



IEEE TÜRKİYE SECTION

**UBMK'25**

**Bildiriler Kitabı**

**Proceedings**

**Editor Eşref ADALI**



**10. Uluslararası Bilgisayar Bilimleri ve Mühendisliği Konferansı**

**10th International Conference on Computer Science and Engineering**

**17-18-19 Eylül (September) 2025 İstanbul - Türkiye**

Media type	Part Number	ISBN	Online ISSN
XPLORE COMPLIANT	CFP25L97-ART	979-8-3315-9975-1	2521-1641
CD-ROM	CFP25L97-CDR	979-8-3315-9974-4	

2495	Combined Method of Citation Sentiment Analysis: Linguistic Features and Machine Learning	Ainur Akhmediyarova ; Dina Oralbekova ; Zhibek Alibiyeva ; Nurzhan Mukazhanov ; Dinara Kassymova ; Alimkulova Janna	514-518
2497	Kazakh Text Classification using Deep Learning Approaches	Nurzhan Mukazhanov ; Tatiana Batura ; Aigerim Yerimbetova ; Mussa Turdalyuly ; Bakzhan Sakenov ; Ainur Bayekeyeva	519-524
2498	Comparative Analysis of BERTurk and Gemma Models for Turkish Text Classification by LoRA Fine-Tuning	Tufan Yiğit Çiftçioğlu ; Metin Bilgin	525-528
2499	An LSTM-Based Model and Algorithm for Uzbek Language Text Generation	Sharipov Maksud ; Qurbanova Rozikajon ; Kurbanova Lola	529-532
2513	A Large-Scale Dataset and Citation Intent Classification in Turkish with LLMs	Kemal Sami Karaca ; Bahaeddin Eravcı	533-538
2529	Automatic Identification of Phrase Structures in Uzbek Texts	Elov Botir Boltayevich ; Abdullayeva Oqila Xolmo'minovna ; Abdullayeva Nazokat Isayevna ; Saodat Israilova ; Nizomova Fotima Bo'rixo'ja qizi	539-544
2532	Detection of Disinformation on Social Platforms: A Review of Computational Approaches and Challenges	Duman Telman ; Aigerim Yerimbetova ; Elmira Daiyrbayeva ; Madina Sambetbayeva ; Bayangali Abdygalym ; Almas Turganbayev	545-550
2547	GenAI-Based Threat Intelligence Model: A Comparative Performance Analysis with Traditional Tools	Ahmet Albayrak ; Arafat Şentürk ; Serdar Arpacı	551-556
2550	Exploring Turkish Idiomaticity with Large Language Models	Özge Umüt ; Atakan Site ; Doğukan Arslan ; Gülşen Eryiğit	557-562
2557	Structured Financial QA with LLMs: Fine-Tuning vs. Code-Augmented Retrieval	Alperen Çağlayan ; Saliha Nur Gökçe ; Değer Ayata	563-568
2558	Bankacılık için Ulusal Elektronik Tebligat Sistemi Otomasyonu <i>National Electronic Notification System Automation for Banking</i>	Mehmet Yasin Akpınar ; Murat Güç	569-574
2570	Beyond Polarity: Unsupervised Sentiment Analysis in Turkish Social Media Texts	Zeynep Bozkurt ; Arzum Karataş	575-580
2574	Linguistic And Computational Solutions For Translating Lexical Lacunae In Neural Machine Translation Systems	Erkinjon Gaybullaev ; Manzura Abjalova ; Munojot Adilova	581-586
2579	Behavior content analysis of subjects in digital environment using NLP	Nikolai Prokopyev ; Pavel Ustin ; Fail Gafarov	587-589
2587	Learning Rate Sensitivity of Multilingual Transformer Based Large Language Models in Turkish Binary Classification	Selma Alav ; Bihter Daş ; Kristin Surpuhi Benli	590-596
2602	Geri Alma-Artırılmış Üretim (RAG) Sistemlerinde Vektör Veritabanı ve LLM Modellerinin Değerlendirilmesi <i>Evaluation of Vector Database and LLM Models in Retrieval-Augmented Generation (RAG) Systems</i>	Hilal Tektöz ; Halit Erdoğan ; Harun Uz ; Hatice Altınok	597-601
2622	Methodological Approaches to the Semantic Description of Explanatory Lexemes within Religious Text Corpora	Manzura Abjalova ; Fayiza Tursunboyeva ; Aminova Nafisa Istamovna ; Yulduz Shodmonaliyeva	602-607
2625	Özbek Dilindeki Deyimlerin Üslup Özelliklerinin Belirlenmesi <i>Determination of Stylistic Features of Idioms in Uzbek Language</i>	Manzura Abjalova ; Umida Rashidova ; Sarvinoz Rasulova ; Sarvinoz Sharipova	608-612
2628	Designing a Domain-Specific Semantic Network for Uzbek Mining Terminology: A Computational Lexicography Approach	Maftuna Sobirova ; Manzura Abjalova ; Mukaddas Abduraxmanova	613-618
2630	Development of a rule-based model and algorithm for predicate identification in Uzbek language texts	Maksud Sharipov ; Ixtiyor Avezmatov ; Hushnubek Adinaev	619-623
2646	Sentiment Analysis of Uzbek Idioms	Manzura Abjalova ; Sarvinoz Sharipova ; Shoxista Istamova	624-629
2658	Performance Analysis of Text Chunking Methods for LLM-Based Document Summarization	Salih Dünder ; Öykü Şahin ; Mehmet Ali Özer	630-634
2660	Automatically Determine Semantic Relations in the Uzbek Language Using machine Learning	Umirova Svetlana ; Alisher Gafforov ; Islamova Dinara ; Kholmuhamedov Bakhtiyor ; Yusupova Orziboni ; Narzieva Mamura	635-640

# Özbek Dilindeki Deyimlerin Üslûp Özelliklerinin Belirlenmesi

## Determination of Stylistic Features of Idioms in Uzbek Language

Manzura Abjalova

Taşkent Devlet Özbek Dili ve Edebiyatı Üniversitesi

Tashkent, Uzbekistan

abjalova.manzura@gmail.com

ORCID: 0000-0002-1927-2669

Umida Rashidova

Semerkant Devlet Üniversitesi,

Semerkant, Özbekistan

umida\_rashidova@bk.ru

ORCID: 0000-0001-8887-2572

Sarvinoz Rasulova

Taşkent Devlet Özbek Dili ve Edebiyatı Üniversitesi

Tashkent, Uzbekistan

sarvina\_ok@mail.ru

ORCID: 0009-0003-8717-2568

Sarvinoz Sharipova

Semerkant Devlet Yabancı Diller Enstitüsü

Semerkant, Özbekistan

shsara232323@gmail.com

Öz— Dilciliğin deyim bilim sahası, yapısal-dilbilgisel özelliklerinin yanı sıra işlevsel-üslûbi imkânların araştırılmasıyla da ilgilenir. Deyim biliminin deyim birimleri işlevsel-üslupsal açıdan araştıran bölümü “deyimbilimsel üslupbilim” diye adlandırılır. Bu makalede, Doğal Dil İşleme (DDI) yaklaşımına dayalı olarak Özbek dilindeki deyim birimlerin üslup özelliklerinin otomatik olarak tanımlanması konusu ele alınmıştır. Deyim birimlerin anlamsal ve üslup özellikleri; edebî ve bedîî, kamusal, resmi ve günlük konuşma tarzlarının karakteristiği bakımından belirlenmiş, bağlamsal kullanımlarına göre sınıflandırma modelleri önerilmiştir. Çalışmada morfolojik tahlil, bağlamsal kullanım (SBERT-uz), istatistiksel yöntemler ve elle tanımlanmış korpuse dayalı makine öğretmenin SBERT usulu kullanılmıştır. Elde edilen sonuçlar, Özbek dilinde üslup açısından belirsiz deyim birimleri otomatik olarak ayırt etmeye ve üslup bakımından tanımlamaya olanak tanımaktadır. Bu makalede, Özbek dilindeki deyim birimlerin üslup bakımından karşılıklı kullanım özellikleri, diferansiyel işaretleri ve bunların kendine özgü yönleri örnekler üzerinden analiz edilmiştir.

*Anahtar Sözcükler — deyim birim; deyimsel unsur; üslup özgünlüğü; bağlamsal tahlil; üslupsal sınıflandırma; Özbek dili.*

*Abstract—* The idiomatic field of linguistics deals with the study of functional-stylistic possibilities as well as structural-grammatical features. The part of idiomology that studies phraseological units from a functional-stylistic point of view is called “idiomatic stylistics”. In this article, the issue of automatic identification of stylistic features of idiomatic units in the Uzbek language based on the natural language reprocessing (NLP) approach is considered. The semantic and stylistic features of idiomatic units are determined in terms of the characteristics of literary and literary, public, formal and colloquial speech styles, and classification models are proposed according to their contextual use. Morphological analysis, contextual usage (SBERT-uz), statistical methods and machine learning procedures based on manually defined corpus were used in the study. The results obtained allow to automatically distinguish and stylistically identify stylistically ambiguous phraseological units in Uzbek language. In this article, the features of stylistic interchangeability of phraseological units in Uzbek language,

their differential signs and their peculiarities are analyzed on the basis of examples.

*Keywords — Idiomatic unit, phraseological unit, idiomatic element, stylistic originality, contextual analysis, stylistic classification, Uzbek language.*

### I. GİRİŞ

Günümüzde doğal dili tekrar işleme (Natural Language Processing, NLP) teknolojileri, muhtelif dillerde stilistik tahlil, duygusal analiz ve metni otomatik sınıflandırma gibi görevleri gerçekleştirmek için yaygın olarak kullanılmaktadır. Özbekçe gibi morfolojik olarak zengin ve bağlama bağlı çok anlamlı dillerde, deyim birimlerin otomatik olarak algılanması ve üslup farklılıklarının belirlenmesi güncel bilimsel konulardan biridir.

Deyim birimler doğası gereği, bağlamda farklı üslup anlamları kazanırlar. Bedîî metinlerde imgeleri güçlendirirken kamusal metinlerde duygusal etkiyi artırır ve resmi üslupta daha az kullanılırlar veya değiştirilmiş bir biçimde görünürler. Bu makalede, NLP araçlarını kullanarak Özbek dilindeki deyimbilimsel birimlerin üslup farklılıklarını belirleme sorunu ele alınmaktadır.

“Deyimbilimsel üslup bilimi” terimi, 1960’lardan beri Milletler Topluluğu dil biliminde kullanılmaktadır. Başlangıçta bu terim V.L. Arkhangelsky [1] ve A.V. Kunin’in eserlerinde kullanılmış ve daha sonra popüler hâle gelmiştir. “Dilbilimsel Üslup Biliminin Sorunları” adlı bilimsel konferansta Prof. A.V. Kunin “Deyimbilimsel Üslup Biliminin Temel Kavramları” [2] başlıklı bir tebliğ ile konuşma yapmıştır. Bu konuşmada, deyimbilimsel üslup biliminin genel üslup biliminin bir bileşeni olduğu ve deyim birimlerinin üslup özelliklerini incelediği vurgulanmıştır.

### II. KAYNAKLARIN İNCELENMESİ

Deyimsel birimler, özellikle deyimler, kalıplar ve metaforlar; doğal dilde karmaşık anlamsal yapıları sahiptir.

Bunların otomatik olarak saptanması ve sınıflandırılması NLP alanındaki güncel sorunlardan biridir. Son yıllarda bu konuda çok sayıda önemli çalışma ve uygulamalı gelişme gerçekleştirilmiştir. Bu bağlamda dil korpuslarına dayalı deyimel birimlerin otomatik saptanması sorunu, incelenmiştir. Sag [3] ve Moon [4] gibi araştırmacılar, karşılıklı bilgi (PMI), t-testi, log-likelihood gibi istatistiksel yöntemleri kullanarak yöntemlere dayalı deyimel birimleri saptamayı önermişlerdir. Bu yaklaşımlar deyimel birimlerin kalıcı ve yarı-kalıcı kombinasyonlar olarak tanınması için önemlidir. Fazly [5], İngilizcedeki deyimel birimleri sınıflandırmak için bağlamsal özellikleri, n-gram modellerini ve doğrusal sınıflandırma usüllerini kullanmıştır. Salton ve diğerleri [6], deyimler ile edebî ifadeler arasındaki farkı RNN (Recurrent Neural Network) kullanarak modellemiştir. Zhou vd. [7], deyim birimlerinin biçimsel çağrışımlarını dönüşümsel modeller esasında anlamsal olarak analiz etmiştir. Liu ve Hwa [8], deyimleri anlamsal olarak tanımlamak ve sanatsal bakımdan deyimel yapılarına göre sınıflandırmak için BERT modelini kullanmıştır. Paradis ve diğerleri [9], bağlamsal kullanımlar yoluyla deyim birimlerin biçimsel özelliklerini analiz etmiş ve stilistik belirteçleri belirlemiştir.

Liu ve diğerleri [10], İngilizce, Almanca ve Fransızca'daki deyim birimlerini karşılaştırarak evrensel bağlamsal özellikleri belirlemeye çalışmıştır. Diller arası kullanımlar (örneğin, XLM-RoBERTa) hakkında yeni olasılıklar keşfetmiştir. Deyim birimleri, Stamatos [11], Kestemont ve diğerleri [12] tarafından önerilen stilistik analiz modellerinde önemli parametreler olarak kabul edilir. Bu yaklaşımlar, deyimlerin resmi, günlük, sanatsal ve kamusal üsluplarda nasıl kullanıldığını ayırt etmede yardımcı olmaktadır.

Özbek dilbiliminde deyimlerin sanatsal doğası üzerine araştırmalar; I. Kochkortoyev, E. Umarov, O. Abdunazarov, H. Qahhorova, M. Yoldoshev, B. Jurayeva, A. Mamatov [13], [14] gibi bilim insanları tarafından yürütülmüştür. NLP'de, Özbek deyimlerinin bir veritabanının oluşturulması [15], [16], bunların çevirisi, özellikle Özbekçe ve Türkçede deyimlerin çevirisindeki sorunlar [17] ve bunların üstesinden gelme yolları, tercüme programını iyileştirme ilkeleri, deyimlerin morfolojik ve semantik olarak eşleştirilmesinin önemi [18], [19] üzerinde durulmuştur. Ancak, deyimlerin üslup açısından tanımı sorunu, dünya dilbiliminde henüz mutabık bir çözüme kavuşmamıştır ve bu konu Özbek dilbiliminde özel olarak incelenmemiştir. L.I. Royzenzon şöyle der: "Deyimlerin üslupsal tanımı, sözcükbirimlerin üslupsal tanımından çok daha zordur. Bu alanda ülkemizde veya yurtdışında net kriterler geliştirilmemiştir" [20]. Deyimbilimin üsluplararası kullanımı kapsamında şu güncel konular ele alınmaktadır: üslupsal anlam ve bileşenleri, deyim birimin üslupsal farklılaşmasına ilişkin kriterler, Özbek deyimbilim sisteminin üslup katmanlarının belirlenmesi, deyimbilimin anlamsal benzerliği ve üslupsal dallanmaları, vb.

### III. YÖNTEMLER

Deyim birimin üslup özgüllüğünü belirlemek için 1) Kural tabanlı; 2) İstatistiksel tahlil; 3) Semantik analiz yöntemi; 4) N-gram modeli ve 5) Stilistik üslubu otomatik tespit etme yöntemlerinden faydalanıldı. Bu bağlamda ilk olarak 1)

Sözlüklü metot; ve 2) sözlüksüz yöntemler kullanılarak bir deyimbilim veri tabanı oluşturuldu.

İlk olarak, veri tabanı için bir Özbek deyimleri bankası oluşturacağız.

#### 3.1 Sözlüklü metot

Bu metot esasında Sh. Rahmatullayev'in "Özbek Dilinin Deyimsel Sözlüğü" (O'zbek tilining frazeologik lug'ati) adlı eserinden 3500 deyim seçildi. Bu sözlükte deyim izahı, cümledeki kullanımına örnek, 1. adımda bir kelime sorusu, bir varyantı ve eşanlamlıları gösterilmiştir.

#### 3.2 Sözlüksüz yöntem

Bu metot esasında Özbek dili multikorpusundan, okul ders kitaplarından ve resmi internet sayfalarından alınan metinlerde geçen 2900 deyim yer verildi. Bu deyimler, bağlamsal semantiklerine göre veritabanındaki parametrelerle eşleştirildi. Özellikle, 1) deyim; 2) deyim açıklaması; 3) semantik grubu; 4) kelime türü; 5) varyantı; 6) eşanlamlıları; 7) zıt anlamlıları; 8) üslubu belirlendi. Geniş çaplı araştırmamız için seçilen toplam 6500 deyim bu 8 parametreye göre analiz edildi. 100 deyim Özbek dilinin bedii, ilmî, resmî ve kamusal üslup tarzına uyarlandı.

Sözlüklü yöntem, deyim birimleri hakkında hazır bilginin bulunması nedeniyle en etkili yöntem olarak düşünüldü. Ancak günümüzde Makine Öğrenimi (Machine Learning) alanında, tekrarlamaya dayalı sözlüksüz bir yöntem olan yinelemeli sinir ağları (RNN) kullanılmaktadır ve bu yöntem metinlerdeki deyimlerin çoğunu kapsayacak niteliktedir.

Örneğin: Sözlükte Özbek dilinde "ölmek" anlamına gelen 10 deyim tespit edilmiştir. Sözlüksüz yöntem kullanılarak; Özbek dili multikorpusu [www.uzcorpara.uz](http://www.uzcorpara.uz)'dan bu deyimlerin 357 semantik eşanlamlısı ve varyantı toplanmıştır.

Seçilen her deyim anlamı analiz edilmiştir. Örneğin, Çizelge I, Özbekçedeki "dunyodan ko'z yummoq" (gözlerini dünyaya kapatmak) deyimini "ölmek" anlamına gelmektedir. Bu deyim üsluplar arası çaprazlama yöntem olarak ele alındığında, ilmî üslupta "vafot etmoq", "halok bo'lmoq"; resmi üslupta "joniga qasd qilmoq" (intihar etmek), "hayotiga chek qo'ymoq" (hayatına son vermek) kamu üslubunda "umriga zomin bo'lmoq" (ömrüne nihayet vermek); bedii üslupta "ikki qo'li ko'ksida ketmoq" (iki eli göğsüne gitmek) gibi eşanlamlıları olduğu ortaya çıkmaktadır.

ÇİZELGE I. "ÖLÜM" SÖZLÜK BİRİMİNİN ÜSLUPLAR ARASI KULLANIMI

№	Üslup türü	Deyim	Örnek
1.	İlmî	Vafot etmoq, halok bo'lmoq	Ahamoniylar sulolasi vakili Kir II- mil.avv. 530-yildagi massagetlar bilan bo'lgan jangda <i>halok bo'ldi.</i> (Tarix. Qadimgi dunyo, 2017, 6-sinf, 118-b).
2.	Kamusal	umriga zomin bo'lmoq	Musiqa festivali to'rt kishining umriga zomin bo'ldi.
3.	Resmî	"joniga qasd qilmoq (suisid)", Hayotiga chek qo'ymoq	AQShda o'z joniga qasd qiluvchilar o'q otar qurollardan (52,1 foiz), o'zlarini osish va bo'g'ish (22,2 foiz), zaharlanish (17,6 foiz) va boshqa metodlar (8,1 foiz)dan foydalanishadi

4.	Bediî	ikki qo'li ko'ksida ketmoq	Qizingni ko'ndir, ko'zing ochig'ida erga tegsin. Bo'lmasa ikki qo'ling ko'ksingda ketadi. I. Rahim. Ixlos.
----	-------	----------------------------	--

Özbek dilindeki “*hayotining ko'kini kesmoq*” (hayatının kökünü kesmek) deyimini “hiçbir çaba sarf etmeden sonucu beklemek” anlamına gelir. Deyim kelimesi kelimesine çevrildiğinde, “*elmali turta, ağzıma gel*” anlamına verir. “*Suvdan quruq chiqmoq*” (Sudan kuru çıkmak) deyimini ise, “zor bir durumdan zarar görmeden kurtulmak” anlamında kullanılmaktadır.

Bir başka örnek (Çizelge II)

ÇİZELGE II. “ALDAMOQ” ANLAMINI KAPSAYAN DEYİMLERİN ÜSLUPLAR ARASI KULLANIMI

№	Üslup türü	Deyim	Örnek
1.	İlmî	Nayrang ishlatmoq, hiyla ishlatmoq	Mustamlakachi davlatlarning hukmron doiralari mustamlaka va qaram mamlakatlar xalqlarining ommaviy harakati ta'siri asosida <i>nayrang ishlatishga</i> majbur edi. (Qadimgi dunyo tarixi) Uch kunlik qattiq qamaldan so'ng Chingizxon hiyla ishlatishga majbur bo'ldi (7-sinf O'zbekiston tarixi)
2.	Kamusal	chuv tushirmoq, domiga ilinmoq, tuzog'iga ilintirmoq, laqqa tushirmoq, pand bermoq, qopqoniga tushirmoq	Jorjiya Meloni kremlparast prankerlarning hiylasiga «laqqa tushib» ular bilan telefonda gaplashdi. Uning ruslar «qopqoniga tushgani» muhokamalarga sabab bo'lmoqda. ( <a href="https://kun.uz/87938449">https://kun.uz/87938449</a> )
3.	Resmî	Ishonchini suiiste'mol qilmoq	Firibgarlikda <i>ishonchni suiiste'mol qilish</i> deganda, aybdorning mulkdor, mulkning boshqa egasi yoki mulkni uchinchi shaxslarga berib yuborish to'g'risida qaror qabul qilishga vakolatli boshqa shaxslar bilan ishonchli munosabatlardan g'arazli maqsadlarda foydalanishi tushunilishi lozim (lex.uz).
4.	Bediî	ko'zini bog'lamoq, ko'zini bo'yamoq, oldiga poxol solmoq, daryoga oborib sug'ormay kelmoq,	Agar ishlari chappasidan kelsa-yu, Shayboniyxonning qopqoniga tushsalar, ajal sirtmog'i eng avval Boburning bo'yniga tushishini hammalari bilardilar. (Yulduzli tunlar 125-b)

[www.kun.uz](http://www.kun.uz) haber sitesinde *aldamoq* 30 kez, *laqqa tushmoq* 2 kez, *tuzog'iga tushirmoq* / *ilinmoq* 7 kez kullanılmıştır.

[www.daryo.uz](http://www.daryo.uz) haber sitesinde *chuv tushirmoq* 97 kez, *tuzog'iga tushmoq* / *ilinmoq* 2 kez kullanıldığı tespit edilmiştir

[www.lex.uz](http://www.lex.uz) resmi evrak sitesinde 4 belgede **aldamoq** sözcük birimi 8 kez, **ishonchini suiiste'mol qilmoq** (güveni kötüye kullanmak) şekli 6 kez kullanılmıştır.

### 3.3 Kurallara dayalı metot

Bu metot esasında deyimlerin kullanım bağlamlarını analiz ederek anlamları belirlendi. Örneğin, “*joniga qasd qilmoq*” iborasi <jon>, <qasd>, <qilmoq> gibi ayrı ayrı anlamlara gelen sözcüklerden oluşmuş olsalar da, bağlam içinde tek bir simge olarak karşımıza çıkarlar ve “o'lmoq” (ölmek) anlamı ile özdeşleştirilir. “İkki qo'li ko'ksida ketmoq” deyimini ise 4 sözcükten oluşmaktadır, ancak veritabanında bu deyim tek simge şeklinde “o'lmoq” sözüyle tanımlanmaktadır. Eski Özbek dilinde kullanılan “*jon taslim qilmoq*”, “*jahondin ketmoq*” deyimleri “o'lmoq” (ölmek) ile değil, 1 sözcük grubu şeklinde, “o'z ajali bilan vafot etmoq” anlamıyla tanımlanmıştır. Sıradaki “*vujudining ayoqi ajal balchiqina botmoq*” deyimini ise “vujud”, “ayoq”, “ajal”, “balchiq”, “botmoq” gibi bağımsız sözcük birimlerden oluşsa da, bu deyim de bir anlam üzerine işaretlenmiş, “o'lmoq, vafot etmoq” olarak tanımlanmıştır. Demek oluyor ki; deyimın anlamı sözcüklerin müstakil anlamına bağlı değildir, yani:

Deyim anlamı ≠ bileşenlerin anlamları

Bu nedenle, her bir deyimveritabanında tek bir belirteç olarak tanımlanır ve semantik bağlamsal anlamıyla tanımlanır:

deyim=bağlamsal anlam

### 3.4 İstatiksel tahlil.

Bu araştırma sürecinde istatistiksel ve semantik analiz yöntemleri kullanılmıştır. Her bir deyimın kullanım sıklığını belirlemek için Özbek dili multikorpus'u kullanılmıştır. Bu korpusun avantajı, stilistik olarak karakterize edilmiş bir veri tabanı içermesidir. Metinler, meta verilerinde belirtilen stilistik parametreye göre bu veri tabanına girilmektedir. Kamusal ve bediî üslubu yansıtan korpus metinlerinde, “ölmek” anlamına gelen “*umriga zomin bo'lmoq*” deyimini 200 kez, “*joniga qasd qilmoq*” deyimini 199 kez kullanılmışken, “*jon taslim qilmoq*” deyiminin sıklığı 1'dir. Buradan anlaşılacağı üzere, üç deyim arasında “*umriga zomin bo'lmoq*” ifadesinin kullanım açısından aktif olduğu ve kamusal- bediî üslubu yansıtan tipik örnek olarak kabul edilmiştir.

### 3.5 Semantik tahlil.

Bu analiz biçimi; sözcük veya deyimın anlamını, semantik kapsamını, duygusal yükünü, bağlamsal bağımlılığını ve üslup işlevini belirlemeyi amaçlayan analitik bir yöntemdir. Kelimenin hangi çağrışıma sahip olduğunu ve hangi konuşma tarzında kullanıldığını belirlemeye yardımcı olur. Bir deyimın semantik analizi yoluyla üslup özgülüğünü belirleme aşamaları:

1. Sözcük veya deyimın leksik-semantik anlamı belirlenir. Bu durumda, kelimenin ana sözcüksel anlamı dikkate alınır (örneğin, “ölmek” - hayotdan ko'z yumish/hayata gözlerini kapamak).
2. Eşanlımlı satır(lar) belirlenir. Aynı anlamı ileten diğer ifadeler dikkate alınır: “vafot etmoq”, “halok bo'lmoq”, “joniga qasd qilmoq”, vb.
3. İbarenin stilistik rengi belirlenir, yani kelimenin ne kadar resmiyet, bediilik, duygusalılık ve kamusalılık tonuna sahip olduğu belirlenir.

4. İbarenin bağlamı belirlenir. Bu deyim hangi türde, metin türünde kullanılıyor? (bilimsel makale, bediî eser, kitle iletişim araçları, resmi belge...)

5. Sözdizimsel üslup belirlenir. Elde edilen bilgilere dayanarak deyimın aşağıdaki konuşma üsluplarından hangisine ait olduğu belirlenir: resmi, bediî, kamusal, ilmi, günlük konuşma. (Çizelge III).

ÇİZELGE III. SEMANTİK TAHLİL YARDIMIYLA “ÖLMEK” ANLAMINA GELEN DEYİM BİRİMLERİN ÜSLUP ÖZGÜNLÜĞÜ

Deyimler	Anlam (Semantik çekirdek)	Çağrışım/imelem	Üslupsal özellik (Stilistik açıklama)
<i>Vafot etmoq</i>	Hayotdan ko'z yumish (hurmat bilan tilga olish)	Resmi, nazik, halis	<b>Resmî üslup</b> — belgeler, basında kullanılan evraklar
<i>Jon taslim qilmoq</i>	O'lish (tasviriy, diniy tusli)	Duygusal, bediî	<b>Bediî/ kamusal</b> — edebiyatta, makalelerde
<i>Halok bo'lmoq</i>	O'lish (katta falokat natijasida)	Tarafsız veya katı, resmî	<b>İlmî</b> — istatistiksel verilerde, resmi raporda
<i>Ikki qo'li ko'ksida ketmoq</i>	O'lish (obrazli, milliy kolorit)	Bediî, figüratif	<b>Bediî üslup</b> — folklorida, edebî eserlerde
<i>Joniga qasd qilmoq</i>	O'z joniga qasd qilish (suisid)	Resmî, toplumsal sorun odaklı	<b>Resmî / tubbi / içtimai</b> — haberler, analizler
<i>Umriga zomin bo'lmoq</i>	O'limga sabab bo'lmoq	Suçlayıcı ahenk, duygusal	<b>Kamusal</b> — gazeteler, makaleler

### 3.6 N-gram modeli

Bu model, deyimlerin anlamlarını, peşi sıra gelen kelime dizisini analiz ederek belirler. Bu modelin bir örneği bigram veya trigram modelleri olabilir. Örneğin, “*jon taslim qilmoq*” deyiminin trigram modeli <jon>, <taslim>, <qilmoq> kelime dizisini analiz ederek oluşmaktadır. “*vujudining ayoqi ajal balchiqina botmoq*” deyiminin kvadrigram modeli ibareyi teşkil eden <vujud>, <ayoq>, <ajal>, <balchiq>, <botmoq> kelime dizisini analiziyle tespit edilmektedir. Bu nedenle, Bölüm 3.3'te açıklandığı gibi, bir deyim kaç kelimeden oluşursa da, tek bir belirteç (token) olarak tanımlanması, deyimlerin bağlamdan tanımlanmasını kolaylaştırır. Aksi takdirde, bir deyimle aynı birime sahip kelime kombinasyonları metinde deyim olarak kabul edilebilir veya tam tersine ifade olarak tanımlanmayabilir. Örneğin, *qo'lini ko'tarmoq* (elini kaldırmak) sözcük birimi hem bir ifade (*o'rtog'iga qo'lini ko'tardi (urdi)* – arkadaşına elini kaldırdı (vurdu)) hem de basit bir kelime kombinasyonu (*qo'lini tepaga ko'tardi* – elini kaldırdı) olarak kabul edilir. Bu tür eşesli sözcük birimleri, ifadeden önce veya sonra gelen kelimeye göre tanımlanır ve bu durumda N-gram yöntemi kullanılır.

### 3.7 Stilistik üslubu otomatik tespit etme (Classification)

Bunun için 1) üslubu belirlenmiş metinler (Labeled corpus) hazırlanır; 2) deyimlerin bulunduğu bağlamsal cümleler elde edilir; 3) bu cümlelere dayanarak, NLP'de SBERT modelleri kullanılarak sınıflandırma yapılır. Örneğin:

Çizelge IV, SBERT analiz sonuçlarının 6 ifadenin stil özgüllüğünün aşağıdaki göstergelerine sahip olduğunu göstermektedir: Stil özgüllüğü aralığında “vafot etmoq”

deyimi en yüksek aralığa (0,7) sahip kamusal ve ilmi stillere, “jon taslim qilmoq” deyimini en yüksek aralığa (0,8) sahip bediî stile, “halok bo'lmoq” deyimini en yüksek aralığa sahip (0,8) resmi ve kamusal stillere, “ikki qo'li ko'ksida ketmoq” deyimini (0,9) bediî stile, “joniga qasd qilmoq” deyimini en yüksek aralığa (0,9) sahip resmî ve bediî stillere ve “umriga zomin bo'lmoq” deyimini de en yüksek aralığa sahip (0,9) resmî ve kamusal stillere aittir. Bu dağılım, “o'lmoq” köküne sahip altı deyimine dayanarak belirlenmiştir. Bu sonuçlar, Özbek dili ifadelerinin bediî, resmî, kamusal ve ilmi metinlerde kullanıldığını desteklemektedir. Stil-ilişkililik puanı 0,7 ve üzeri göstergelere dayalı SBERT modeli ile başarılı bir şekilde belirlendi.

ÇİZELGE IV. STİLİSTİK SINFLANDIRMA MODELİ İÇİN GEREKLİ VERİLER

Deyimler	Lemma	Resmî	Bediî	İlmî	Kamusal	Sonuç
vafot etmoq	o'lmoq (ölmek)	0,38	0,28	0,52	0,65	Kamusal
jon taslim qilmoq	o'lmoq	0,41	0,72	0,33	0,65	Bediî
halok bo'lmoq	o'lmoq	0,78	0,42	0,69	0,76	Resmî / kamusal
ikki qo'li ko'ksida ketmoq	o'lmoq	0,31	0,82	0,29	0,68	Bediî
joniga qasd qilmoq	o'zini o'ldirmoq (kendini öldürmek)	0,85	0,82	0,65	0,78	Resmî / Bediî
umriga zomin bo'lmoq	o'lmoq	0,85	0,78	0,52	0,82	Resmî / kamusal

#### IV. SONUÇ

Deyimse birimlerin üslup özelliklerinin belirlenmesi, Doğal Dili İşleme (NLP) bakımından güncelliğini koruyan alanlardan biridir. Bu süreç, deyimlerin anlamsal katmanını analiz etmeyi, duygusal, çağrışımsal ve bağlamsal özelliklerini tanımlamayı ve bunları üslup açısından sınıflandırmayı kapsamaktadır.

Semantik tahlil yöntemi kullanılarak, deyimleri eşanlamlı dizisi, duygusal yükü, imge düzeyi ve kullanım bağlamı belirlenir. Sonuç olarak, deyim ait olduğu ifade tarzını (resmî, bedîî, ilmî, kamusal) belirlemek mümkündür. Özbek dilinde deyimlerin ifade tarzının otomatik olarak belirlenmesi (örneğin, “vafot etmoq”, “joniga qasd qilmaq”) ve bunların deyim birimi olup olmadığının belirlenmesi bir NLP sorudur. Bu, özellikle metnin üslup analizi, metin türlerinin sınıflandırılması, ibarelerin resmî veya sanatsal niteliğinin belirlenmesi ve eşanlamlı deyimlerin farklılaştırılması için önemlidir. Makale, SBERT modeline dayalı olarak 0,7 ve üzeri indeksli sonuçlara dayalı olarak 100 ifadenin sanatsal, biçimsel, bilimsel ve kitlesel stillerden birine özgü olduğunu buldu. Sonuçta sözlüklü ve sözlüksüz yöntemler kullanılarak Özbek dilinde toplanan 100 ibare üzerinde yapılan araştırma sonucunda, 25’inin bedîî, 15’inin resmî ve 20’sinin içtimaî nitelik taşıdığı belirlenmiştir.

NLP’de deyim biçimlerinin belirlenmesi şu aşamaları kapsamaktadır: 1) deyim birimlerin otomatik belirlenmesi - istatistiksel (PMI, t-score), kural tabanlı ve yerleştirme yöntemleri kullanılarak; 2) semantik tahlil - eşanlamlı gruplara ve semantik çekirdeğe dayalı deyimlerin anlamsal analizi; 3) üslupsal sınıflandırma - deyimlerin bağlamsal görünümüne dayalı özel olarak eğitilmiş NLP modelleri kullanılarak üsluba göre sınıflandırılması. Böylece, “vafot etmoq” resmî, “jon taslim qilmoq” bedîî-kamusal, “halok bo’lmoq” ilmî, “ikki qo’li ko’ksida ketmoq” bedîî, “joniga qasd qilmoq” resmî veya içtimaî, “umriga zomin bo’lmoq” ise kamusal biçiminin karakteristiği olarak otomatik şekilde tanımlanabilir.

Bu yaklaşımlar, Özbek dili için üslup analizi, otomatik tercüme, metni üsluba göre tahlil etme ve duygusal tonu belirleme gibi görevlerde etkili bir şekilde kullanılmaktadır.

#### KAYNAKÇA

- [1] V.L. Arxangelskiy, “Ustoychivie frazi v sovremennom russkom yazike,” Rostov-na-Donu: izd-vo Rstovskogo un-ta, 1964. – S. 135.
- [2] A.V. Kunin, “Osnovnie ponyatiya frazeologicheskoy stilistiki” // Problemi lingvisticheskoy stilistiki (materiali konferensii). – M.: 1969. – S. 71-75.
- [3] I. A. Sag, T. Baldwin, F. Bond, A. Copestake, & D. Flickinger, “Multiword expressions: A pain in the neck for NLP,” Proceedings of the Third International Conference on Intelligent Text Processing and Computational Linguistics, 2002. pp. 1–15.
- [4] R. Moon, “Fixed Expressions and Idioms in English: A Corpus-Based Approach,” Oxford University Press. 1998.

- [5] A. Fazly, P. Cook, & S. Stevenson, “Unsupervised Type and Token Identification of Idiomatic Expressions,” Computational Linguistics, 35(1), 2009. 61–103.
- [6] G. Salton, R. Ross, & J. Kelleher, “An Empirical Study of the Impact of Idioms on Neural Machine Translation,” Proceedings of the 3rd Workshop on Computational Approaches to Linguistic Code Switching, 2014. pp. 36–45.
- [7] W. Zhou, Q. Liu, & W. Che, “Phrase-level Idiomaticity Detection with Transformer Models,” Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, pp. 4485–4496.
- [8] Q. Liu, & R. Hwa, “Representations of Multiword Expressions for Identifying Idiomaticity,” Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL), 2020. pp. 4229–4240.
- [9] C. Paradis, C. Willners, & D.A. Carrodegua, “Stylistic variation and idiomatic expressions: A corpus-based study of expressive figurative language,” Corpus Linguistics and Linguistic Theory, 18(2), 2022. pp. 327–351.
- [10] Liu, S., Wang, W., & Li, M. (2021). Cross-lingual Multiword Expression Identification with Embedding Transfer and Lexicon Projection. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 2562–2573.
- [11] E. Stamatatos, “A survey of modern authorship attribution methods,” Journal of the American Society for Information Science and Technology, 60(3), 2009. pp. 538–556.
- [12] M. Kestemont, J. Stover, M. Koppel, F. Karsdorp, & W. Daelemans, “Authenticating the writings of Julius Caesar,” Expert Systems with Applications, 80, 2018. pp. 1–10.
- [13] A. Mamatov, Frazelogik stilistika masalalari, Toshkent: Navro’z, 1991. – 120 b.
- [14] A. Mamatov, Ocherki po fransuzskoy frazeologicheskoy stilistiki, Tashkent: O’qituvchi, 1989. – 118 s.
- [15] M. Abjalova, U. Rashidova, A. Hasanov “Creating a phraseological base for the Uzbek↔Turkish automatic translation program,” International Scientific and Practical Conference on Actual Problems of Mathematical Modeling and Information Technology. AIP Conf. Proc. 3147, – pp. 040023-1–040023-5; <https://doi.org/10.1063/5.0210705>
- [16] M. Abjalova, U. Rashidova, “Özbekçe↔Türkçe Otomatik Çeviri Yazılımı İçin Deyimlerin Veritabanını Oluşturma Sorunu,” Journal of Turkic Civilization Studies. – Türkiye, 2024. – S. 178-187.
- [17] M. Abjalova, U. Rashidova, Eşref Adalı, “Özbekçe↔Türkçe Otomatik Çeviri Yazılımı için Deyimlerin Veritabanını Teşkil Etmede Karşılaşılan Güçlükler,” 9th International Conference on Computer Science and Engineering (UBMK), 26-28 October, 2024. Antalya, Türkiye, – pp. 1236-1240.
- [18] M. Abjalova, S. Sharipova, “Semantic and Grammatical Issues in Translating Idioms with Automatic Translation Systems,” 9th International Conference on Computer Science and Engineering (UBMK), 26-28 October, 2024. Antalya, Türkiye, – pp. 64-69.
- [19] M. Sharipov, E. Kuriyozov, O. Yuldashov, O. Sobirov, “Uzbek Verb Detection: Rule-based Detection of Verbs in Uzbek Texts,” International Conference on Computational Linguistics, Language Resources and Evaluation, LREC-COLING 2024 - Main Conference Proceedings, 2024. pp. 17343 – 17347.
- [20] L.I. Royzenon, Frazelogiya v novix zapadnoslavianskix slovaryax // Problemi frazeologii (issledovaniya i material). – M.-L.: Nauka, 1964. – S. 254.
- [21] Sh. Rahmatullayev, O’zbek tilining frazeologik lug’ati. Toshkent: O’qituvchi, 2001. – 407 b.