

## GRAMMATICALIZATION OF “BAŞLAMAK”

Nadzeya Dardykava

Post-graduate student, Minsk State Linguistic University, Belarus, n.dardykova@gmail.com

### Abstract

The present paper is concerned with the grammaticalization of the Turkish aspectual verb *başlamak*. The paper analyses the level of autonomy of *başlamak* with Ch. Lehmann's parameters of grammaticalization. These parameters are based on three criteria, i.e. weight, cohesion and variability, which act on syntagmatic and paradigmatic axes. The weight of a sign viewed paradigmatically is its integrity or semantic and phonological complexity. Viewed syntagmatically, it is its structural scope, i.e. the extent of constructions it helps to form. For *başlamak* no signs of phonological attrition (gradual loss of phonological substance) were found. Semantically it has inceptive meaning and acts both as main and functional verb with no change of meaning. *Başlamak* forms constructions with verbal nouns ending in *-maya/-meye*, which generally belong to dynamic situation types. The cohesion of a sign on paradigmatic level is depicted by its paradigmaticity, i.e. a degree to which a sign enters a paradigm. On syntagmatic level the cohesion of a sign is associated with the degree to which it depends on or attaches to other signs. *Başlamak* is the main means of expressing initial phase in the Turkish language. Verbal constructions with it are built according to the schema where *başlamak* follows a verbal noun, ending in *-maya /-meye*, easily divided into logical parts, not phraseologically related. Paradigmatic variability is the freedom with which the language user chooses a sign which is concerned with the restrictions on its usage. *Başlamak* is used with no limitations in the structures and forms peculiar to the Turkish language. Syntagmatic variability, in the case of a grammaticalized sign, concerns mainly its positional mutability with respect to those constituents with which it enters into construction. *Başlamak* is located in the postposition to its complement and syntactically serves link verbs functions majorly acting as a part of complex predicate. The analysis shows that development of *başlamak* doesn't show many signs of desemantization, obligatorification, coalescence and other processes characteristic to grammaticalization. Being semantically neutral and syntagmatically and paradigmatically variable it occupied its slot in expressing initial phase of action in the Turkish language.

**Keywords:** aspectual verbs, initial phase, grammaticalization, Ch. Lehmann's parameters, the Turkish language.

### 1 INTRODUCTION

The present paper takes up the issue of grammaticalization of *başlamak* and verbal constructions with it. *Başlamak* is one of the high-frequency inceptive verbs in the Turkish language and information about the level of its grammaticalization may help to give general conclusions about the grammaticalization of the category of inceptivity in the Turkish language. The analysis was carried on the material extracted from TS corpus (<https://tscorpus.com>) and O. Pamuk's "Other colors" (Pamuk 1999), "Snow" (Pamuk 2002), "The Museum of Innocence" (Pamuk 2008), books (approx. 357,000 words total). These two different types of

source materials were used in order to increase the credibility of the research by adding the possibility to compare the results. Grammaticalization of *başlamak* was investigated through synchronic perspective within Ch. Lehmann's parameters (Lehmann 2002), a set of criteria that used to determine the level of autonomy of a linguistic sign.

The organization of this paper is as follows. Section 2 provides some background information on grammaticalization theory. Section 3 presents more information about Ch. Lehmann's parameters and their application to our material. Section 4 provides results and makes the overview of *başlamak* through each of the six parameters. In section 5 we give conclusions and in section 6 we acknowledge financial support of this research.

## 2 GRAMMATICALIZATION THEORY

The term "grammaticalization" presumably appeared first in the "L'évolution des formes grammaticales" article of A. Meillet who defined it as the "passage of an autonomous word to the role of grammatical element" (Meillet 1912, p. 131). Modern studies in grammaticalization began in early 1970s with the works of T. Givon whose slogan 'Today's morphology is yesterday's syntax' opened a new perspective for understanding grammar (Givon 1971, p. 413). T. Givon provided numerous examples, when lexical items function as grammatical categories, i.e. tense, aspect and modality markers derive from a small group of verbs, e.g. the development of the Turkish continuous tense suffix *-yor* from the base form *yürü-* 'to walk' (Demirci 2008, p. 134-135). Approximately at that period the definition of grammaticalization which is now considered classical and which we follow in our study appeared:

Grammaticalization is an increase of the range of morpheme advancing from a lexical to a grammatical form or from a less grammatical to a more grammatical status (Kurylowicz 1975, p. 52).

Since then many landmark research volumes and collected works have appeared (e.g. Traugott and Heine (eds.) 1991, Hopper, Traugott 1993, Lehmann 2002, Heine, Kuteva 2002, Heine (ed.) 2012). They helped to bring theoretical grounds to grammaticalization theory by suggesting principles and criteria of language change, i.e.:

1. Layering. Grammaticalization doesn't proceed with eliminating old forms and substituting them with the new ones, on the contrary old and new forms coexist and interact with each other.
2. Divergence. Divergence means that the grammaticalization of form doesn't always involve the disappearance of its lexical uses.
3. Specialization. Specialization is a central aspect of grammaticalization which presupposes reduction of the variety of forms with the selected smaller number of forms assuming more general grammatical meanings. For example, there are several durative forms in the Turkish language, but the ones which possess more specific meaning, like *-ip gitmek* 'to last from present to future' or *Egelmek* 'to last from past to present' are gradually disappearing of usage giving way to more regular morphological forms (Kondratjev 1992, pp. 43-46).
4. Persistence. Persistence presupposes that as long as a grammaticalized form continues to have a grammatical role, some traces of its original lexical meanings tend to adhere to it, and details of its lexical history may be reflected in constraints of its lexical distribution. Above mentioned forms serve a good example to it as far as they originate from motion verbs *gitmek* 'to go' and *gelmek* 'to come' the viewpoint differences between which clearly affected the meaning of grammaticalized durative forms.
5. De-categorialization. Adverbs, auxiliaries, prepositions and other "minor" categories always derive from the prime categories like noun or verb and never reverse (Hopper 1996, pp. 230-231).

Nowadays grammaticalization theory encompasses a range of approaches and theoretical orientations which show that grammaticalization is far from being a refined concept and various definitions may be proposed. Firstly, grammaticalization is delineated in terms of different schools of thought within linguistics. Generative linguists find parallels between language evolution and language acquisition stating that both processes frequently involve the same semantic changes (Croft 2001; Diessel 2012). The adherents of school of pragmatic linguistics invoke discourse and context-induced principles of language change. Cognitive approaches operate on the principle that language is built up of conventionalized form-meaning pairings that represent all grammatical knowledge (Auwera, Almen, Mon 2015, p. 640).

Secondly, grammaticalization may be approached from diachronic and synchronic perspectives. Viewed diachronically, it involves etymology and the taxonomy of the possible language change. Tracking

grammaticalization through synchronic variation is based on the idea that any diachronic prediction has its synchronic consequences which can be revealed through its distribution.

Finally, different frameworks are applied to different linguistic items under analysis, e. g. the development of discourse markers is viewed through their pragmatic functions, while the evolution of pronouns is more likely to be treated with conceptual manipulations. The present paper takes up Ch. Lehmann's approach to grammaticalization, which is oriented to estimate the level of grammaticalization of an item using a clear-cut range of criteria.

### 3 CH. LEHMANN'S PARAMETERS OF GRAMMATICALIZATION

Ch. Lehmann's parameters of grammaticalization are a set of criteria which help to determine the level of autonomy a linguistic item (Lehmann 2002, pp. 108-160). The autonomy, i.e. the freedom with which the sign is used, is converse to its grammaticality. As long as grammaticalization means a process in which something becomes or is made more grammatical, the more autonomous the item is the less the degree of its grammaticalization is.

Measuring the autonomy consists of three aspects. First, in order to be autonomous, the item must have certain prominence and be distinct from other members of the class, i.e. have certain **weight**. Second, autonomy decreases with rising **cohesion**, the extent to which the item systematically contracts certain relations with other items. Third, the item is the more autonomous the more **variability** it enjoys; this means a momentary mobility or shiftability with respect to other signs (Lehmann 2002, p. 109).

These three aspects work both on paradigmatic and syntagmatic axes and make parameters of grammaticalization (see Table 1).

Table 1. Ch. Lehmann's parameters of grammaticalization.

| parameter/axis     | paradigmatic             | syntagmatic             |
|--------------------|--------------------------|-------------------------|
| <b>weight</b>      | integrity                | structural scope        |
| <b>cohesion</b>    | paradigmaticity          | bondedness              |
| <b>variability</b> | paradigmatic variability | syntagmatic variability |

As Ch. Lehmann notes (Lehmann 2002, p. 110), the weight of the item, viewed paradigmatically, is its **integrity**, its substantial size, both on the semantic and phonological sides. Viewed syntagmatically, it is its **structural scope**, that is, the extent of the construction which it enters or helps to form. In our study we enlarged the criteria for analyzing the structural scope by investigating the semantic composition of verbal constructions with *başlamak*, i.e. the situation types it forms. Each situation type (dynamic or stative) is represented by different lexical classes of verbs (Quirk 1985, p. 200–213). And the study of frequency of collocation of *başlamak* with different lexical verbs may reveal how semantics of situation types interacts with initial phase meaning.

The cohesion of an item with other items in a paradigm will be called its **paradigmaticity**, that is, the degree to which it enters a paradigm, is integrated into it and dependent on it. In case of *başlamak* it mainly concerns its place among other means for expressing inception. The cohesion of an item with other signs in a syntagm will be called its **bondedness**; this is the degree to which it depends on, or attaches to, such other signs. As long as only verbal constructions are in the focus of our attention, only the degree to which *başlamak* is dependent on the verbal noun is taken into account.

The **paradigmatic variability** of a sign is the freedom with which the language user chooses a sign. In constructions with *başlamak* it is concerned with restrictions imposed by the Turkish language system on the expression of initial phase meaning. The **syntagmatic variability** of a sign is its flexibility according to other elements of syntagm. The flexibility of *başlamak* is considered within the properties of its functioning in a clause.

In this way, Ch. Lehmann's parameters are not an invariable list of properties, on the contrary, they are specific and the full set of them will not work equally on one item. Furthermore, some of the parameters correlate positively, others – negatively. With increasing grammaticalization, cohesion of the item increases as well, but its weight and variability decreases (Lehmann 2002, p. 110-111).

## 4 RESULTS AND DISCUSSION

Having studied structural and semantic characteristics of *başlamak* and verbal constructions with it, we present the acquired results respectively within each of Ch. Lehmann's aspects of grammaticalization. First we touch upon the weight of *başlamak* which is embodied in its integrity and structural scope properties. Then we refer to the level of its cohesion to other items which is determined by the parameters of paradigmaticity and bondedness. And finally we investigate paradigmatic and syntagmatic variability of *başlamak*.

### 4.1 Integrity

In the framework of grammaticalization integrity is associated with the changes in semantic and phonological aspect of a linguistic item. For *başlamak* no signs of phonological attrition (gradual loss of phonological substance) were found. Semantic integrity of an item is mainly concerned with its semantic representation and desemantization or bleaching, i.e. loss of semantic content (Lehmann 2002, p. 114). The verb *başlamak* originates from Old Turkic body part noun *baş* 'head' which developed into the verb with the help of derivational affix *-la-* and acquired 'be at the head' meaning (Nişanyan sözlük). According to "Dictionary of Modern Turkish" (Güncel Türkçe Sözlük) and "Comprehensive dictionary of Turkish" (Büyük Türkçe Sözlük) at this stage of its development *başlamak* has initial phase meaning. In general terms, new phasal meaning is likely to follow SPACE -> TIME metaphorical transfer and have body part term (*baş*) as its conceptual source.

### 4.2 Structural scope

Structural scope is associated with the diversity of semantic subjects and situation types *başlamak* collocates with. Analysis shows that in literary texts 465 collocations in 403 contexts were registered. The number of collocations differs from the number of contexts due to the fact that in certain cases *başlamak* collocates with more than one verbal noun and can be attributed to more than one subject. To make the figures of the research compatible we compared 465 extractions from literary texts with 500 TS Corpus instances of *başlamak*. Analysis shows that *başlamak* prefers human beings as its subject, but numbers differ: 410 extractions from literary texts versus 335 corpora instances. The acquired constructions generally belong to either dynamic or stative situation types (acc. to Quirk 1985, p. 200–213)

Dynamic situation types includes collocation of *başlamak* with the verbal nouns denoting:

- 1) activity: *aramak* 'to search', *kullanmak* 'to use', *oyunmak* 'to play'; including the so-called reciprocal actions: *buluşmak* 'to meet', *dövüşmek* 'to fight', *öpüşmek* 'to kiss', *görüşmek* 'to see each other';
- 2) communication: *anlatmak* 'tell', *sormak* 'ask', *konuşmak* 'speak', *tartışmak* 'argue'; or speech and sound production: *okumak* 'to read', *şakılamak* 'to twitter', *ses çıkarmak* 'to make a sound';
- 3) motion: *gitmek* 'to go', *yürümek* 'to walk', *zıplamak* 'to jump'.

The correlation of these types of collocations in the two material sources is illustrated in Fig. 1.

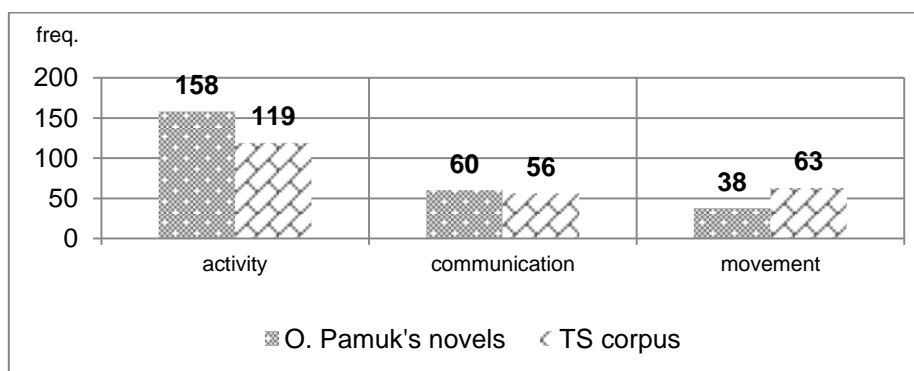


Fig. 1. The combination of *başlamak* with lexical types of verbal nouns denoting dynamic and stative situation types in O. Pamuk's novel and TS Corpus.

Stative situation types are represented by:

- 1) intellectual states: *düşünmek* 'to think', *sanmak* 'to suppose', *özdeşleştirmek* 'to associate', *hatırlamak* 'to remember';
- 2) states of emotions and attitudes: *kızmak* 'to get angry', *kıskanmak* 'to envy', *korkmak* 'to fear', *sevmek* 'to love', *öfkelenmek* 'to be furious', *inanmak* 'to believe', *ilgilenmek* 'to be interested';
- 3) states of perception: *görmek* 'to see', *işitmek* 'to hear', *hissetmek* 'to feel';
- 4) states of bodily sensations: *acımak* 'to hurt', *sarsılmak* 'to shudder', *üşümek* 'to feel cold'.

The frequency of occurrences of stative situation types in O. Pamuk's novels and TS Corpus is shown in Fig. 2.

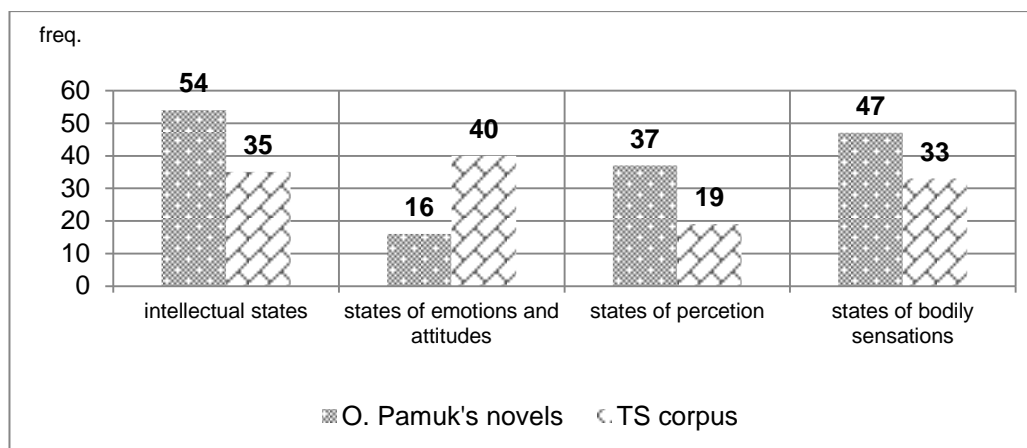


Fig. 2. The combination of *başlamak* with lexical types of verbal nouns denoting stative situation types in O. Pamuk's novel and TS Corpus.

As we see stative situation types are better represented in O. Pamuk's novels, the only exception are emotional states are covered more in TS Corpus. The difference in numbers can't be treated as substantial one as originally we had less corpora instances where verbal constructions with *başlamak* were used with agentive subject (410 extractions from literary texts versus 335 corpora instances).

There were also registered cases of collocation of *başlamak* with verbs, belonging to stance situation type, which is intermediate between the stative and dynamic categories and includes verbs like *oturmak* 'to sit, to reside', *yaşamak* 'to live' (4 extractions both from corpus and literary texts).

Cases when *başlamak* is used with non-agentive subjects make 165 instances in TS Corpus and the times less in the literary texts (53 extractions). Constructions with this type of subjects predominantly belong to one of the following groupings:

- 1) subjects denoting natural phenomena with motion verbs *kar sepiştirmeye başladı* 'snow started to fall', *kar/yağmur yağmak* 'it started to snow/rain', *duman tütmeye başladı* 'it started to exhaust fumes';
- 2) subjects denoting physiological processes with motion verbs: *yürek/ kalp atmaya başlıyor* 'heart starts to beat', *kan dolaşmaya başladı* 'blood started running';
- 3) subjects denoting mechanisms and appliances with verbs denoting its functional features: *gong/horna çalmaya başlıyor* 'gong/horn starts to ring', *perde kapanmaya başlıyor* 'curtains start to close'; or its physical properties: *enkaz paslanmaya başladı* 'carcass started to rust';
- 4) subjects denoting material objects or abstract concepts with verbs belonging to diverse lexical types (activity, motion, existence etc.): These are mostly the cases of personification, i.e. the attribution of a personal nature or human characteristics to something non-human (English Oxford Living Dictionaries).

To sum it up, the numbers of occurrences of both dynamic and stative situation in O. Pamuk's novels and TS Corpus do not differ substantially which leads to the idea that expressing initial phase may be more important for dynamic situations in general and activity situations in particular. While expressing the beginning of state may be of lesser relevance to the speaker.

### 4.3 Paradigmaticity

Paradigmaticity is the formal and semantic integration of a single unit into the paradigm. In our case, it is associated with the place of *başlamak* among other means of expressing inceptive meaning. The term 'inceptive' will be used as the Oxford English dictionary defines it 'expressing the beginning of an action' (English Oxford Living Dictionaries). These means are mainly lexical and may be divided into two groups. The first group comprises phasal adverbials (*ilk defa* 'for the first time', *yeni* 'newly') and expressions (*ortaya çıkmak* 'come to exist', *baş göstermek* 'appear') which have initial phase meaning in their semantics. The members of the first group are formed within major classes of words which furnish items that enter into grammaticalization channels. The second group are aspectual verb constructions, in our case understood as conventional form-meaning pairings, in which grammatical meaning is expressed by semi-auxiliary inceptive verb and lexical meaning is expressed by the verbal noun. These type of constructions appear to be a more regular means of expressing inceptivity than members of the first group. In the Turkish language there are at least three aspectual verbs with initial phase meaning: *başlamak*, *koyulmak* 'set out' and *girişmek* 'take up'. Analysis shows that *başlamak* has pure inceptive meaning while constructions with *koyulmak* and *girişmek* may get attemptive meaning: *kitabın gizli mantığını bulmaya girişti* 'he tried to find the hidden logic of the book', *bankaların batmasını engellemeye koyuldu* 'he tried to prevent the collapse of banks'.

Another factor that is closely connected to the paradigmaticity and hence grammaticality of the item is its frequency, i.e. the more grammatical the item is, the higher frequency it possesses (Lehmann 2002, p. 127). The three verbs differ majorily in their frequency: according to the data extracted from TS Corpus *başlamak* has 1808,53 instances per million words while *girişmek* and *koyulmak* appear only in 381,01 and 20.88 instances per million words respectively.

In this way basing on its frequency and semantics, *başlamak* can be considered to take central position in expressing inceptivity in Turkish (as well as many other Turkic languages (Juldashvili 1965, p. 52)).

### 4.4 Bondedness

Bondedness is attributed to the level of cohesion between the two elements in the construction. Constructions with *başlamak* are built according to the schema where *başlamak* follows a verbal noun, ending in *-maya /-meye*, with no possibility insertion of other elements between them. Decreasing or impossible insertability is an important syntactic criterion testifying increasing level of bondedness of the item and accordingly its rising grammaticalization.

Another criterion of bondedness is the level of coalescence between the two elements of the construction. It may proceed according to either of the two following principles. One possibility is that the position of a grammaticalized item in a syntagm is defined by grammatical relations. The second possibility is that its position in a syntagm is defined by sequential order relations (Lehmann 2002, p. 132). As long as *başlamak* is syntactically attached to a verbal noun and performs grammatical function of denoting initial phase of action, its coalescence in the syntagm clearly follows the first possibility.

One more process that affects bondedness of is cliticization, i.e. weakening or loss of boundaries between the elements of construction. Analysis shows that verbal constructions with *başlamak* are easily divided into logical parts and are not phraseologically related, which means that the meaning of the construction is reduced to the meanings of its constituents.

### 4.5 Paradigmatic variability

Paradigmatic variability is "the freedom with which the language user chooses a sign" (Lehmann 2002, p. 123) which in our case is concerned with the restrictions on the usage of that item. *Başlamak* is used with no limitations in the structures and forms peculiar to the Turkish language. Furthermore, unlike phasal verbs in, for example, the English and the Russian languages it may be used in passive voice in verbal constructions: *İstanbul'da yeni yapılan pek çok binaya aile adları verilmeye başlanmıştı* 'people started to attach one's family name to one's newly constructed apartment building'.

One more important factor to be considered are verb forms that *başlamak* prefers. As TS corpus word frequencies show, it is used predominantly in past-time or participial forms (See Table 2) with other forms not being to that extent common.

Table 2. General frequency of forms of *başlamak* according to TS Corpus (per 1 million words).

| No | Form              | General frequency<br>(per million words) |
|----|-------------------|--|
| 1. | <i>başladı</i>    | 175                                      |
| 2. | <i>başlayan</i>   | 174,5                                    |
| 3. | <i>başlıyor</i>   | 82,91                                    |
| 4. | <i>başlayacak</i> | 78,21                                    |
| 5. | <i>başlar</i>     | 64,59<br><i>etc...</i>                   |

Results from literary texts partially overlap with the corpora results with past-tense *di-* form of *başlamak* appearance in almost half of the contexts (201 out of 403). But *-an*, *-iyor*, *-acak* and *-ar* forms are not frequent at all being used only in one out of ten contexts. Such different results may only testify certain inclination of *başlamak* to be used in past-tense reference.

#### 4.6 Syntagmatic variability

Syntagmatic variability, in the case of a grammaticalized unit, concerns mainly its positional mutability with respect to those constituents with which it enters into construction (Lehmann 2002, p. 140). *Başlamak* is located in the postposition to its complement and syntactically may serve diverse functions in different types of clauses, but is more frequently used as a predicate.

### 5 CONCLUSIONS

*Başlamak* is a semantically neutral functional verb, the semantics of which interacts with the lexical semantics of the verbal noun it is used with. It adheres to past-tense forms (*-di*) and is found mostly within dynamic types of situations, which may testify that for a speaker a beginning of action is more important than a beginning of state. The development of *başlamak* does not show many signs of desemanticization, obligatorification, coalescence and other processes characteristic to grammaticalization. Being syntagmatically and paradigmatically variable it occupied its slot in expressing inceptivity.

### 6 ACKNOWLEDGMENTS

This research was financially supported by Ministry of the Education of the Republic of Belarus (Grant NO. 16rp /2017).

### REFERENCE LIST

- Croft, W. (2001). Explaining language change. An evolutionary approach.
- Demirci, K. (2008). Dilbilgiselleşme üzerine bir inceleme. Türk Dünyası Sosyal Bilimler Dergisi, vol. 45.
- Diessel, H. (2012). Grammaticalization and language acquisition. The Oxford handbook of Grammaticalization.
- Givon, T. (1971). Historical syntax and synchronic morphology: An archeologist's field trip. Papers from the Regional Meeting of the Chicago Linguistic Society, vol.7
- Heine, B. (2012). The Oxford handbook of Grammaticalization.
- Heine, B. and Kuteva, T. (2002). Word lexicon of grammaticalization.
- Hopper, P. J. (1996). Some recent trends in grammaticalization. Annual Review of Anthropology, vol. 25.

- Hopper, P. J. and Traugott, E. (1993). Grammaticalization.
- Juldashev, A. A. (1965). Analiticheskie formy glagola v Tjurkskih jazıkah [Analytical verb forms in the Turkic languages].
- Kondratjev, V. G.(1992). Osnovnye tendencii v razvitii grammaticheskogo stroja tureckogo jazyka [Major trends in the Turkish language development].
- Kurylowicz, J. (1965). The evolution of grammatical categories. *Diogenes* ( 51).
- Lehmann, Ch. 1982 (2002). Thoughts on gramaticalization.
- Meillet, A. (1912). L'évolution des formes grammaticales. *Scientia – Rivista di Scienza* (12).
- Pamuk, O. (1999). Öteki renkler.
- Pamuk, O. (2002). Kar.
- Pamuk, O. (2008). Masumiyet Müzesi.
- Quirk, R., Greenbaum, S., Leech, J. and Svartvik, J. (1985). A Comprehensive Grammar of the English Language.
- Traugott, E. and Heine, B. (1991). Grammaticalization.
- Van der Auwera, J., Olmen, D. V. and Mon, D. D. (2015). Grammaticalization. Handbook of Cognitive Linguistics.
- Büyük Türkçe Sözlük. [http://www.tdk.gov.tr/index.php?option=com\\_bts&view=bts](http://www.tdk.gov.tr/index.php?option=com_bts&view=bts).
- Güncel Türkçe Sözlük. [http://www.tdk.gov.tr/index.php?option=com\\_gts&view=gts](http://www.tdk.gov.tr/index.php?option=com_gts&view=gts).
- Nişanyan sözlük. <http://www.nisanyansozluk.com>.
- English Oxford Living Dictionaries. <https://en.oxforddictionaries.com>
- TS Corpus. <https://tscorpus.com>.