E. Adali ⓘD

Istanbul Technical University, Istanbul, Turkey
email: adali@itu.edu.tr

THE LOGIC OF TURKISH LANGUAGE FOR NLP

One of the applications of NLP is Question and Answering (QA) system. In this field there are many researches and applications. Some competition is organized to increase the performance of the systems. Turkish is a language designed as if by a scientific committee. In Turkish language, rules are clear and transparent. In order to prove these features of Turkish, rules regarding phonology, morphology and syntax features have been deduced, a logical expression has been written for each rule and a flow diagram has been drawn. Indexing of pages written in Turkish is somewhat difficult than any India-European pages. Although an English word can get one prefix and one suffix Turkish words can take many suffixes. Therefore, we need to discover the root or stem of the word.

The rules of Turkish were tested with the programs developed in accordance with the flow diagram. As a result of the tests, it has been proven that the rules of Turkish are intact. Finally, a study was conducted to artificially create a sentence and a program was prepared.

Keywords: Turkish, Logic, Phonology, Morphology, Syntax.

1. Inroduction

Language and its transcription can be considered as the most important developments in human history. When the anatomical structure of humans is examined, it is suggested that they started talking approximately 160,000 years ago. Until recently, it was said that the history of writing dates back 6,000 years, but M. Ercan, head of excavations at Göbeklitepe, said that the scene on an obelisk found during the excavations in Göbeklitepe could be the "first pictograph" because it describes an event thematically [1]. It is estimated that Göbeklitepe was built 12,000 years ago. There are opposing views on how people form their languages. While some of these views say that God taught the language, others say that the language developed spontaneously over time.

It is an undeniable fact that language and writing have a very important contribution to the development of societies. Confucius (551 BC – 479 BC) explains the relationship between language and society as follows:

"... If I were to take over the administration of a country, the first thing I would do would undoubtedly be to review its language. Because if the language is flawed, words cannot express the thought well. If the thought cannot be expressed well, tasks and services cannot be performed properly. In places where duties and services cannot be performed properly, *customs, rules and culture are disrupted. If customs, rules and culture are broken, justice will go astray. If justice goes astray, the confused people will not know what to do or where the situation will lead. That's why nothing is as important as language!"* [2,3].

Prof. Dr. David Cuthell says the following about Turkish:

- "Turkish is such a language that it is as if a hundred graduate mathematics professors came together and created Turkish.

- Turkish is such a language that it is a language of emotion, thought, logic and philosophy in itself.

- While other languages are based on learning the static meanings of words, Turkish is based on finding and extracting these meanings, that is, on dynamic meaning. "

Prof. Dr. Here is what Johan Vandewalle said about Turkish in his own words

- "...I believe that a native speaker of Turkish thinks in short sentences and, while speaking, creates complex structures by connecting these short sentences together in various ways. Although I have studied many languages belonging to different language groups, no language has ever struck me as much as the complex sentence structures in Turkish. I can say that I have not come across a structure that fascinates me.

- If you allow me to be a little emotional, sometimes I say to myself, 'I wish Chomsky had

learned Turkish when he was young...' I am sure that then contemporary linguistics would have been shaped according to Turkish, not English..."

Max Müller made the following evaluation about Turkish in his book "Lectures on The Science of Language" published in 1862 [4]:

- *"Even if you have no desire to use Turkish grammar, reading Turkish grammar is truly a great pleasure. The way in which the numerous grammatical forms are brought out, the regularity of the inflectional system, the transparency and comprehensibility of the whole structure should impress anyone who feels the wonderful power of human intelligence that shines in the language."*

- *On the contrary, the grammar of Turkic languages has a transparent structure and an order in which we can examine its inner workings as if we were watching the construction of honeycombs in a crystal beehive. A well-known orientalist says, 'We can imagine that Turkish is the result of the work of some distinguished scholars.' However, the self-contained communities in the Tatarstan steppes could not have designed something so perfect with their innate laws or instinctive powers alone."*

2. Methods

2.1. Phonology of Turkish

The most basic component of language is sounds and the letters corresponding to these sounds. It is known that the first alphabet in its current sense was developed by the Phoenicians, sailors from the Semitic race. There are 22 letters in this alphabet, which is thought to have been used in 1000 BC, and all of them are consonant letters [8,11]. The Phoenician alphabet was taught to the nations bordering the Mediterranean by this seafaring nation. After these transfers, Arabic, Greek and later Latin alphabets were formed. The Old Latin alphabet dates back to the VI century BC. It is said that it was derived from the Etruscan alphabet in the century.

Göktürk alphabet was used in Orkhon monuments. There are 26 consonants, 4 vowels and 8 compound letters in this alphabet. Letters are written separately from right to left. Words are separated by placing ' : ' between them.

Turks who came to Anatolia began to use the Arabic alphabet under the influence of Islam, which they adopted. The three vowels of the Arabic alphabet were not enough for Turkish, which is rich in vowels. This lack of Arabic alphabet has been the source of important errors in the text [13]. For example, the words *gel, bald* and *rose* had to be written with the same letters (K and L).

2.2. Sounds of Turkish

There are a total of 29 letters in the Turkish alphabet. These letters are "a, e, ı, i, o, ö, u, ü" vowel, "b, c, ç, d, f, g, ğ, h, j, k, l, m, n, p, r, s, þ, t, v, y, z" are called consonants. The Turkish alphabet, which was developed based on the Latin alphabet, is a phonetic alphabet, so one letter corresponds to one sound.

There is vowel and consonant harmony in Turkic languages. The vowel harmony of Turkish is seen in Figure 1 with its Finite State Machine (FDM) representation. Consonants are clustered as shown in Table 1.

Table 1 – The Clusters of Consonants

Voiceless (hard)	ç, f, h, k, p, s, ş, t
Voiced (hard equivalent)	b, c, d, g, ğ, j, v, z
Voiced (constantly fluent)	l, m, n, r, y
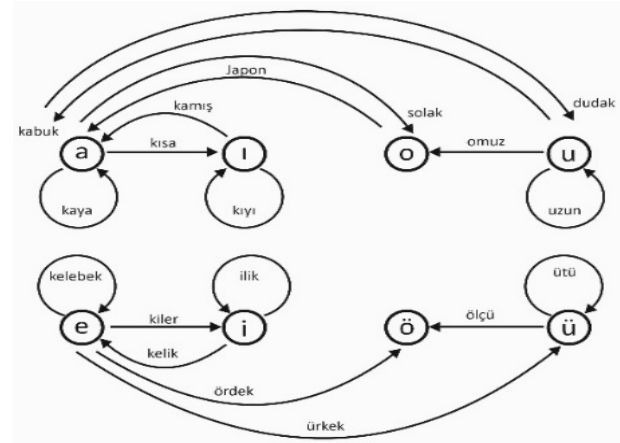
Figure 1 – The vowel harmony of Turkish

The consonant harmony rules are as follows:
- The probability of a hard consonant coming after a hard consonant is higher than the probability of a soft consonant coming.
- The probability of a hard consonant being followed by a consonant without a hard equivalent seems to be higher than consonants with a hard-consonant equivalent.
- The probability of a hard consonant being followed by a soft consonant with a hard counterpart is in third place.
- After a soft consonant that does not have a hard equivalent, a soft consonant that does not have a hard equivalent comes first.
- After a soft consonant that does not have a hard equivalent, a hard consonant with a hard equivalent can come in the second place and a soft consonant

with a hard equivalent can come in the third place. First comes a consonant that has no hard equivalent.

- Consonants in a word are generally chosen from the same cluster.

- After the dental consonants, one of the palatal consonants may most likely come, and then one of the labial consonants.

- After the palate consonants, a dental consonant may most likely come, and then one of the labial consonants.

2.3. Voice Changes

The softening of discontinuous hard consonants is considered important in sound changes. The basic rules of softening are as follows:

- When a suffix starting with a vowel is added to words ending with a discontinuous hard consonant (**p, ç, t, k**), the hard consonants soften and turn into (**b, c, d, g**).

- If the consonant **k** falls between two vowels, it becomes **k, ğ**.

Rules for words consisting of more than one syllable:

- The last letter of words ending in **-ıt** does not change.

- The last letter **t** of words ending in **-ut** does not change, but *angut* is an exception.

- Some words consisting of more than one syllable do not comply with the above rule.

- The last letter t of words ending in **-çüt** does not change, but the last letter of the words *öğüt* and *söğüt* turns into d. The constancy in the word *ölçüt* is determined by the three penultimate hard consonants. *Ölçüt* is a single example.

- Only one-word *kelkit* ending in **-it** has been found. This word is used colloquially. Therefore, it would not be right to take it as general rules.

Other cases where the last letter does not change:

- If there is structure and voice suffix at the end of the word, the last letter of the word does not change even if the suffix starts with a vowel.

Words taken from foreign languages adapt to the sound characteristics of Turkish over time, but some do not.

Rules for monosyllable words containing two or three letters:

- The rule that the letters **ç, k, p, t** between two vowels do not change is a general rule applied to single vowels with two or three letters.

- Six nouns and five verbs do not comply with the general rules:

uç ⇒ uc, but ⇒ bdu, tat ⇒ tadı, gök ⇒ göğü, dip ⇒ dibi, kap ⇒ kabı

tat-, et-, dit-, git- güt-

Softening rules for syllable with four letters:

The softening rules for vowels containing four letters are as follows:

Two consonants can occur consecutively in the same sound, but these consonants obey the following rule:

- The leading consonant must be **l, n, r, s** or **ş; ç, k, p** or **t** after **l; ç, k, t** after **n; ç, k, p, s** or **t** after **r; t** may come after **s** and **ş**.

- In single vowels with four letters, the last letter is softened by complying with these prerequisites, but if the last two letters are **rk** and **nt**, it is not softened.
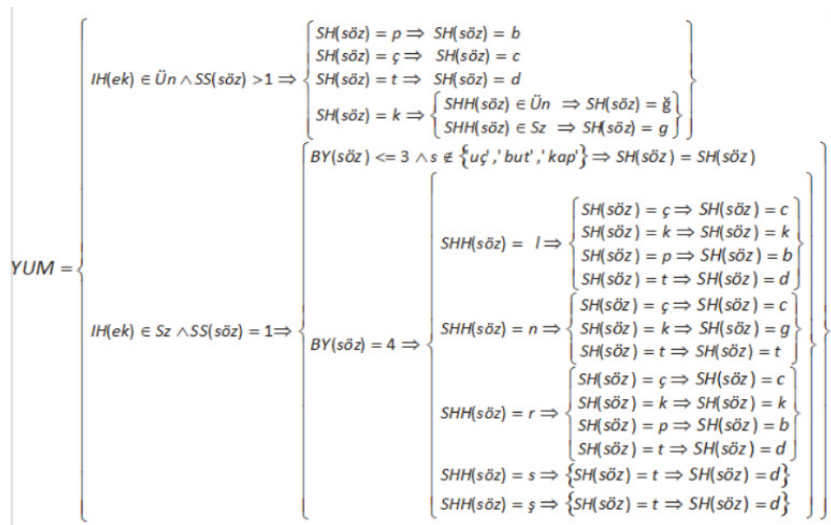
Softening rules are seen in Figure 2.

Figure 2 – The logic of softening rule of hard consonants (YUM: Softener, IH: First letter of, SH: Last letter of, SHH: Letter before last, Ün: Vowel, SS: Number of slybles, BY: Length of word)

2.4. Syllable Structure

The syllable structure of Turkish is regular and is as follows:

```
    [S] (o)        [SZ]    (at)    [SZZ]    ( ü s t )
[ZS] (su) [ZSZ] (tam)      [ZSZZ] (Türk)
    [S + ZS]  (aba)                  [S + Z S Z ]
(ocak)    [S + ZSZZ]      (avurt)
          [SZ + ZS]       (ayna)   [ S Z  +  Z S Z ]
(aşkın)   [SZ + ZSZZ]     (aldanç)
          [ZS + ZS]       (baba)            [ZS + ZSZ]
(tabak)   [ZS + ZSZZ]     (kazanç)
          [ZSZ + ZS]      (bohça)  [ Z S Z  +  Z S Z ]
(bostan)  [ZSZ + ZSZZ] (başkurt)
          [ZSZZ + ZS       (Türkçü) [ZSZZ    +    ZSZ]
(Türklük)
```

3. The Feature of Morphology of Turkish

Turkish is an agglutinative language and derivational and inflectional suffixes are added to the end of the root word. Inflectional suffixes create bonds between words and make a significant contribution to the understandability of the language [12]. Derivational suffixes enable the derivation of new words. Case suffixes included in inflectional suffixes make an important contribution to showing the direction of the action. The inflection of nouns clearly defines the connections between words by giving plural, possessive and genitive suffixes to nominal words. Inflectional suffixes added to verbs show who, when and in what situations an action was performed.
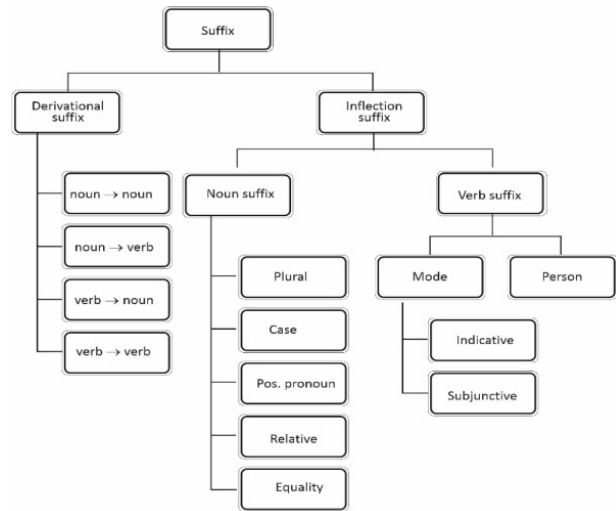
Figure 3 – Derivational and Inflection suffixes

Derivational suffixes give new meanings to root or stem words. Thus, new words can be derived. Since the root does not change in newly coined words, the meaning of the derived word is derstood by the speaker.

Suffixes can follow each other and there is no limit on their number. Research shows that all words in Turkish have approximately 0.94 suffixes; It shows that words with suffixes receive 1.85 suffixes.

It is known that there are about 170 derivational suffixes in Turkish and about 70 of them are frequently used. Inflection and derivational suffixes are shown in Figure 3.

Noun Conjugation. The noun conjugation rule is seen in Figure 4.
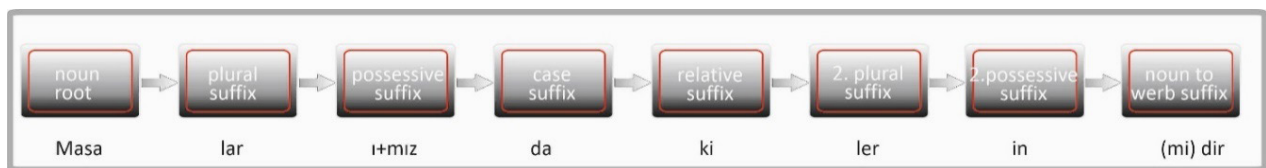
Figure 4 – Noun conjugation rule

In order for a joint to reach the next joint, some adaptations must be made in accordance with the sound characteristics of Turkish. In this stage, which we call the conflict process,

- vowel consonant harmony,
- adding fusing letters,
- softening of hard consonants and
- binding vowel harmony rules apply.
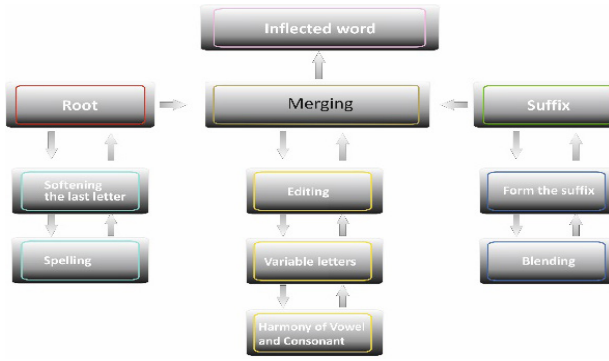The integration process is shown in Figure 5.

Figure 5 – Integration processes

Plural Suffix. The plural suffix is *+lAr*. *A* becomes a or e depending on the last letter of the root. The logical relationship of the plural suffix is as follows:

$$A = (e \vee i \vee \ddot{o} \vee \ddot{\imath}) \wedge (a \vee e)$$

Accordingly, it is written as Söz: word, S: Last vowel of

$$s\ddot{o}z + \begin{cases} S(a \vee \imath \vee o \vee u) \Rightarrow + lar \\ S(e \vee i \vee \ddot{o} \vee \ddot{u}) \Rightarrow + ler \end{cases}$$

Question Suffix. The question suffix is in the form *+mH*. H varies depending on the last letter of the root. The logical relationship of the question suffix is as follows:

$$H = (a \vee \imath) \wedge (\imath \vee i \vee u \vee \ddot{u}) = \imath$$
$$H = (o \vee u) \wedge (\imath \vee i \vee u \vee \ddot{u}) = u$$

$$H = (e \vee i) \wedge (\imath \vee i \vee u \vee \ddot{u}) = i$$
$$H = (\ddot{o} \vee \ddot{u}) \wedge (\imath \vee i \vee u \vee \ddot{u}) = \ddot{u}$$

$$s\ddot{o}z + \begin{cases} S(a \vee \imath) \Rightarrow 'm\imath\,?' \\ S(e \vee i) \Rightarrow 'mi\,?' \\ S(o \vee u) \Rightarrow 'mu\,?' \\ S(\ddot{o} \vee \ddot{u}) \Rightarrow 'm\ddot{u}\,?' \end{cases}$$

$$H = (a \vee \imath \vee o \vee u) \wedge (\imath \vee i) = \imath$$

Possessive Suffix: Kaş*+(H)possessive suffix*.

Table 2 – Suffix, person, samples

Suffix	Person	Samples
+(H)m	1.tekil	kuşum
+(H)n	2.tekil	kuşun
+(s)H	3.tekil	kuşu
+(H)mHz	1.çoğul	kuşumuz
+(H)nHz	2.çoğul	Kuşunuz
+lArH	3.çoğul	Kuşları

Vowels binding rule:

$$BA\breve{G} = \begin{cases} a \vee \imath \Rightarrow \imath \\ e \vee i \Rightarrow i \\ o \vee u \Rightarrow u \\ \ddot{o} \vee \ddot{u} \Rightarrow \ddot{u} \end{cases} \quad ya\ da\ BA\breve{G}(s)$$

$$s\ddot{o}z = \begin{cases} BR\ 1.tek \Rightarrow YUM(s\ddot{o}z) \begin{cases} SH(s\ddot{o}z) = \ddot{U}n \Rightarrow +\ "" \\ SH(s\ddot{o}z) = Sz \Rightarrow +BA\breve{G}(s\ddot{o}z) \end{cases} +m \\ BR\ 2.tek \Rightarrow YUM(s\ddot{o}z) \begin{cases} SH(s\ddot{o}z) = \ddot{U}n \Rightarrow "" \\ SH(s\ddot{o}z) = Sz \Rightarrow + BA\breve{G}(s\ddot{o}z) \end{cases} +n \\ BR\ 3.tek \Rightarrow YUM(s\ddot{o}z) \begin{cases} SH(s\ddot{o}z) = \ddot{U}n \Rightarrow + s \\ SH(s\ddot{o}z) = Sz \Rightarrow +"" \end{cases} + BA\breve{G}(s\ddot{o}z) \\ BR\ 1.\c{c}o\breve{g} \Rightarrow YUM(s\ddot{o}z) \begin{cases} SH(s\ddot{o}z) = \ddot{U}n \Rightarrow +"" \\ SH(s\ddot{o}z) = Sz \Rightarrow +BA\breve{G}(s\ddot{o}z) \end{cases} +m + BA\breve{G}(s\ddot{o}z) +z \\ BR\ 2.\c{c}o\breve{g} \Rightarrow YUM(s\ddot{o}z) \begin{cases} SH(s\ddot{o}z) = \ddot{U}n \Rightarrow + "" \\ SH(s\ddot{o}z) = Sz \Rightarrow +BA\breve{G}(s\ddot{o}z) \end{cases} +n + BA\breve{G}(s\ddot{o}z) +z \\ BR\ 3.\c{c}o\breve{g} \Rightarrow I + \begin{cases} S(s\ddot{o}z) = Ka \Rightarrow a \\ S(s\ddot{o}z) = \dot{I}n \Rightarrow e \end{cases} +r + BA\breve{G}(s\ddot{o}z) \end{cases}$$

Figure 6 – Logical relationship of possessive suffixes (YUM: softener, Ün: vowel, Sz: Consonant, BR: Personal, tek: singular, çoğ: Plural, SH: Last letter of, Ka: Hard vowel, İn: Soft vowel)
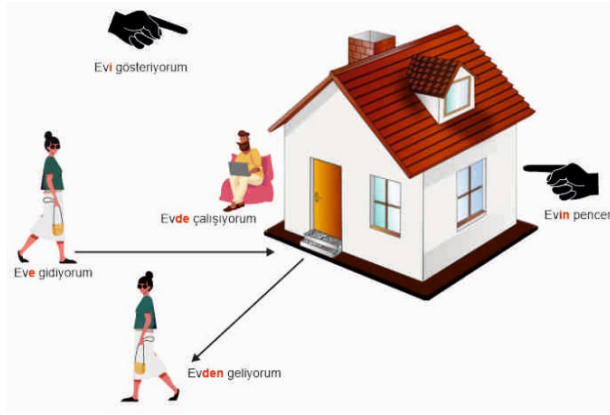
Figure 7 – Meaning of case suffixes

Cases. There are 6 main and three special case suffixes in Turkish:

subjective case		ev	
+(y)H	(I case: accusative)	ev**i**	
+(y)A	(e case: dative)	ev**e**	
+DA	(de case: locative)	ev**de**	
+DAn	(den case: ablative)	ev**den**	
+(n)Hn	(in case: genetive)	ev**in**	

There are also these case suffixes:

+DAki	(relative suffix) kapı	kapı**daki**
+(y)lA	(instrumental) uçak	uçak**la**
+CA	(equality) ben	ben**ce**

- When adding case suffixes to nouns, the phonetic harmony rules of Turkish are followed.

- Common nouns ending in a consonant receive case suffixes according to vowel harmony and consonant similarity.

- Additionally, softening rules are applied for consonants at the end of words.

The logical relationship regarding case suffixes is seen in Figure 8.
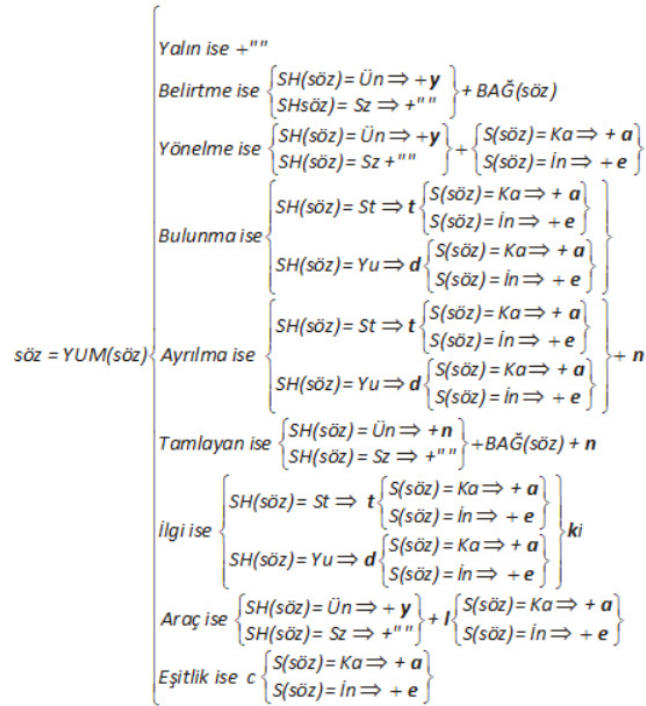
Figure 8 – Logical relationship of case suffixes (Yalın: subjective, Belirtme: accusative, Yönelme: dative, Bulunma: locative, Ayrılma: ablative, Tamlayan: genitive, ilgi: relative, Araç: instrumental, Eşitlik: equality, ise: if, St: Hard consonant, Yu: Soft consonant)

Verb Conjuguation. Verbs are examined in two sets: Auxiliary verbs and main verbs.

Auxiliary Verb Conjuguation. In Turkish, auxiliary verbs allow noun words to be used as predicates [10]. Suffixes are shaped according to the tense and are defined for four tenses and are drawn through the following stages:

1. Creating the negative of the root
2. Creating the suffix
3. Softening
4. Editing variable letters within a suffix
5. Create the verb

Tense and samples of auxiliary verbs are given in Table 3

Table 3 – Auxiliary Verbs

Tenses	Time suffixes	Personel suffixes	Examples
Past tense indefinite	+(y)mHş	+Hm, +sHn, +"", +Hz, +sHnHz, +lAr	öğrenci**mişim**
Past tense	+(y)DH	+m, +n, "+", +k, +nHz, +lA	öğrenci**dim**
Present and future	+""	+(y)Hm, +sHn, +DHr, +(y)Hz, +sHnHz, +lAr	öğrenci**yim**
Optative - condition	+(y)s(A)	+m, +n, +"", +k, +nHz, +lAr	öğrenci**ysem**

Tenses	Time suffixes	Personel suffixes	Examples
suffix +ken	+""	+(y)ken	öğrenci**yken**
suffix +mışçası	+""	+(y)mHşcAsHnA	öğrenci**mişçesine**

4. Results

Conjugation of Verbs. Conjugation of Turkish verbs is regular and there are no examples of exceptions to the rule. On the other hand, in English and French, the conjugations of verbs that are used frequently and number more than a hundred are irregular [14,15]. Conjugate of illegal actions can only be used by memorization. The fact that verb conjugations are regular in Turkish allows it to be conjugated in accordance with the rules when a new verb enters the language.

Turkish verbs have the feature of indicating **direction** and are **logical**:

Eren tells Bartu

Bartu is listening to Eren.

The inflection pattern of a verb in Turkish is as follows:
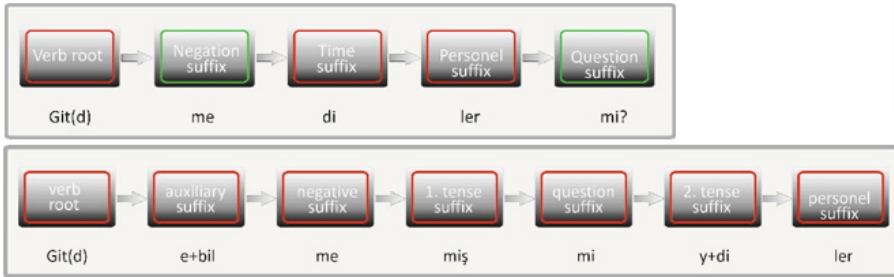
Figure 9 – The inflection pattern of a verb in Turkish

Tenses	Time suffix	Personel suffixes	Examples
Past tense, indefinite	+mHş	+Hm, +sHn, +"", +Hz, +sHnHz, +lAr	**sev**mişim
Past tense	+D	+Hm, +Hn, +H, +Hk, +Hniz, +HlAr	**sev**dim
Present simple tense	+(H)yor	+Hm, +sHn, +"", +Hz, +sHnHz, +lAr	**sev**iyorum
Future tense	+(y)AcAk	+Hm, +sHn, +"", +Hz, +sHnHz, +lAr	**sev**eceğim
Aorist	+(A/H)r *	+Hm, +sHn,+ "", +(y)Hz, +sHnHz, +lAr	**sev**erim
Past perfect tense indefinite	+mHşD	+Hm,+Hn,+H,+Hk,+HnHz,+HlAr	**sev**miştim
Past perfect tense	+DHyD	+Hm,+Hn,+H,+Hk,+HnHz,+HlAr	**sev**diydim
Present past tense	+(H)yorD	+Hm,+Hn,+H,+Hk,+HnHz,+HlAr	**sev**iyordum
Future past time	+(y)AcAkD	+Hm,+Hn,+H,+Hk,+HnHz,+HlAr	**sev**ecektim
Aorist perfect tense	+(A/H)rD *	+Hm,+Hn,+H,+Hk,+HnHz,+HlAr	**sev**erdim
Past indefinite Inferential	+mHşmHş	+Hm, +sHn, +"", +Hz, +sHnHz, +lAr	**sev**mişmişim
Present Inferential	+(H)yormHş	+Hm, +sHn, +"", +Hz, +sHnHz, +lAr	**sev**iyormuşum
Future Inferential	(y)AcAkmHş	+Hm, +sHn, +"", +Hz, +sHnHz, +lAr	**sev**ecekmişim
Aorist Inferential	(A/H)rmHş *	+Hm, +sHn, +"", +Hz, +sHnHz, +lAr	**sev**ermişim
Past ind. Inferential conditional	+mHşsA	+m, +n, +"", +k, +nHz, +lAr	**sev**mişsem
Past ind. Inferential conditional	+DHysA	+m, +n, +"", +k, +nHz, +lAr	**sev**diysem
Present conditional	+(H)yorsA	+m, +n, +"", +k, +nHz, +lAr	**sev**iyorsan
Future conditional	+(y)AcAksA	+m, +n, +"", +k, +nHz, +lAr	**sev**eceksem
Aorist conditional	+(A/H)rsA *	+m, +n, +"", +k, +nHz, +lAr	**sev**ersem

Figure 10 – Indicative Moodes

Moodes	Time suffix	Personel suffixes	Examples
Wish - Conditional	+sA	+m, +n, +"", +k, +nHz, +lAr	**sev**sem
Optative	+A	+yHm +sHn, +"", +lHm, +sHnHz, +lAr	**sev**eyim
Obligational	+mAlH	+yHm, +sHn, +"", +yHz, +sHnHz, +lAr	**sev**meliyim
Imperative mood	+""	+"", +"", +sHn, +"", +HnHz, +sHnlAr	**sev** (2.tek)
Perfect of wish - conditional	+sAyD	+Hm, +Hn, +H,+Hk, +HnHz, +HlAr	**sev**seydim
Perfect of optative	+AyD	+Hm, +Hn, +H, +Hk, +HnHz, +HlAr	**sev**eydim
Perfect of obligational	+mAlHyD	+Hm, +Hn,+H, +Hk, +HnHz, +HlAr	**sev**meliydim
Inferential of wish-conditional	+sAymHş	+Hm, +sHn, +"", +Hz,+sHnHz,+lAr	**sev**seymişim
Inferential of optative	+AymHş	+Hm, +sHn, +"", +Hz, +sHnHz, +lAr	**sev**eymişim
Inferential of obligation	+mAlHymHş	+Hm, +sHn, +"", +Hz, +sHnHz, +lAr	**sev**meliymişim
Conditional of obligation	+mAlHysA	+m,+n,+"",+k,+nHz,+lAr	**sev**meliysem

Figure 11 – Subjunctive Modals

Conjugating a verb is done through the following steps:
1. Preliminary preparation for suffixes
2. Establishing the verb framework
Creating the reflexive verb stem
Creating the corresponding verb root
Creating the causative verb stem
Creating the passive verb stem
3. Auxiliary verbs
haste helper verb

Qualification helper verb
Persistence helper verb
4. Creating the negative of the verb
5. Fix in verb root
6. Creating and combining tense and individual suffixes
7. Softening verb root
8. Fix variable letters in verb
9. Creating the question form of the verb
10. Creating the captured verb

When adding time and individual suffixes to a verb root, the last letter of the root verb is changed. The rule for this change is as follows:

- In the first stage, whether the last letter of the root is **a** or **e** is examined. If the last letter is **a**, it is changed to **ı** and if it is **e**, it is changed to **i**.

- In the second stage, if the last letter of the root is a consonant, the last vowel is checked. If the last vowel is a or ı, ı is added to the root, if e or i, i is added, if **o** or **u**, **o**, if **ö** or **ü**, **ü** is added. In a sense, the binding vowel rule applies.

- Among the Turkish verb roots, it was necessary to change the last letter of the four verb roots ending in **t** (*dit, et, git, güt, tat*).

When adding time and individual suffixes to a verb root, a softening process is applied to the suffixes. Indicative and optative moods of Turkish are shown in Figure 10 and Figure 11.

Table 4 – Words Derivation from Root Eye

Turkish	English	Class	Turkish	English	Class
Göz	Eye	Noun	Gözlemci	Observer	Adjective
Gözlük	Eyeglasses	Noun	Gözlemcilik	Observation	Noun
Gözlükçü	Optician	Adjective	Gözde	Favourite	Adjective
Gözlükçülük	Opticians	Noun	Gözgü	Mirror	Noun
Gözcü	Watchman	Adjective	Gözgülük	Mirror stand	Noun
Gözcülük	Opthamolgy	Noun	Gözgücü	Mirror maker	Adjective
Gözlem	Observation	Noun	Gözgücülük	Mirror makers business	Noun
Gözlemlemek	To observe	Verb	Gözlü	Who have eyes	Adjective

Derivational Suffixes. TDK sources show that the number of root verbs in Turkish is 758, the number of main verbs is approximately 6523, and the number of basic noun words is approximately 15,468. There are those who, with a cursory comment without evaluating the structure of Turkish, say that the vocabulary of Turkish is weak. Those who think this way say that many words in foreign languages do not have equivalents in Turkish. In Turkish, which is an agglutinative language, many new words can be derived from one root, and the first person who hears each new word can deduce its meaning. Table 4 shows the words derived from the root eye.

noun + suffix ⇒ noun   verb + suffix ⇒ verb
noun + suffix ⇒ verb   verb + suffix ⇒ noun
noun + suffix ⇒ adjective verb + suffix ⇒ adjective
noun + suffix ⇒ adverb   verb + suffix ⇒ adverb

A new word can be derived by adding one or more suffixes to a root word. 36 features were determined to derive new words. New words can be derived by choosing one of these features. Figure 12 shows 36 features and the subject to be derived.

The derivational suffix is added to the root by going through the following steps:

1. Preparation phase: Is the word Turkish?; The meaning to be attributed; Class of the root
2. Suffix formation phase: Sound harmony rules; fusion letter
3. Softening phase
4. Adding the suffix to the root

For example: Root: **arpa**; Suffix: **lik.**

First letter of the suffix: l, not variable. Instant addition: l

The second letter of the suffix is H: variable; The suffix is monosyllable, so the new value of the letter H is found as **ı** by using the following logical relation. Suffix becomes: lı.

$$Seslem\_sayısı(ek) = 1 \ ise \begin{cases} S(Kök)\,(a \vee \imath)\,ise\ H \rightarrow \imath \\ S(Kök)\,(e \vee i)\,ise\ H \rightarrow i \\ S(Kök)\,(o \vee u)\,ise\ H \rightarrow u \\ S(Kök)\,(ö \vee ü)\,ise\ H \rightarrow ü \end{cases}$$

(Seslem_sayısı: Number of syllable, kök: root)

Last letter of the s uffix: k, not variable. Instant addition: **lık.**

- Place name
  - Position regarding action
  - Passing place
  - Place, region suitable for action
  - Place name reminiscent of its qualification
  - Cute settlement
  - Relatively small settlement
  - Name given due to the feature it contains
  - Place name that reminds the living people
  - Place name related to the intended use
  - Place name suitable for the purpose of use
  - Specified location
- Hand tool name
  - Hand tool that contributes to the action
  - Hand tool that helps action
  - Hand tool used in action
  - Hand tool used to implement the action
  - Object related hand tool
- Tool name
  - Tool used in action
- Device name
  - Action related device
- Name of profession or career stage
  - Profession name
  - Career stage
  - Community of colleagues
  - The work done by the professional
  - Professional name related to the job he does
  - Profession name at the degree of specialization
- Adopting a movement of thought
  - Adopting an idea
  - Opinion based on thought
  - Those who think the same way
- Togetherness
  - Being together
  - Cluster of similar things
- Equality, partnership, affiliation
  - Carrying common features
  - A community of people with common character
  - Having common qualities
  - Coming from the same place, root or lineage
  - Resembling an object

- Similarity
  - Resemblance
  - Similar quality
  - Similar behavior
  - Reminiscent behavior
- Illness
  - Disease name
- Herb
  - Plant name
- Organ
  - Organ name
- Equal distribution
  - Delivery numbersı
- Sorting
  - Ordinal numbers
  - Do the action in order
  - Doing two actions together
- Time
  - A certain moment
  - In a certain period
  - Within a certain period of time
  - A certain time
  - At a certain moment
  - While in a certain situation
  - Always
  - When it happened
  - Time period
- Action that states not to do
  - The act of doing indicated by the noun
- The act of doing that is indicated by the verb
  - The act of doing indicated by the verb
  - The act of doing indicated by the reflexive noun
  - The act of doing that the adjective indicates
  - Verb make from verb
  - Action of making from reflexive verb
- Verb indicating being
  - The act of becoming what the noun indicates
- Person doing the action
  - The person or object doing the action
  - A person who does the action without caring
  - A person who performs an action without thinking
  - The person who performs the action

- Temperament, behavior, habitual quality
  - Maintaining temperament and behavior as a habit
  - Action-related outcome behavior
  - Action-related temperament, behavior
- Contempt, pity, love
  - Contempt
  - Disdain
  - Pity and love
  - Relatively weak
  - Humiliation
  - Relatively small
- Language and dialect
  - Language and dialect
- Ongoing
  - Noun that describes the result of an ongoing action
  - Known situation
- Prayer and wish
  - Prayer and wish
- Food name
  - The name of the food given by those who eat its name.
  - A dish named after its cooking method.
  - Direction
  - Direction
- Seven, Düşkün
  - Seven, yakınlık gösteren, düşkün
- Hosting a feature
  - Qualifying adjective related to a noun
  - Possessing the qualities indicated by the noun
  - The property of resembling something
  - Being related to the subject indicated by the noun
  - Qualifying adjective related to noun
  - Being related to the organ or object indicated by the noun
  - Approaching the quality indicated by the adjective
  - Qualifying the outcome of the action
  - Explaining the result of the action

- Resulting from action
  - The place, environment created as a result of the action
  - The tool created as a result of the action
  - Desire resulting from action
  - Passive subject of action
  - Document or opportunity created as a result of an action
  - Action-like function
  - Abstract concept related to action
  - Action related concept
  - Tool used in action
  - Result concept that evokes action
  - The resultant product resulting from the action
  - Technical terms that remind of action
  - Subject performing the action
  - The unit or part that emerges with the realization of the action
- The object that emerges when the action takes place
  - The relevant process for the action to be carried out
  - The situation that occurs when the action occurs
  - How to do the work related to the action
  - The act of doing action-related work
  - Place where the action is performed
  - The object resulting from the action
  - The structure, product formed as a result of the action
  - Reflexive event that confuses the action
  - Event that evokes action
  - Result related to action
  - Structure, subject that evokes action
  - Reflexive situation that confuses the action
- Having some qualities
  - Having some qualities
  - Noun that resembles the appearance of an object or person
- Structure
  - Structure

Figure 12 – Classification of derivational suffixes

The Syntactic Features of Turkish. The vocabulary in a language is limited. In addition, the sentence structure and rules of each language are clear. Based on these facts, it can be said that the sentences that can be written in a language can be derived artificially and their number will be limited. However, human ability shows that an infinite number of sentences can be produced in accordance with the rules of the language. Because those who produce sentences are individuals, and it is seen that individuals can produce different types of sentences. Each new sentence produced has a meaning and the listeners understand it. It is accepted that this ability arises from human creativity.

Linguists Wilhelm von Humboldt and Noam Chomsky also say that the number of sentences in a language is infinite. The human ability to produce new sentences and understand new sentences is called Language Acquisition Device (LAD) by N. Chomsky.

The EDB approach assumes that every individual innately knows the principles and variables that determine the structure of the language he speaks. Principles are abstract and do not change from language to language. Every language has a structure and rules, so words cannot be ordered randomly in a sentence constructed to express something. In other words, words are arranged in a certain order to create a sentence to express something. Because people know these rules, they can understand what is said.

It is thought that Chomsky did not sufficiently examine Turkish when proposing the EBD theory. Because in Turkish, especially the syntax is flexible. **In other words, the sentences are not in formulaic form. The ability of a flexibly constructed sentence not to lose its meaning is an ability that Turkish case suffixes have brought to Turkish.**

Arabayı sabunla yıka. (Araba: car, sabun: soap, yıka: wash)

Arabayı yıka sabunla.

Sabunla arabayı yıka.

Sabunla yıka arabayı.

Yıka arabayı sabunla.

Yıka sabunla arabayı.

While a bond is established between complement and complement words in Turkish, this bond is not established in English. Turkish morphological and syntactic features enable a close bond to be established between the words in the sentence. The bond established is clearly visible. The establishment of the bond and the elements that establish the bond

are clear and are added to the words in accordance with certain rules. **These features provide Turkish with the flexibility to change the place of words within the sentence, but without disrupting the meaning of the sentence.**

As a result of Turkish being regular and logical, a sentence can be derived by the computer. For this reason, the Turkish sentence;
- basic structure,
- subject characteristics
- sentence features and

- It is sufficient to know the properties of the predicate.

The sentence structure in Standard Turkish is as shown in Figure 13.
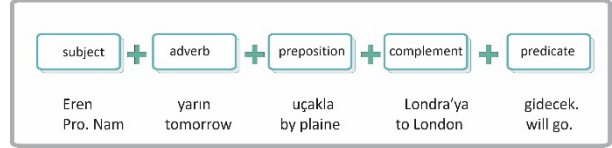

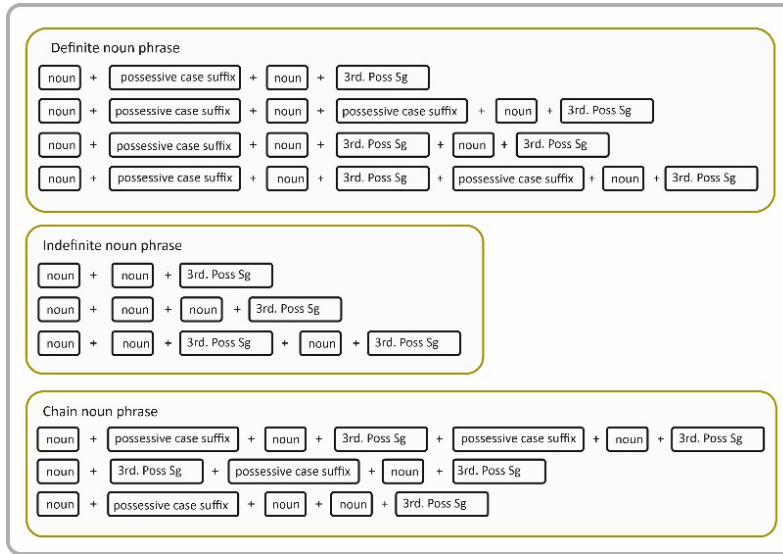**Figure 13** – Structure of Standard Turkish Sentence

Figure 14 – Noun Compount Subject

There are three types of noun phrases in Turkish (Figure 14):
- Definite noun phrase
- Indefinite noun phrase
- Chained noun phrase

The subject structure in the form of adjective phrase is given in Figure 15. As seen in Figure 15, the subject can be formed with four different adjective phrases. The complement structure is drawn in Figure 16.

Sentence generation software consists of the following authoring environment, editor and interpreter. This structure can be seen in Figure 17.
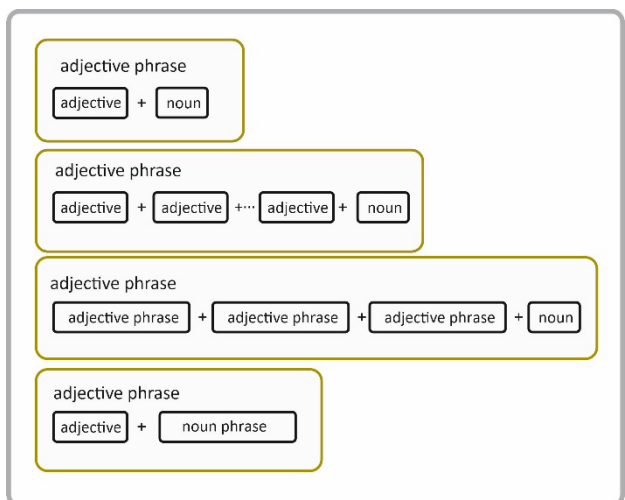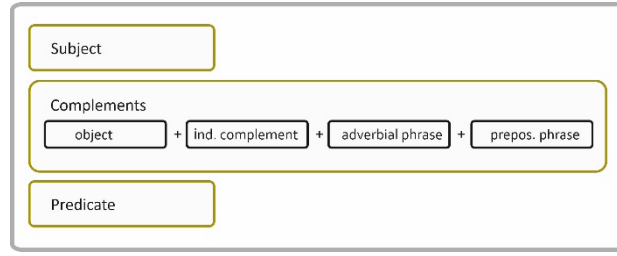

Figure 15 – Adjective Compount
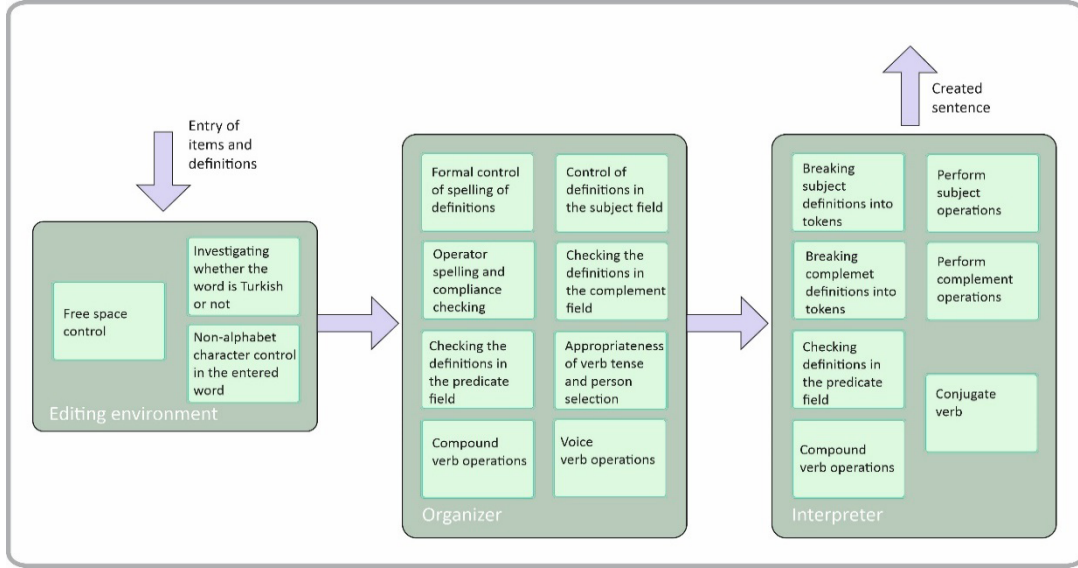
Figure 16 – Complement

Figure 17 – Structure of the sentence-generating program

Organizer. In this environment, the elements that should be included in the subject and complement are defined. In addition, it is possible to choose the framework and compound action operations related to the verb root and verb. Naturally, the time, person, positive and negative situations of the action are determined. In addition to ensuring that definitions and selections can be entered into the program, the editor also checks whether the entries comply with the spelling and logic rules. In this context, it carries out the following inspections.

- Whether or not the fields to be defined and selected have been entered,
- Whether or not there is any character other than the alphabet in the entered words,
- Whether the words entered are Turkish or not,
- Formal control of the spelling of definitions,
- Whether the operators are written correctly and used appropriately,
- Whether all definitions within the subject are made correctly or not,

- Whether all the definitions in the complement are made correctly or not,
- Whether the predicate is Turkish or not,
- Suitability of the options selected for voice operations,
- Suitability of options selected for combined verb,
- Whether the verb mode has been chosen and the appropriateness of the individual's choice for this mode

Interpreter. After entering the necessary information regarding sentence formation, the interpretation phase begins. At this stage, first the definitions entered in the subject and complement fields are analysed. Within the scope of the analysis process, the definitions entered in the relevant fields are divided into tokens.

A[Eren kedi ben] OY[sarışın kız] OY[güzel kedi] AY[okul yol] OT[Alpay güzel kız] OT[Alpay güzel kız -] ATD[kedi göz] OT[biz büyük ev -] ATD[geniş kapı] ZA[han kapı kol] ZB[han kapı kol]

5. Conclusion and Evaluation

The grammar rules of Turkish were created thousands of years ago and the rules have not been changed until today. There is nothing unusual in Turkish. The fact that the structure of the language is regular makes it possible to write the rules with logical relations. A flow chart can be drawn by using logic relations and then a program can be written. The following programs were written according to the rules of Turkish:

- Addressing

- Testing whether the word is Turkish or not

- Soften the last letter of the word and edit the suffix accordingly

- Conjugating a noun with all possibilities

- Shoot additional action

- Action shooting

- Adding a suffix to a root word according to the purpose

- Sentence creation

All root words of Turkish (nouns and verbs) have been tested on the written programs and the soundness of the rules of Turkish has been demonstrated.

References

1. https://arkeofili.com/bilinen-ilk-resim-yazisi-gobeklitepede-olabilir/
2. S. Eker, Türk Dili, Airforce Univ., 1999, Ankara
3. A.Güzel, Üniversiteler için Türk Dili (I-II) Course book, Başkent Univ., 2020
4. M. Müller, Lectures on The Science of Language, given to Royal Institution of Great Britain in 1861, [Ebook 32856], 17 July, 2010
5. E. Adalı, Türkçenin Mantığı, Akademik Bilişim Yayınları, 2024
6. J. Weizenbaum, ELIZA A Computer Program For the Study of Natural Language Communication Between Man And Machine MIT Cambridge, Communication of the ACM Volume 9 / Number 1 / January, 1966
7. X. Li. D. Roth, *Learning Question Classifiers*, COLING '02 Proceedings of the 19th international conferenceon Computational linguistics – Volume 1 Pages 1-7 , 2002
8. X. Li. D. Roth, *Learning Question Classifiers: The Role of Semantic Information*, Natural Language Engineering 1 (1): 000–000. Cambridge University Press. 2004
9. *Scoring, term weighting and the vector space model*, Cambridge University Press. 2009
10. *Probabilistic information retrieval*, Cambridge University Press. 2009
11. M. Sahlgren, *The Word-Space Model, Using distributional analysis to represent syntagmatic and paradigmatic relations between words in high-dimensional vector spaces*, Doktora Tezi, Stockholm University, 2006
12. A. Delibaş, *Doğal Dil İşleme ile Türkçe Yazım Hatalarının Denetlenmesi,* Yüksek Lisans Tezi, İTÜ Fen Bilimleri Ens. 2008
13. N. Coşkun, *Türkçe Tümcelerin Ögelerinin Bulunması,* Yüksek Lisans Tezi, İTÜ Fen Bilimleri Ens. 2013
14. Kemal Oflazer "Two-level Description of Turkish Morphology"
15. Hasim Sak, Tunga Gungor, Murat Saraclar "Morphological Disambiguation of Turkish Text with Perceptron Algorithm".

Information about author:

Eşref Adali – PhD, professor at Istanbul Technical University (Istanbul, Turkey, e-mail: adali@itu.edu.tr)