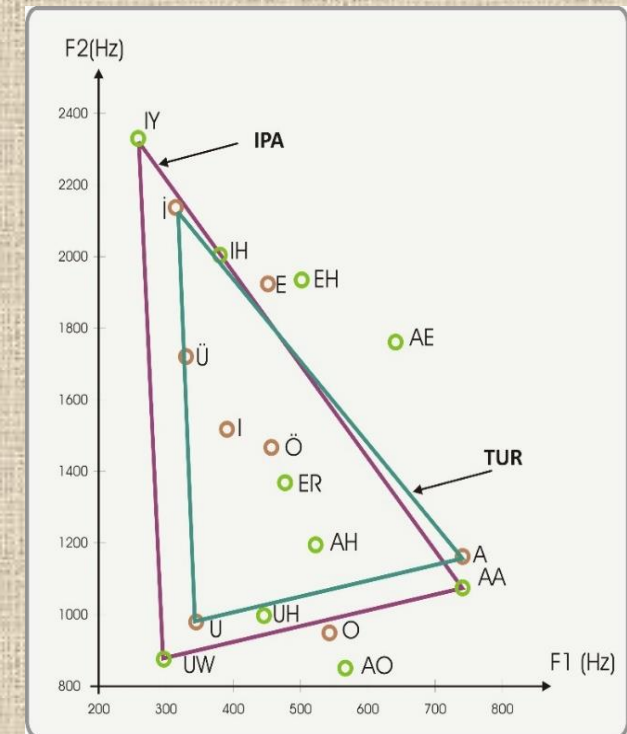

tenteredi, tintoridö

The first word is easier to say and less tiring on our jaws. it is easy to say because the arrangement of the vowels is in accordance with Turkish sound rules.

Formants of Turkish Vowels, IPA Formants

- The basic frequency and formant frequencies of male and female voices are different. In addition, the frequency values given in this table are the average values obtained from many subjects, not one person.

- Figure shows the formant values of Turkish vowels together with the IPA's formant values. As seen in this comparative way, the vowels in the phonetic alphabet of IPA are close to the vowels of Turkish.

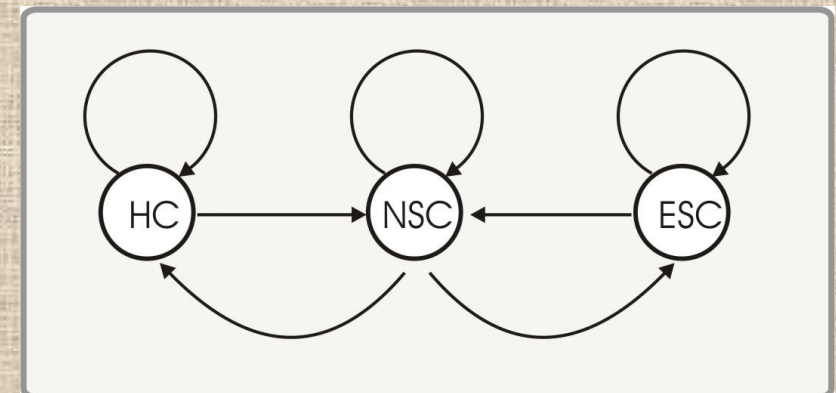| letter	gender	fo (Hz)	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)
a	female	236	771	1338	2998	4168
	male	130	642	1128	2714	3707
e	female	231	578	2205	2961	4128
	male	127	470	1866	2563	3715
ı	female	233	492	1629	2976	4232
	male	128	396	1500	2479	3782
i	female	245	430	2591	3325	4308
	male	138	306	2111	2897	3751
o	female	243	564	959	2976	3794
	male	130	483	860	2733	3668
ö	female	212	543	1636	2764	3947
	male	124	469	1510	2439	3554
u	female	247	452	961	2940	3825
	male	141	379	980	2490	3558
ü	female	234	424	1938	2742	3694
	male	139	333	1769	2337	3342

Harmony of Consonants

- In Turkish, consonants also follow certain rules.

- Consonants are divided into subsets.

- Hard and soft consonants according to the vibration of sound beams

- Continuous and discontinuous consonants, depending on whether the vocal path is blocked or open after the vocal chords.

- Lip, palate, tooth and larynx consonants, depending on where the sound is produced.

Hard Consonants	(HC)	ç, f, h, k, p, s, ş, t
Soft Consonants with No Hard Equivalent	(NSC)	l, m, n, r, y
Soft with Hard Equivalent	(ESC)	b, c, d, g, ğ, j, v, z

	Voiced consonants			Volatile consonants	
	Continuous			Discontinuous	
	fluent	infiltrating	explosive	infiltrating	explosive
Double lip	m		b		p
Tooth-lip		v		f	
Tooth	n, l, r	z	d	s	t
Tooth-palate		j	c	ş	ç
Front palate	n		g		k
Back- palate	ñ	ğ	g		k
Glottal				h	
Half vowel	y				

Frequency of Use of Letters in Turkish

Vowel and Consonant Usage Rates of Different Languages

Languages	Vowel letter frequency (%)	Frequency of consonant use (%)
Italian	47,617	52,383
Finnish	45,65	54,355
French	44,811	55,189
Spanish	44,23	55,77
Turkish	42,82	57,18
Czech	41,52	58,47
Polish	39,20	60,803
German	38,238	61,762
English	38,1	61,9
Swedish	36,19	59

Frequency of Use of Letters in Turkish

Letter	Freq. %	Letter	Freq. %	Letter	Freq. %	Letter	Freq. %
a	11,46	ı	4,56	ü	1,92	c	0,92
i	9,32	t	3,60	ş	1,53	p	0,87
e	9,07	m	3,51	z	1,5	ö	0,77
n	7,42	y	3,32	g	1,15	f	0,49
r	7,04	s	3,15	h	1,11	j	0,05
l	6,4	u	3,14	ğ	1,047		
K	4,65	b	2,67	ç	1,046		
d	4,6	o	2,58	v	1,01		

According to the results of this study, the ratio of vowel used in Turkish to all letters is 42.82% and consonants 57.18%.

Easy of Saying, Nice Sounding, Energy Conservation

Easy to Speak

Nice Sound

Energy Conservation

The harmony of vowels and consonants in Turkish facilitates the production of these sounds in the sound production organ.

According to a research conducted by the Department of Phonology of the University of Trier in Germany:

- Vowels are the melody sounds of a language.

- The aesthetics of a language is directly proportional to the number of its vowels.

- The longer the duration of the vowel sounds brings fluency to the language.

- The results of this research show why Turkish sounds good. There are 13 vowel sounds in Turkish. In addition, the duration can be doubled when used with the thick ordinary vowel sounds "ğ".

The rules in the arrangement of vowels and consonants in Turkish are the natural result of the structure of our vocal organs. When we want to make a sound without tiring or forcing our vocal organs, we need to follow the vowel and consonant rules. When we evaluate it from this point of view, we can say that Turkish is an energy-saving language. The jaw of a Turkish-speaking person does not get tired.

Vocabulary-I

The effectiveness of a language can also be measured by the richness of its vocabulary. The size of a language's vocabulary is directly proportional to the person's ability to think. When investigating the richness of a language's vocabulary, looking only at the number of words in the dictionary does not give correct results.

- Hami-Sami languages have the ability to derive new words from a root. KTB

  ketebe, kitab katib, katibe, mekteb
  (*clerks*, book, *clerk, lady clerk*, s*chool)*

- Turkish words only have suffixes.
- English can also have prefixes.

- These people are the ones who used to use negative suffixes to words according to Arabic and Persian rules in the past.

  mevcut, namevcut (*available, absent).*
  In the same habit, they also advocate giving words a gender feature.
  memur, memure (*officer, lady officer*).

Turkish Word Derivation Ability

Turkish	English	Turkish	English
Göz	Eye	Gözlemci	Observer
Gözlük	Eyeglasses	Gözlemcilik	Observation
Gözlükçü	Optician	Gözde	Favourite
Gözlükçülük	Opticians	Gözgü	Mirror
Gözcü	Watchman	Gözgülük	Mirror stand
Gözcülük	Opthamolgy	Gözgücü	Mirror maker
Gözlem	Observation	Gözgücülük	Mirror makers business
Gözleme	Observing		

Prefixes are not used in Turkish

Turkish	English
Olanaksız	Impossible
Yeniden canlandırmak	Reactivate
Tepkin	Reactive
Düzeltmek	Reform
Mutsuz	Unhappy
Önlem	Precaution
Ön yargılı	Preconceived
Özürlü	Disabled

Vocabulary-II

While evaluating the richness of a language's vocabulary, it is not enough to look only at the noun noble words, it is also necessary to examine the verb noble words. Apart from the basic form of every action in Turkish, such as active, passive, reflexive, active, and causative states provide a significant advantage over other languages.

Predicate type	Turkish	English
Active	Görmek	To see
Passive	Görünmek	To be seen
Passive	Görülmek	To be seen
Reflexive	Giyinmek	To dress
Reciprocal	Görüşmek	To see each other or to discuss
Causative	Görüştürmek	To bring someone to see or to discuss each other
Causative in second degree	Görüştürtmek	To have somebodies to see or to discuss each other
Passive causative	Görüştürülmek	To be brought to see or to discuss somebody

- Turkish is that words do not change according to gender.

    melik-melike (kink-quin), rahip-rahibe (priest-nun), memur-memure (officer-lady officer)

    öğretmen hanım (lady teacher) or hanım öğretmen (lady teacher).

- In languages where words are given femininity and masculinity, for example, in French and Arabic,

    French, *la mur* (wall), *le port* (door), the word wall is female and the word door is male. It is really funny and difficult to understand why such a distinction is necessary, perhaps it is meaningless.

Efficiency

- Adding a suffix to words in agglutinative languages give new meanings to words. These features not only enrich the vocabulary of the language, but also make the language more efficient. Active and passive forms of actions are seen in all languages. However, forming reflexive, active, causative, secondary causative, causative passive forms of actions is not as easy and regular as in Turkish. It is not even possible in some languages.

- *kırmak* (*breaking*) is to (*smash hard things by hitting or crushing*). It is possible to come across many meanings of the word to break in developed languages such as Turkish. For example:

  - *Soğuk hayvanları kırdı (öldürmek) – Cold kill animals*
  - *Fiyatları kırmak (indirim yapmak) - Breaking prices (discounting)*
  - *Kalbimi kırdı (gücenmek) - It broke my heart (offended)*
  - *Pulunu kırmak (tavla oyununda) - Breaking your checker (in backgammon game)*
  - *Buğdayı kırdırmak (kaba öğütmek) - Crushing the wheat (coarse grinding)*
  - *Direksiyonu kırmak (Çevirmek, yönünü değiştirmek) - Breaking the steering wheel (turning, changing direction)*
  - *Soğuğun belini kırmak (Soğuğun etkisini ciddi olarak azaltmak) - Breaking the waist of the cold (seriously reducing the effect of the cold)*
  - *Senet kırmak (senedi paraya çevirmek) - Issuing bills (converting the bill into cash)*
  - *Dersi kırmak (dersten kaçmak) - Breaking the lesson (avoiding the lesson)*

Clarity

A certain order is expected in the order of the words in the sentence. In every language, certain rules are followed in ordering the words in sentences. If this order is changed in some languages, the meaning of the sentence is distorted or lost. In Turkish sentences, the words are arranged in a certain order. However, there is no loss of meaning when the order is changed. However, the difference in emphasis occurs. For example, the meanings of these three sentences are the same, the emphasized word is different.

- *Elini sabunla yıka.* - Wash your hand with soap. (emphasize hand)

- *Sabunla elini yıka.* - Wash your hands with soap. (emphasize soap)

- *Yıka elini sabunla.* - Wash your hand with soap. (emphasize washing)

The sentence structure in Indo-European languages is rigid. Therefore, when the places of the words in the sentence are changed, the meaning of the sentence changes or deteriorates. The flexibility of sentence structure increases the meaning power of Turkish and provides skill and convenience to the speakers. However, this feature causes additional difficulties in natural language processing studies.

Being Rules – Cases

- It is as if a language congress was held 5000 years ago and the rules of Turkish were determined. These rules have not changed until today. This feature of Turkish astonishes linguists. The unchanging rules of Turkish have never been broken in phonics and morphology. An exception can be given in the rules of morphology. As it is known, when possessive suffixes are added to a word ending with a vowel, the letter "**n**" is added in between. According to this rule, when we should say **su+n+un** but say *suyun*. This situation, which we can consider as the rule breakdown, actually stems from the fact that the word *su* was used as *suy* in the past.

- We can also examine languages in terms of case suffixes. The meaning of the case suffixes in Turkish is very precise. This certainty is not at the same level in Indo-European languages. For example:

  - *Müziği dinliyorum*. - I'm listening to the music. (I am listening to music)
  - *Partililere konuşuyorum*. - I am speaking to the party members. (I am speaking to members of parties)
  - *Seni düşünüyordum*. - I was thinking of you. (I think of you)

- As it is known, there are six cases of nouns in Turkish: Lean, **-i** (indicating), **-e** (orientation), **-de** (presence), -den (leaving), **-in** (affiliation). These attachments can be thought of as **vectors** because they indicate direction. It cannot be wrong in the meanings given by the case suffixes. For example, *someone is listening*; *is told to someone*. Note that the same direction indicator (**to**) is used for these two sentences in English.

Being Rules-Verbs

The roots of Turkish verbs never change. For example, if we write the action to go for all times

git+tim, **git**+mişim, **gid**+iyorum, **gid**+eceğim, **gid**+erim)

we can clearly see this feature. In the example we have given, the letter "t" can turn into "d" due to the phonetic rule, and this is a rule. There is not even a letter similarity between the present and the past of the act of going in the Indo-European language family, Persian, French and English. For example,

- I go (present tense)   I went (past tense) (English)
- Je vais (present tense)   Je suis allé (past tense) (French)
- Hahem reft (present tense)   reftem (past tense) (Persian)

Such irregularities are not encountered in the shooting of Turkish verbs .

Being Rules-Syllable - I

- It can be said that there are 4840 syllables in Turkish. Accordingly, it can be said that when 4840 syllables are voiced, Turkish text vocalization can be realized.

- There is no need to use a dictionary to separate a Turkish word into its syllables, it can be accomplished using a very easy algorithm.

- Syllable ending with a vowel are called open vowels, and those ending with consonants are called closed vocalizations. Univocal syllable in Turkish must also conform to one of these six forms. Most of the monophonic root words are closed syllable. *Al, vur, sev, at, kuş, aş.* Turkish syllable are as follows:
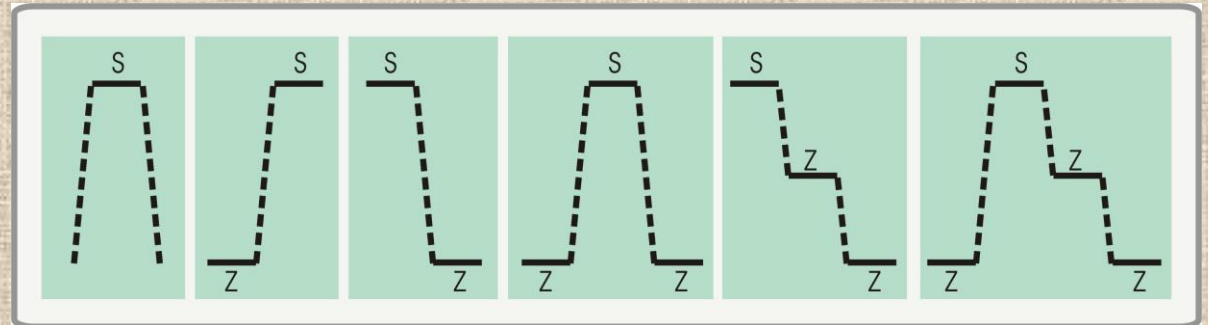
1. *The first of the two consonant sounds side by side in the word forms a syllable with the vowel before it, the second with the vowel after it:* **bir-lik**
2. *The first two of the three side by side consonants in the word form the syllable with the vowel before it, and the third with the vowel after it:* **kork-maz, Türk-çe**
3. *The first two of the four side by side consonants in the word form the syllable with the vowel before it, and the last two with the vowel after it.*

Syllable Patterns in Turkish and Their Possible Numbers

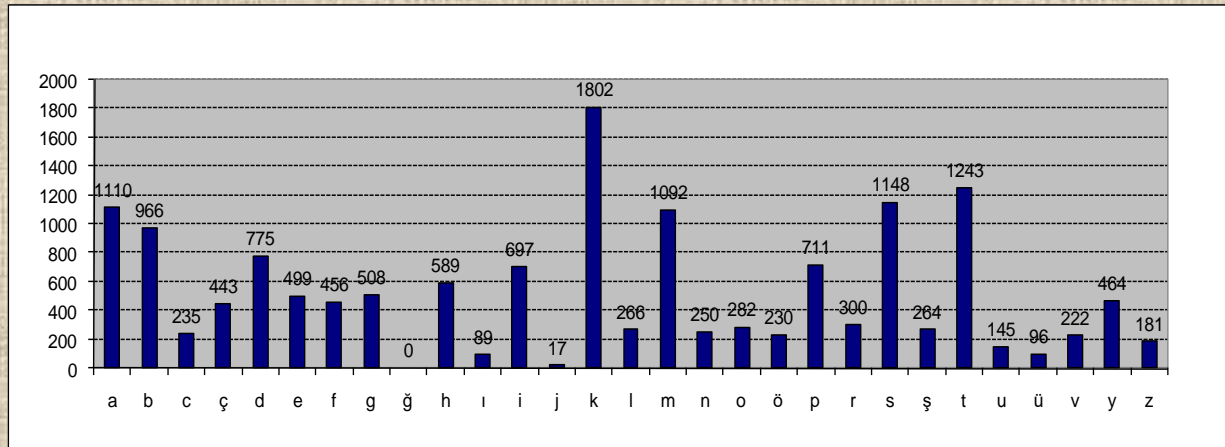| Syllable structure	S	SZ	SZZ	ZS	ZSZ	ZSZZ
Possible number	8	128	112	152	2312	2128

(S: vowel and Z: consonant)

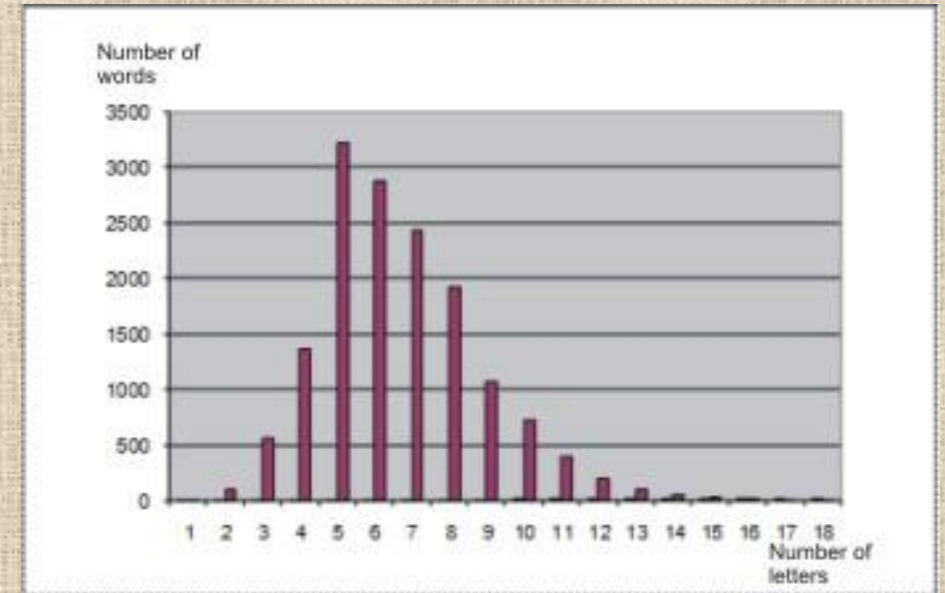### Peak ups and downs in Turkish syllable
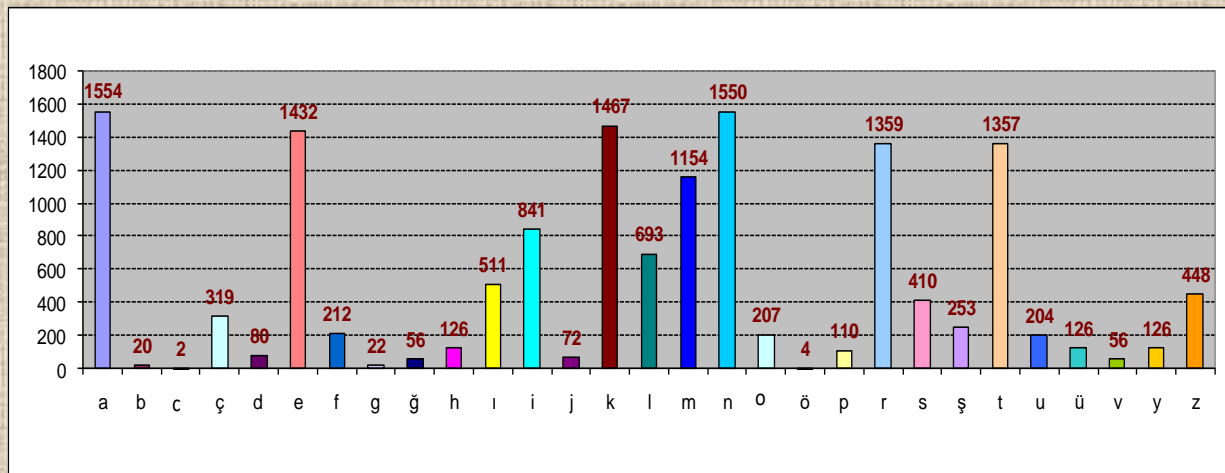
Being Rules-Syllable - II

Can an artificial dictionary be created in accordance with the phonetic rules of Turkish?

Statistical results of Turkish words according to their first letters

Statistical results of Turkish words according to their last letters

Statistics of Turkish word lengths

Being Rules-Syllable - III

The syllable structures of words consisting of two syllables are as follows:

[S + ZS]	*(aba)*	*[S + ZSZ]*	*(ocak)*	*[S + ZSZZ]*	*(avurt)*
[SZ + ZS]	*(ayna)*	*[SZ + ZSZ]*	*(aşkın)*	*[SZ + ZSZZ]*	*(aldanç)*
[ZS + ZS]	*(baba)*	*[ZS + ZSZ]*	*(tabak)*	*[ZS + ZSZZ]*	*(kazanç)*
[ZSZ + ZS]	*(bohça)*	*[ZSZ + ZSZ]*	*(bostan)*	*[ZSZ + ZSZZ]*	*(başkurt)*

As a result, Turkish words consisting of one and two syllables can be formed in 18 different ways. Meanwhile, in addition to the basic sound features described in this section, the following basic rules should be kept in mind:

- *Since three consonants cannot be side by side, the SZZ and ZSZZ syllables cannot form the initial syllable.*
- *Two vowels are not considered side by side.*
- *If there are two consonants in the same notation, this consonant pair is one of the following: lç, lk, lp, lt, nç, nk, nt, rç, rk, rp, rs, rt, st, sht*
- *"j" does not exist in Turkish-origin words, so it does not need to be included in consonants.*
- *There are no Turkish words that begin with "ğ". If there is a suffix that starts with a consonant and ends with a consonant, then the second notation does not begin with "ğ".*
- *While forming syllables, the rule of resembling hard consonants should be applied.*
- *The letter "o" can only be in the first sound of the word.*
- *When the suffix that begins with a vowel ending with a consonant, they cannot be the same letter.*

Being Rules-Syllable - IV

There may be 3.039.208 words in Turkish consisting of two syllables.

Monosyllabic words		Two syllable words			
structure	number	structure	number	structure	number
S	8	S + ZS	342	ZS + ZS	6.498
SZ	128	S + Z SZ	5.202	ZS + ZSZ	98.838
SZZ	112	S + ZSZZ	4.788	ZS + ZSZZ	90.972
ZS	152	SZ + ZS	4.914	ZSZ + ZS	88.794
ZSZ	2312	SZ + ZSZ	74.574	ZSZ + ZSZ	1.347.534
ZSZZ	2128	SZ + ZSZZ	68.796	ZSZ + ZSZZ	1.243.116

The number of Turkish words using ZSZZ as the second syllable seems high, but words that fit this pattern are rare in our language. Even if all words that fit this pattern are ignored, the number of two-syllable words becomes 1,636,324.