
Word-Forming Process in Azeri Turkish Language

Morvarid Jaza'ei¹, Mohammadreza Oroji^{2*}, Behzad Rahbar³, Sakine Jafari⁴

¹Ph.D. Student, Department of Linguistics, Islamic Azad University, Zanjan, Iran

^{2*}Assistant Professor, Department of Linguistics, Islamic Azad University, Zanjan, Iran

³Assistant Professor, Department of Linguistics, Allameh Tabataba'i University, Tehran, Iran

⁴Assistant Professor, Department of Linguistics, Islamic Azad University, Zanjan, Iran

Received: 25 February, 2020

Accepted: 29 November, 2020

Abstract

The subject intended to study the general methods of natural word-forming in Azeri Turkish language. This study aimed to reach this purpose by analyzing the construction of compound Azeri Turkish words. Same'ei (2016) did a comprehensive study on word-forming process in Farsi, which was the inspiration source of this study for Azeri Turkish language word-forming. Numerous scholars had done various studies about this issue in different languages; however, Same'ei's (2016) point of view and his application of variables like 'x' and 'y' were really attractive and considerable although absence of such study was quite visible to the Turkish language. Accordingly, this study considered 'x's and 'y's to represent the components of the compound words which were collected from reliable Turkish bilingual or monolingual dictionaries, ongoing dialogues, and grammatical sentences. Different patterns were introduced to show the structures of the recognized compound words. The compound words were assessed by the process, and also were contrasted considering their syntactical chain and have been analyzed by analogy of the words, as Same'ei (2016) did. Accordingly, the structural differences between the Azeri Turkish and Farsi languages were considered, and words structure were successfully analyzed with their components via the desired variables.

Keywords: Azeri Turkish, Compounding, Morphology, Word-formation

INTRODUCTION

Language is a special and social phenomenon that acts as a source of affection and a means of exchanging thoughts, intentions, and opinions among the people of the community. Language is the result of humans' collective life, and like other institutions and social forms is in transformation and is the basis of the human development in societies. Knowledge, technology, art, and economy as well as other

aspects of life grow up in the society and this growth and development directly reflects in the language.

In the Azeri Turkish language, word analysis and word formation process have always been a part of grammatical studies. European scholars such as Scalise and Vogel (1977) have done studies for compounding process and American scholars like Lewis (2000), Kornfilt (1997), McCarthy (1970), and Hieber (2007) have conducted extensive studies to clarify Turkish grammar and have addressed

*Corresponding Author's Email:
Mohammadreza.oroji@iauz.ac.ir

word-building issues in Turkish language; also some Turkish scholars like Erguvanli (2001), Kazimov (2010), and Ergin (2009) have published increasingly perfect works about Azeri Turkish language. Some works of Iranian scholars studying Azeri Turkish word formation process and Azeri Turkish grammar are Rahimi (2009), Hadi (1995), Shahmarasi (2008), and Qaramaleki (2006). When examining the mentioned studies, it is considerable that there is a big gap in the precise definitions of Azeri Turkish Morphology, and there are no reliable studies about Azeri Turkish words and compounds and their origins and constructions. Same'ei (2016) has conducted a comprehensive study on the processes of word processing in Farsi and has provided a comprehensive theoretical idea for word forming and compounding process in Farsi language; same study could be done on the Turkish language, and in particular the Turkish language of Azerbaijan, with the use of the 'x' and 'y' models, as Same'ei (2016) has put forward. The question posed as the research pursuit was:

Based on Same'ei's notion (2016), how can Turkish word-making process be analyzed?

The main aim of the present study was to introduce the framework of Turkish word forming process with regards to Same'ei (2016) framework in Farsi. Attempts were made to implement a similar framework for the Azeri Turkish language and the idea that Same'ei (2016) has provided for the Farsi language was the researchers' inspiration source for their analysis.

It seemed that the work would be applicable for interested students, academicians, teachers, translators, speakers, journalists, broadcasters, play writers, artists, movie makers, directors, and everybody who is interested in languages and their specification.

Compound words may form by simple merging of two elements, and the exact relation among the elements is highly variable. Jackendoff (2002) claims that they may be regarded as 'protolinguistic fossils', a structural

type that has survived from the earliest forms of human language and are characterized by syntactically less constrained relations between the elements. Dressler (1987) advises that languages may have compounding without affixation, but the reverse does not seem evident.

Pinker (1999) studies the issue and mentions that the lexical processing system contains a duality in its procedure, while regular forms may be computed; irregular ones may be stored in the mental lexicon.

The Turkish language is located in the 'post-agglutinative languages' class and is part of 'agglutinative' group. The Turkish language forms nouns and verbs, also new terms and conditions by attaching suffixes; even some grammatical words, may derive from switching the voices (vowels) in the stem of words. Alibekiroğlu (2017) has studied the mentioned voices by details.

There isn't any compiled monograph for word formation process in Azeri Turkish language, yet. However, there are plenty of studies about word forming processes which are relying on word appearance like studies conducted by Adali (2004), Aksan (1983), Banguoğlu (2000), Bilgin (2002), Ergin (2009), Erkman-Akerson (2008), Koç (1996), and Toklu (2003).

In these studies, the main emphasis is that the Turkish language structurally keeps the word root constant and creates new words by adding some constructional suffixes which have different tasks and apply different forms to these unchanging fixed roots. And it is always emphasized that the Turkish language is one of the most typical examples of joined (agglutinative) languages. The unquestionable fact is that suffixes have the first rating order in the word forming process. However, like many other languages, Turkish applies different ways to meet new entities and concepts and to encounter loan words.

The researchers observed that there are few studies about Turkish morphology and ways of making words. Some Iranian and foreign scholars have studied the phenomenon by details,

some of the foreigner scholars are Turks and some are Europeans or Americans. Tekin (1967) mentions that, there are two kinds of stems:

i. Simple (primary) stems, which cannot be further analyzed, e. g., *av* (hunt) and *kal* (stay);

ii. Derivative (secondary) stems, which are obtained from the simple stems either by suffixation or by composition, e. g., *avcı* (hunter) and *kalan* (who stays);

iii. He continues that "A peculiar feature of the morphological system is the sharp distinction between the nominal and verbal stems expressing closely related ideas." (p. 103)

iv. The identical nominal and verbal stems are extremely few, and according to his study they are restricted to *ac* (hungry) and *ac* (to feel hungry), *bədiz* (design, decoration) and *bədiz* (to design, to decorate), *qərı* (old, aged) and *qarı* (to grow old). He continues that, the derivational suffixes fall into two classes:

v. Suffixes which form new stems from the nominal stems,

vi. Suffixes which form new stems from the verbal stems (p. 104).

Menges (1994) divides morphological system into a nominal and a verbal section. He claims that there is a strict distinction between nominal and verbal stems in Turkic languages, so that the word formation suffixes are divided into two opposing groups:

i. one of suffixes are attached only to nominal stems,

ii. another of them are just attached to verbal stems,

iii. only there are rare suffixes which can serve for either category,

iv. a complete word is generally characterized by the sound harmony (p, 157).

Ergin (1994) claims that there are four ways to encounter a new entity and concept:

i. Taking a foreign word

ii. Making a word group

iii. Searching and finding a dead word within the language

iv. Making a new word by adding suffixes to the roots and stems (Ergin, 1994, pp. 219 – 221).

Ahanov (2013) and Karaağaç (2012) focus on word derivation process and introduce the following categorization:

i. Insertion method: Adding a suffix to the root or stem and related tasks (deriving a new meaningful word),

ii. Bending method: There are certain patterns in the language and new words are outputs of these patterns,

iii. Repetition method: It doesn't have frequent use, but there are words which are made in this way, same as: *kara kara* (repetition of 'black'), *taza taza* (repetition of 'fresh'), *ağaç mağaç* (repetition of 'tree'),

iv. Combination method: Deriving a word by merging morphemes, *karagöz* (black eyed),

v. Auxiliary words method: This type of word may establish by uniting two different words from different grammatical levels, but one of these two word units becomes the connecting unit and would expand the meaning domain, so the produced word can name a bigger area, such as: *yaza-bilme* (being able to write), *tanıştırmak* (introducing),

vi. Word order method: Some syntactic occurrences may happen during the use of language. Some people who don't speak Turkish use this method by obeying the syntactic structure, like *taş adam* (stone man), *kapı adam* (strong man),

vii. Emphasis method: The method in which the accent is used as a connected unit of production and attraction. Turkish emphasis doesn't use meaning as a distinctive mechanism, for example: the Russian *зѣмѣк* (chateau), *замѣк* (lock) (Ahanov, 2013,

pp.300-313) and (karaağaç, 2012, pp. 251 – 253).

Word derivation method is also a widely used method with an expanded list of abbreviated words which are mostly English (USA) or Farsi (NAJA) or sometimes Turkish in essence like (GÜNAZ for Güney Azərbaycan), and (GAMOH for Güney Azərbaycan Milli Uyanış Hərəkatı); shortenings is a method of saving in writing and pronouncing and should not be considered as a way of making new words.

Looking at the examples offered by Eker (2013) it can be seen that the suffixes in the words *üste* (upward), *girdi* (entered), and *nur-dan* (from light) have been transformed into the language by the method of sampling, there is also the effect of exemplification, too (Karaağaç 2012, p. 114). Compilation and screening are practices of bringing words into the standard written language in order to reuse them. It does not form new words.

Apart from these three ways, it is seen that new meaningful words may be derived and produced by changing the vowels and consonant voices in the bases of Turkish words (roots and stems). This method used by Turkish to form words is called 'the method of twisting in word derivation' by Ahanov (2013) and Vardar (Vardar, 2002); the method is going to be introduced because the Turkish use this method to yield a new word, so comparing and discussing the method would be useful. It is the most distinctive feature of the language, to produce a new meaningful word or to determine the task of the word by changing the voice (Ahanov, 2013, p. 302) and the names of the gender (masculine, female, neutral), the number, the person, the time, the mode, etc. in the act of forming the language more meaningful by changing the sound structure in the word bases (Vardar, 2002, p. 52). For example, there are certain patterns in the Arabic language, and in these patterns new meaningful words are derived by changing the sound (more famous voices):

Root letters forming the verb 'to write' are) ك ت ب (k-t-b), and some of the derived

words are /كُتِبَ/ 'book', /كاتب/ 'a man writer', /كتبت/ 'a woman writer'.

In English, there are words such as 'man' and 'men'; 'mouse' and 'mice'; 'drink' and 'drank'; 'these' and 'those'; as can be seen in the examples, the words are either assigned to tasks by the method of changing the sound or new meaningful words are derived from the words.

There are a number of words that are not underestimated by the 'method of changing the voice' which is a twist-like method in Turkish language. This method has been widely discussed by Kara (2011) in his work titled 'Separate Fallen Words'. This study is a monograph that deals with the emergence of new meaningful dictionary units in Turkish, or foreign words with the method of changing the sound.

In Hacıeminoğlu (1991) work, noun may deliver some information about the verb and the root. In the words noun- *ayrılmak* (to leave), noun+ak; *ayak* (foot), the noun- points out the root, so a- is the verb (p.18). Let's consider the fact that the root of the verb is a- and the first derivation would be noun-. Examples of noun- are *ayır-* (separate), *ayrıl* (get separated) or other example can be *at-* (throw it), *az-* (depart from the road and the direct departure) and *aç-* (open) (away from a fixed point), which express the different aspects of the meanings they carry and has been transformed into the informative words by the method of voice exchange.

Hadi (1995) in his study 'Turkish is an Art' believes that "Turkish suffixes are fixed in number and as a master key can attach to every root, and the suffixes follow the vowels of the root" (p. 49). So, he believes if we get familiar with suffixes and know the grammatical structures, we would be able to differentiate Turkish words and even create new words. This study has divided stems into two categories of 'nominal stems' and 'verbal stems'; and has divided suffixes into two categories of 'nominal suffixes' and 'verbal suffixes'. It should be mentioned that verbal suffixes are suffixes which attach to verbs in order to inflect them, considering tense, number and

person; nominal suffixes attach to roots and make new words (adjectives and nouns).

Oflazer and Sarachlar (2018b) present a short survey of some important functions of Turkish, which are challenging and amazing for speech processing and natural evolving of the language. They introduce Turkish morphology as a complex one and believe its fusion with syntax is even more complex, that is because of agglutinating morphology, vowel harmony, and free constituent order in syntax. They add words form by a kind of really productive affixation of several suffixes to the root words in lexicon of about 30,000, except proper nouns.

The Turkish lexicon is extensively under the effect of Arabic, Farsi, Greek, Armenian, Russian, French, Italian, German, and recently English languages. Turkish words can take several inflectional and derivational suffixes; they can form a word which has translation of almost a sentence in English:

yaz + *a* + *bil* + *acək* + *sə* + *k* = if we will be able to write it

All morphemes have allomorphs, which is different considering the containing vowels in the root word:

sən + *dən* (from you) vs. *on* + *dan* (from her / him)

Multiple derivations in a given word are not an unusual occurrence. The researchers introduce the word *aparabildiklərimizdəndi* (It was one of things we could take with ourselves), which has a root and nine suffixes.

Oflazer (2017) claims that "the morphological analysis of a word can be represented as a sequence of tags corresponding to the overt (or covert) morphemes." He mentions the word *uzaklaştırılacak* (one, who will be sent away), and introduces five tags which have added to the root *uzak* (far) which are:

- (1) + Adj,
- (2) + Verb + Become,
- (3) + Verb + Cause,
- (4) + Verb + Pass + Pos,
- (5) + Adj + FutPart + Pnon.

The first tag indicates that the root is an adjective; the second tag indicates a derivation into a verb whose semantics is 'to become'; the third tag indicates that a causative verb has been derived; the fourth tag indicates the derivation of a passive verb; the fifth tag indicates a derivation into future participle.

Oflazer and Yeniterzi and Tür (2017) claim Morphological disambiguation is the task of determining the contextually correct morphological parses of tokens in a sentence. A morphological disambiguator takes in a set of morphological parses for each token, generated by a morphological analyzer, and then selects a morphological parse for each, considering statistical and / or linguistic contextual information. This task can be seen as a generalization of the part-of-speech tagging problem, for morphologically rich languages.

Arısoy and Saraçlar (2017) claim that the morphological analysis of Turkish language in word segmentation both at the lexical and surface forms can be applied as sub-lexical language modeling units. They believe that "discriminative language models, which outperform generative models for various tasks, allow for easy integration of morphological and syntactic features into language modeling" (p. 66).

So the agglutinative character of Turkish may cause high 'out-of-vocabulary' rates which can be decreased by applying sub-lexical units for finding special features of the language.

Yeniterzi (2017) and Tür and Oflazer (2017) persist in the 'named-entity recognition'; they believe this issue is an important subject for different natural language processing and issues like, data attraction, responding to questions, machine translation and so on. They claim that during the last decades, Turkish named-entity recognition has absorbed linguists' attention in terms of system development and resource development. They have offered a brief definition for named-entity recognition, but shortly it can be said that it is a process of identifying and categorizing the named-entities, like individual, place, material, and organization names, or date, time, curren-

cy, and percentage expressions in unplanned texts. They claim that "this is an important initial stage for several natural language processing tasks including information extraction, question answering, and sentiment analysis".

METHODS

Same'ei (2016) has conducted study about word processing issue, in Farsi, and has provided a theoretical idea. The same model was conducted on the Turkish language of Azerbaijan, with the use of the 'x's and the 'y's of the model.

The subject of this work is "describing the natural processes of word formation in Farsi language." This work would reach its goal through the analysis of data and would gain processes and rules governing the fabrication of Farsi words. Some phrasal, syntactic, and semantic issues of word-formation in Farsi would be discussed in this regard.

This review specifically addresses the major processes of word-making in a variety of verbal names, adjectives, and adverbs. Syntactic stereotypes (terms and expressions), prefixes and combinations of verbs, and compounding terms are not included in this analysis.

The main difference between the two groups of words is that the rules governing the construction of nouns, adjectives, and adverbs are different from the rules governing syntactic constructs, so that it is possible to talk about constructs and rules which are different from syntactic constructs. And the mentioned different nature leads us to devote an independent part to the word-making, apart from the syntax.

Furthermore, the rules governing the construction of terms, prefixes and conjunctions, as well as compound letters, are syntactical, and therefore is different from the mentioned rules. For example, the relationship between *færa* (beyond) and *ræftæn* (going) in *færa ræftæn* (going through), or *hærf* (conversation) *zædæn* (do) in *hærf zædæn* (to make conversation) is a relationship that is established between the dependent adverbial with the verb and the dependent object with the verb. So that

these affiliates; they can also be expanded, like any other independent syntactic units: *færatær ræftæn* (going more through), *hærfi zædæn* (to utter a conversation). The same rule is applicable to compound letters: the preposition of *be sooye* (to) may be used as *be* in *sooy* (to this way) or *be an sooy* (to that way) (Same'ei, 2016, pp, 3-4).

Then, Same'ei offered some brief definitions for number of terms which are necessary to define in his point of view, those definitions are as follows:

Word In linguistics, the concept of the 'word' or 'vocabulary' is often omitted because of its ambiguity (after (Chomsky, 1970) article, the issue has been changed). Some linguistic theories have not taken any independent levels into the word. In some languages, there are phenomena which are regarded as a word in those languages; however, those phenomenon may do not adapt to the definitions provided by linguists, or their equivalent in another language may correspond to the chain which is larger or smaller in that language; for example, in the Turkish language, the chain of four morphemes *ev-lər-in-də* is a 'word', but the Farsi equivalent of *dær khane shan* is two words, and the English equivalent of 'in their house' is three words.

Being Natural The patterns of constructing words can be considered from two perspectives: 'natural patterns' and 'fabricated patterns.'

In natural patterns we are faced with the mechanisms of word forming in the natural language function. In this context, one can see how the language has responded to her vocabulary needs - language responses to these needs are mainly done in three ways:

- i- Borrowing lexicons;
- ii- Creating grammatical and semantic changes in existing words;
- iii- Making new words.

All of these methods occur naturally in the language, but the third one is based on certain patterns that form part of the grammatical structure of the language.

Description Our method here is 'descriptive' and not 'prescriptive', we do not say how the word should be made, but we just want to show what the term is.

Phonetic System The phonetic system or phonemic section is a part of the language system in which the patterns and rules govern the relationship between the phonemes to form the phonetic chains and produce the phonological linguistic signs, and their perception and analysis, on the other hand.

Semantic System The semantic system, or semantic part, is the part of the language that defines and explains the general rules governing the semantic interpretation of linguistic messages.

Grammatical System The grammar system is the central part of the language. In this section, the smallest units with phonetic faces and semantic content, namely linguistic signs, are identified, and the combination of these units leads to the formation of larger units - words, groups, and sentences. Subcategories of the grammatical system are 'vocabulary' and 'syntax'. Among these two, there are two sub-sections of 'word-formation' and 'inflection' which are classified as the title of 'morphology' by the whole syntactic tradition as well as the new linguistics.

Vocabulary Each of the language tokens stored in the lexicon is a lexical unit. But how can the language signs stored in lexicon or lexical units may be recognized? Which of these three terms can be called a lexical unit: *mirævæm* (I am going), *rævænde* (goer), *ræftæn* (to go)? The word *mirævæm* is not a fixed or single form and varies according to the needs of the linguistic message. It can be

said that this word and other similar ones are various forms of a lexical unit. The word *rævænde* is not a lexical unit, because it does not have distinct syntactic and semantic features and cannot easily be placed on one of the semantic lexical networks.

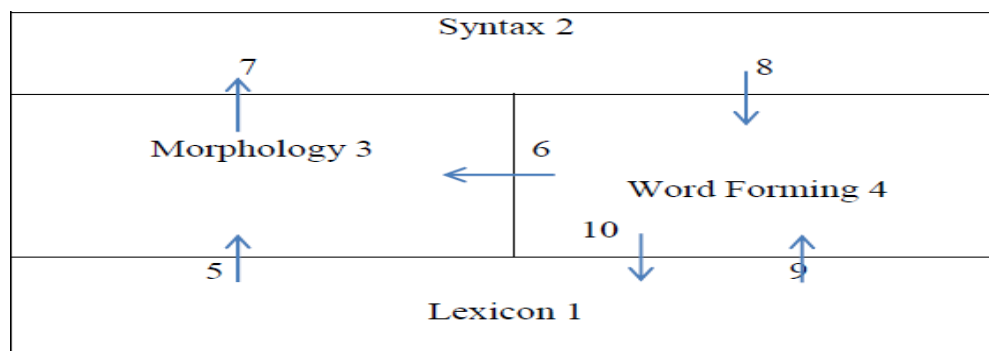
Syntax The syntax is the function field of some finite and limited rules, which refers to vocabulary units presence and sequence in the synchronic axis to form a language message.

Morphology Morphology is a part of the grammar system and through its limited and finite rules, lexical units may transform into forms and sentences that are applicable to the construction of syntactic chains. In this section, for example, the infinitive of *ræftæn* may change to the forms of *mirævæm*, *ræfte bashæm* (If I have gone), *miræftæm* (I was going) and so on; so that it becomes a message that the speaker intends to produce in proportion to the formal and semantic construction.

Word Forming Word forming is part of the grammatical system that produces new words with the help of a set of existing rules in the language, without the words necessarily being fixed in the lexicon and becoming a lexical unit.

Thus, word processing is a subset of the grammatical system, which is the domain of the two categories of rules: the rules of word-forming and the rules of lexicalization.

The grammars form The four parts of the grammar and the general relationship between them can be seen in the diagram below:



The diagram is derived from Same'ei (2016, p, 3) The diagram illustrates:

- 1-The lexicon is a fixed language form, with an open list, each of which is a lexical unit.
- 2-Syntax is the combination of some limited rules for creating syntactic chains between lexical units.
- 3-Morphology is the transformation of lexical units into possible terms.
- 4-The word-forming section refers to the ways of constructing new words.
- 5- 6-Input of the lexical units and the newly created words in the word-making section, which should be converted to applicable terminology to construct new sentences.
- 7-The output section of the spells to enter the syntax section or syntactic chains.
- 8- 9-The word-forming section uses words to create new words or syntactic chains, or uses lexical units as stems for vocabularies.
- 10-Prepared words in the word-forming section, directly enter the sentence making process, or are stored in the lexicon (Same'ei, 2016, pp. 1-33).

RESULTS AND DISCUSSION

The Turkish language is one of the languages of the Ural-Altai language groups and is an 'Agglutinative' language. Mahdixan (1960) argues that, in the agglutinative languages, the root of the verb remains fixed, and pronouns and suffixes are attached to it. Thus, root of the verb can always be distinguished in its various spatial states. In these languages, linguistic units are also effective in meaning and construction, but their construction is very clear. The components of the structure are also fixed. In these languages, in addition to the root of the words, which are composed of one or more syllables, there are also links between the syllables. In these languages, the main words always contain a single entity, and if the second-

ary meaning is added, the original form of that word does not change (pp, 7-15).

Shahmarasi (2008) claims, in his bilingual dictionary's introduction part, Azerbaijani people are ethnically dependent on Oghuz Turks. The Western Turks are the inhabitants of the eastern Caspian Sea, Iran, Turkey, Azerbaijan, Iraq, and the Balkans. Turkic tribes would speak 29 accents, 22 of which were accentuated with grammar and independent literature (p, 10).

In general terms, if a variable element (x) and another variable element or a fixed element along with (a) has been formed with two possible sequences, then there would be three templates or the process of word-formation in the formulation of the rules; each of these rules brings together some specific rules. If we set the constant element by (a) and the second variable element next to the variable (x) by (y) or (x'), then trivial processes can be formulated in the following forms.

One constant element and one variable element

In this process, one constant element is associated with one variable element (stem).

[a , x]

Based on the possible sequences of the two elements, as well as the constant element's duplicating ability, this process finds three different modes.

Stem and the preceding element: [a –x]

In this process, the constant element is placed before the variable. The fixed element, however, includes units that are called 'prefix' in the linguistic tradition, as well as other units that are not prefix, but their presence in a given texture, may cause forming a certain predictable lexical result. These results are related to the syntactic characteristics of the rules of this process, as well as their semantic characteristics.

It is necessary to mention that, there is no fixed element with the characteristics of linguistic 'prefix' in Turkish language; so the fol-

lowings are examples of the [a -x] process, but the fixed elements (a)s are not prefixes.

Od günü (eve day), Odlac (torch)	[od -x]
Barmaqca (little toe), Barmaqıq (toes's cover)	[barmaq -x]
Canlı (alive), Cansız (soulless body)	[can -x]
Qol qanad (arm and wing), Qolçaq (doll)	[qol -x]

Stem and the following element

In this process, the constant element is placed after the variable. It also includes what is

traditionally called 'suffix', and what traditionally is not called a 'suffix'. The rules below are typically [x- a] rules:

Kökdaş (two with same origin), Qardaş (sibling)	[x -daş]
Göbələklık (mushroom land), Güvənlik (security)	[x -lik]
Yerdən (from land), Göydən (from sky)	[x -dən]
Yağlı (oily), ballı (with honey)	[x -lı]

Stem and double fixed elements

In this state, one fixed element with two parts surrounds the variable element. The linguistic tradition calls the previous element as a prefix

and the following one as a suffix, but has no explanation for their simultaneous occurrence. The following forms are among the [a1 -x - a2] rule:

Göz açmak (opening eyes), Göz atmak (having a look)	[göz -x -C _i]
Əkin qarqası (kind of crow), Əkin lələsı (kind of tulip)	[Əkin -x -C _i]

the simultaneous creation of new words. In

Same'ei (2016) claims that The triple modes of [x , a] process, in terms of form, are completely similar to a process called 'derivation', in the existing traditions; but in terms of the application and inclusion of the elements in it, it is much broader than derivation. The basis of the process of derivation is on the 'affixation' theory that can be considered as a bound lexical morphology, which creates semantic and procedural changes while accompanying a stem or the basis of the compound (p. 44).

The 'affixation' theory is a limited theory, which is based on the historical rules of the language, rather than the existing mechanisms of the of word-making rules in the language; it focuses more on justifying the structure of existing vocabularies than on the mechanisms of

most optimistic terms, the vocabulary is a collection of units that have been built and compiled in the past and stored in the lexicon; the lexicon is storage of diachronic forms. Therefore, 'affixes' are representations of history of the languages; they are not representations of present language. "And since the history of language is more diverse than language in present, it is difficult to achieve a coherent role and meaning for the function of an affix, within the language" (Same'ei, 2016).

The theory of 'constant element' focuses on the formal and semantic relation between lexical units: which formal similarities, with relatively similar semantic consequences, can lead to the creation of a model that not only involves the development of new lexical units, but also may release a proper analysis for the meaning of the new modified words. The theory of constant element not only concerns on the formal / semantic constructions of the lexical units, but also the development of the rules for newly constructed word.

Two heterogeneous variable elements

In this process, two variable elements, or two different stems, are with together.

[x , y]

However, none of the two elements are fixed and the rule is to establish a relationship between the two. In a process with a constant element, that is [x, a], the central role is for the constant element, which is 'a', and the rule is developing the grammatical properties and the semantic role of this element in the whole construction. In *qab dəsmal* (kitchen towel), *dəsmal* – as the fixed element – persuades its adjective characteristic and semantic property to the whole word. But in [x , y], there is no such element, the relationship between the two elements has the duty. The relationship is a type of companion relation, which is considerable in syntactic rules. The relation between *rasta* and *bazar* in *rasta bazar* (straight bazar) is exactly like the relations observed in the words of *məhəllə bazarı* (local bazar) or *cumə bazarı* (Friday bazar).

Therefore, one can ask whether these constructs are not, in essence, syntactical constructs that have come up with words and have been frozen in the lexicon. The proof of such a process is possible, but first the historical stages of these units should be traced; it should get clarified if there were periods on which these syntactic units were in the form of syntactic chains and were used frequently. But finding such initial cases are difficult and impossible in many cases. However Turkish has *rasta bazar* and *şeytana bazar* (evil's bazar) (it is supposed that they are diverse forms of *rasta bazarı* and *şeytana bazarı*), is it necessary for the word *uzun bazar* (long bazar) to pass the same way and stage of being *uzun bazarı*? Does *baş barmağı* precede the common word

of *baş barmak* (the big finger)? If so, what about *sarı gəlin* (the bride of light)? It seems that there are rules in these constructions which are about words' structure not about phrases' structure (but not all of them). The present hypothesis is that, these rules themselves have roots in syntax, not their products. It means that it is possible to form a rule in which two nouns which can make a phrase are free to build a phrasal word without any overt symbol like 'sı'; the output for such a rule on a paradigmatic axis can be *canbir* (two with a same soul) or *quşbala* (the little bird). And along with this rule, there are several terms that, although seemingly look like outputs for the mentioned rule, but in fact are products of syntactic compounding process. The important point is that, in synchronic terms, the separation of products from these two processes is virtually impossible.

Two identical variable elements

In this process, two variables lie alongside each other, but unlike the process of [x , y], their relationship is not justifiable in the syntagmatic axis, but they have a paradigmatic relationship, that is, they belong to a same grammatical or semantic class.

[x , x']

The variable pairs in the constructions of the words *dəvə qüşü* (ostrich), *ölüm qalım* (dying or living), *basabas* (from beginning to end), *burun buruna* (nose to nose), *dağ daş* (mount and stone) are both belong to the same grammatical level or semantically they belong to similar area, while in the words *sarı gəlin*, *baş barmak*, or *rasta bazar* the containing words are chosen from different grammatical or semantic classes.

Formally, in some instances, there is a linking element between the two variables. The linking element can be considered as part of the process, which is sometimes eliminated through phonological or phrasal processes.

<i>Azdan çoxdan</i> (a little or a lot)	/az – dan – çox – dan/
<i>Alım satım</i> (buying and selling)	/al – ım – sat – ım/
<i>Qız qıza</i> (just girls)	/qız – qız – a/
<i>Baş başa</i> (just two people)	/baş – baş – a/

There is a form which belongs to this process

In this process the elements are similar; a variable repeats in compound [x , x'], like the words: *yuhayu* (washing and washing), *vürhavür* (hitting and hitting), *gəlhagəl* (coming and coming).

There is another form of the process [x , x'], too

In which the variable repeats incompletely or diversely; in other words, a successive variable falsifies in the compound which is similar to the main variable and places next to it. This fake variable is kind of duplex, which lacks independent meaning and application. This type of construction, which is very similar to reduplication, is a common phenomenon in the informal Turkish language. The difference between this process and the reduplication is that in the latter process the part of the base form repeats to create a new word, but in the duplexing process, in fact, the phoneme of the word repeats partially to intensify the meaning of the original word, but it does not necessarily has a new meaning. We observe this situation in complete reduplication, as well. The following terms are some examples:

Taraq turuq (hammer's voice), *sağ salım* (safe and sound), *qıp qırmıze* (really red).

It is one of the complexities of the phonetic rules in the Turkish language that the imitation of the phonemic face of the original word adheres to the phonetic rules. Kalbasi (1992) has tried to categorize this outcome in the Persian language, but it is very difficult to achieve an acceptable generalizations and predictable rules.

Compounding Processes of [x , y] and [x , x']

The processes of [x , y] and [x , x'] are traditionally called compounding processes and is in op-

posite side of the derivation process. In tradition, in the compounding process to stems mix up and there is no place for affixes, and the relationship between two stems is not mentioned. Recent theories consider both relations as a category of syntactic processes. However as mentioned about the [x , y], perhaps the [x , x'] process has origins in syntax but these rules have no relation with syntactical issues, except in some formal considerations. How can we explain ranges such as *yuhayu*, *qıp qırmıze*, and *bazak duzak* (repeat of the word, make up) through syntactic processes? Of course distinguishing 'compound' words with frozen syntactic forms is not as easy as distinguishing [x , y], because separate stems are present in their structure. If the words *dəvə qüşü* and *qıp qırmıze* are frozen nominal group or they are outputs of the [x , x'] rule is never provable.

Process' relationships

Based on the mentioned uncertain tokens, both in [x , y] and [x , x'], can be considered on a continuum in which one of the poles is entirely a constructive word, such as derivations and extensions, and the other pole consists of a syntactic processes such as nominal groups. In fact, there are several boundary divisions to justify this continuum. In this case, *yer göbələğe* / *yer göbələk* is closer to be a syntactic result and *yerəlama* is close to be a constructional word. These two chains are both words, based on the word recognition criteria, but the presence of occasional grammatical symbols in words such as *yer göbələğe* close them to the syntax, while word *yerəlama* lacks such features.

This continuity is considerable in [x , y] and [x , x'], too. Many of the units in this model are

References

- Adali, O. (2004). *Türkiye türkçesinde biçimbirimler*. İstanbul: Papatya Yayıncılık.
- Ahanov, K. (2013). *Dil bilimin esasları*. Ankara: TDK.
- Aksan, D. (1983). *Sözcük türleri*. Ankara: TDK.
- Alibekiroğlu, S. (2017). Türkçede Sözlük Türetme Yollar. *Dergi Park*, 35-46.
- Arısoy, E. S., M. . (2017). Turkish Speech Recognition. *Theory and Applications of Natural Language Processing*, 93-114.
- Banguoğlu, T. (2000). *Türkçenin grameri*. Ankara: TDK.
- Bilgin, M. (2002). *Anlamadan anlatıma türkçemiz*. Ankara: Kültür Bakanlığı.
- Chomsky, N. (1970). Remarks on Nominalization. In R. A. Jacobs & P. S. Rosenbaum (Eds.), *Readings in English Transformational Grammar* (pp. 184-221). Boston: Ginn.
- Dressler, U. W. (1987). 'Morphological Islands: Constraint or Preference?'. *R. Steele and T. Treadgold, Language Topics*.
- Eker, S. (2013). *Çağdaş Türk dili*. Ankara: Grafiker Yayın.
- Ergin, M. (1994). *Türk dili*.
- Ergin, M. (2009). *Türk dil bilgisi*. İstanbul: Bayrak.
- Erguvanli, T. (2001). *The verb in Turkish*. Amsterdam: John Benjamins.
- Erkman-Akerson, F. (2008). *Türkçe örneklerle dile genel bir bakış*. İstanbul: Multilingual.
- Hacieminoğlu, N. (1991). *Türk dilinde yapı bakımından fiiller*. Ankara: Kültür Bakanlığı Yayınlar.
- Hadi, I. (1995). *Turkish is an art*. Tabriz: Ahrar.
- Hieber, D. W. (2007). *A grammar of Turkish*. New York: College of William.
- Jackendoff, R. (2002). *Foundations of knowledge*. Oxford: Oxford University Press.
- Kara, M. (2011). *Ayrı düşmüş kelimeler*. Ankara: Bayrak Matbaası.
- Karaağaç, G. (2012). *Türkçenin dil bilgisi*. Ankara: Akçağ.
- Kazimov, Q. Ş. (2010). *Müasir Azerbaycan dili*. Bakı: Elm ve Tahsil.
- Koç. (1996). *Türkçe verintiler sözlüğü*. Ankara: TDK.
- Kornfilt, J. (1997). *Turkish*. New York: Routledge.
- Lewis, G. (2000). *Turkish grammar*. Oxford: University Press.
- Mahdixan, M. (1960). *A Persian Guide to the Turkish Language*. London: E. J. W. Gibb Memorial.
- McCarthy, K. M. (1970). *The linguistic adaptation of loanwords in modern standard Turkish*. Michigan: Xerox.
- Oflazer, K. (2017). Morphological Processing for Turkish. *Theory and Applications of Natural Language Processing*, 21-52.
- Oflazer, K. S., M. . (2018b). Turkish and Its Challenges for Language and Speech Processing. *Theory and Application of Natural Language Processing*, 1-19.
- Pinker, S. (1999). *Words and rules - the ingredients of language*. New York: Basic Books.
- Qaramaleki, H. (2006). *Analyzing the Turkish language in a mathematical method*. Tehran: Andisheye Arya.
- Rahimi, F. (2009). *Türkçenin mucizesi*. Tehran: Andisheye No.
- Same'ei, H. (2016). *Word forming in Farsi language*: Bahar.
- Scalise, S. V., I. . (1977). Why Compounding? (pp. 1-18). University of Bologna and University of Delaware: University of Bologna and University of Delaware.
- Shahmarasi, P. (2008). *Turkish - Farsi dictionary*. Tabriz: Akhtar.
- Toklu, O. (2003). *Dilbilime giriş*. Ankara: Ançağ.
- Vardar, B. (2002). *Açklamalı Dilbilim Terimleri Sözlüğü*. İstanbul: Multilingual.

Yeniterzi, R. T., G. Oflazer, K. . (2017). Turkish Named-Entity Recognition. *Theory and Application of Natural Language Processing*, 115-132.

Biodata

Ms Morvarid Jaza'ei is a Ph.D. graduate of Linguistics from Islamic Azad University, Zanzan Branch, in 2019. This article has been accomplished standing on her PhD thesis titled "Word-Forming and Compounding Process in Azeri Turkish Language Based on the Same'ei's Principle", supervised by Dr. Mohammadreza Oroji. Her field of interest is morphology.

Email: Morvaridj@live.com

Dr Mohammadreza Oroji holds Ph.D. in General Linguistics from Allameh Tabataba'i University, Tehran, Iran; he currently is an assistant professor in Azad University, Zanzan Branch.

His fields of interests are syntax, phonology, and morphology.

Email: Mohamadreza.oroji@iauz.ac.ir

Dr Behzad Rahbar holds Ph.D. in General Linguistics and he currently is an assistant professor at Allameh Tabataba'i University, Tehran, Iran. His fields of interests are sociolinguistics, morphology, and semiotics.

Email: Behzadrahbar@atu.ac.ir

Dr Sakine Jafari holds PhD in TEFL and is an assistant professor in the department of Linguistics at Islamic Azad University, Zanzan, Iran. Her research interests are second language acquisition, language teaching, and learner autonomy.

Email: Sjafari591@yahoo.com