



UBMK'25

**Bildiriler Kitabı
Proceedings**

Editör 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



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	

10. Uluslararası Bilgisayar Bilimleri ve Mühendisliği Konferansı (UBMK'2025)

10th International Conference on Computer Science and Engineering

17-18-19 Eylül 2025 -İstanbul-Türkiye

17-18-19 September 2025 - İstanbul-Türkiye

Telif Hakkı

Bu elektronik kitabın içinde yer alan tüm bildirilerin telif hakları IEEE'ye devredilmiştir. Bu kitabın tamamı veya herhangi bir kısmı yayıncının izni olmaksızın yayımlanamaz, basılı veya elektronik biçimde çoğaltılamaz. Tersî davranışta bulunanlara ABD Telif Hakkı Yasalarına göre ceza uygulanır.

Copyright and Reprint Permission

Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. Copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For reprint or republication permission, email to IEEE Copyright Manager at pubs-permission@ieee.org

All right reserved. Copyright C 2025

IEEE Catalog Number : CFP25L97-ART

ISBN : 978-8-3315-9975-1

Additional copies may be ordered from:

Curran Associates, Inc

57 Morehouse Lane Red Hook, NY 12571 USA

Phone: (845) 758 0400

Fax: (845) 758 2633

E-mail: curran@proceeding.com

UBMK'2025'ye Hoşgeldiniz

Welcome to UBMK'2025

Sevgili Katılımcılar:

UBMK uluslararası nitelikli konferans serisi, 1990 yılından beri düzenli olarak yapılmakta olan Bilgisayar Mühendisliği Bölüm Başkanları toplantılarında alınan bir kararla on yıl önce başlamıştır. Konferansın 10.su IEEE-UBMK-2025 bu yıl 17-18-19 Eylül, 2025 günlerinde İstanbul Teknik Üniversitesinin ev sahipliğinde düzenlenmiştir.

IEEE-UBMK-2025 konferansına bu yıl Almanya, Amerika Birleşik Devletleri, Azerbaycan, Fransa, Irak, İngiltere, İsveç, İtalya, Kanada, Kazakistan, Kırım, Kırgızistan, Rusya, Özbekistan, Tataristan, Tayland, Ürdün ve Türkiye'den 610 dolayında bildiri gönderilmiş ve bu bildiriler Türk ve yabancı 250 hakem tarafından değerlendirilmiştir.

Her bildiri en az iki hakem tarafından incelenmiş ve uzlaşma olmadığı durumlarda üçüncü bir hakemin değerlendirmesine başvurulmuştur. Bildiri başına düşen ortalama hakemlik 2,3 olmuştur. Bu değerlendirmelerin sonunda 327 bildirinin sözlü olarak sunulması uygun bulunmuştur. Kabul edilen ve sunulan bildiriler içerik ve kalite ölçünlerini sağlaması durumunda IEEE Xplore'da yayımlanacaktır.

Konferans çalışmalarında, Bilgisayar Mühendisliği Bölüm Başkanları Danışma Kurulu olarak görev almışlardır. Bildirilerin değerlendirilmesi Bilim Kurulu üyeleri tarafından yapılmıştır. Konferansın düzenlenmesi ise Yürütme Kurulunun önerileri doğrultusunda, Düzenleme Kurulu tarafından yapılmıştır.

Son olarak, konferansın başarılı bir şekilde yürütülmesi için tüm olanaklarını sunan İstanbul Teknik Üniversitesi Rektörü Sayın Prof. Dr. Hasan Mandal'a teşekkür ediyoruz. Ayrıca Düzenleme Kuruluna, bildirileri titizlikle değerlendiren Bilim Kurulu Üyelerine ve değerli araştırmalarının sonuçlarını bilişim camiası ile paylaşan bildiri sahiplerine teşekkürlerimizi iletiriz.

Prof. Dr. Eşref ADALI
UBMK-2025 Konferans Başkanı ve Bildiri Kitabı Editörü

Dear Participants:

The UBMK international conference series started nine years ago with a decision taken at the Computer Engineering Department Heads (BMBB) meetings, which have been held regularly since 1990. The 10th edition of the conference, UBMK'25, was held this year on October 17-18-19, 2025, hosted by İstanbul Technical University.

This year, approximately 610 papers were submitted to the IEEE-UBMK-2025 conference from Germany, the United States, Azerbaijan, France, Iraq, the United Kingdom, Sweden, Italy, Canada, Kazakhstan, Crimea, Kyrgyzstan, Russia, Uzbekistan, Tatarstan, Thailand, Jordan, and Turkey, and these papers were evaluated by 250 Turkish and foreign referees.

Each paper was evaluated at least by two referees, and in cases where there was no consensus, a third referee was consulted. At the end of these evaluations, 327 papers were accepted for oral presentation. Accepted and presented papers will be submitted for inclusion into IEEE Xplore subject to meeting IEEE Xplore's scope and quality requirements.

During the conference, Heads of Information Engineering Departments took part in the Advisory Board. The evaluation of the papers was made by the members of the Scientific Committee. The conference was organized by the Organizing Committee in line with the recommendations of the Executive Committee.

Finally, we would like to thank İstanbul Technical University Rector Prof. Dr. Hasan Mandal for his continued support for the success of the conference. In addition, we would like to thank the Organizing Committee, the Scientific Committee Members who carefully evaluated the papers, and the owners of the papers who shared the results of their valuable research with the informatics community.

Prof. Dr. Esref ADALI
UBMK'25 Conference Chair and Proceedings Editor

Düzenleyenler / Organizer



İTÜ



Destekleyenler / Sponsors



Özbek Ağızları Derlemi Geliştirme Çalışması (Nurota İlçesi Kıpçak Ağzı Örneğiyle)

Development of an Uzbek Dialects Corpus (With an Example of the Kipchak Dialect from Nurota District)

Khamroeva Shahlo Mirdjonovna
Dept. of Computational Linguistics and
Digital Technologies Tashkent State
University of Uzbek Language and
Literature
Tashkent, Uzbekistan
shaxlo.xamrayeva@navoiy-uni.uz

Muftillayeva Ruxsora
Toshmuhammad qizi
Tashkent State University of Uzbek
Language and Literature
Tashkent, Uzbekistan
ruxsoramuftillayeva@gmail.com

Yuldasheva Dilorom Nigmatovna
Department of Uzbek Linguistics
and Journalism
Bukhara State University
Bukhara, Uzbekistan
d.n.yuldasheva@buxdu.uz

Öz — Bu çalışma, Özbek ağzlarının sistematikleştirilmesi ve dijital teknolojiler aracılığıyla korunmasına odaklanmaktadır. Araştırmada, ağzılara ait bir derlemin oluşturulmasının teorik ve pratik yönleri ele alınmakta olup, Nurota bölgesinin Kıpçak ağzı örnek olay olarak incelenmiştir. Çalışmada, ağız derlemelerinin oluşturulmasında uluslararası deneyimlere yer verilmiş ve Özbekistan'daki mevcut çalışmalar vurgulanmıştır. Sonuç olarak, Kıpçak ağzına dayalı özel bir derlem oluşturulmuş ve bu ağzın temel dil özellikleri sistematik olarak analiz edilmiştir.

Anahtar Sözcükler — Ağız derlemi, ağız bilimi, Kıpçak ağzı, ağız çalışmaları, dijitalleştirme, bölgesel konuşma çeşitleri.

I. GİRİŞ

Günümüz dünya dilbiliminde dijital teknolojilerden yararlanmanın kolaylığı ile yapay zekâ desteğinde dillerin daha mükemmel öğrenilmesi, ağzların korunması ve gelecek nesillere aktarılması meseleleri dilbilimsel açıdan araştırılmaktadır. Günümüz dilbilimcileri ağız veri tabanlarının oluşturulmasına yönelik pratik çalışmalar yürütmektedir. Dünya dilbiliminde, dillerin metin içerisinde muhafaza edilmesi yoluyla yaşatıcılığını artırmak ve ağız derlemelerini oluşturmak güncel konulardan biri olarak kabul edilmektedir. Özbek dilbiliminde, yakın yıllarda yürütülen derlem dilbilimi ile ilgili araştırmalar bu alanı öğrenmek ve incelemek isteyenler için temel kaynak görevini görmekte ve bu alanın gelecek vadeden bir disiplin olduğunu ortaya koymaktadır. Özellikle bilgisayarlı dilbilim, metinlerin sözlükbilimsel işlenmesi ve lingvistik istatistiksel analiz konularında A. Pulatov, S. Muhammedova, S. Karimov, S. Rızaev, A. Norov, M. Ayimbetov, M. Kholova [1] gibi araştırmacıların çalışmaları mevcuttur. Ayrıca, M. Kholova'nın çalışmaları ağız derlemine yönelik pratik adımlar olarak değerlendirilebilir. Bu çalışmada ise biz, ağız derlemelerinin oluşturulmasının dilbilimsel ve uygulamalı yönlerini ele alacağız.

II. KAYNAK ARAŞTIRMASI

Günümüzde dilbilim alanında dillerin modern teknolojiler ve yapay zekâ yardımıyla öğrenilmesi mümkündür. Diyalektlerin hem dilbilimsel hem de uygulamalı açıdan incelenmesi, sonuç olarak ağız birimlerinin korunması ve gelecek nesillere aktarılması konuları araştırılmaktadır. Özbek bilgisayarlı dilbilimi ve derlem bilimi (korpus

dilbilimi) alanında bir dizi çalışma yürütülmektedir. Diyalekt derlemelerinin oluşturulması, büyük hacimli dilsel veri tabanlarının ve çeviri sözlüklerinin hazırlanmasında, Özbekçenin otomatik işlenmesi, makine çevirisi ve derlem tabanlı çeviribilimsel teorik meselelerin derinlemesine incelenmesinde önemli bir rol oynamaktadır [2]. Modern bilgi teknolojileri çağında herhangi bir alanı bilgisayarlar ve yapay zekâ olmadan düşünmek zordur. Bugün, diyalektleri incelemenin modern bir konsepti ortaya çıkmaktadır. Diyalektleri incelemenin bu modern yaklaşımı, birçok yeni sorun ve meseleyi kapsayacak şekilde yeniden yapılandırılmaktadır [3]. Bu bağlamda, ağzıları bir araya getirmek, onların ortak transkripsiyonlarını oluşturmak, ağzıları araştırmak ve sözlüklerini oluşturmayı sağlayacak diyalekt derlemeleri üzerine çalışmalar yapılmaktadır. Ağız derlemeleri konusunda Avusturya, Çin (Mandarin diyalekt derlemi), Macaristan, Almanya, Portekiz, Çekya, Finlandiya, Polonya, Litvanya, Letonya, Gürcistan ve Rusya'nın Moskova, Saratov, Kazan diyalekt derlem okullarında yürütülen araştırmaları örnek göstermek mümkündür. Özellikle Portekiz derlemi, 1972-1974 yıllarında Finlandiya'nın Helsinki bölgesindeki halkın diline dayalı derlem, Corpus Oral y Sonoro del Español Rural (İspanyol derlemi), Bulgarian dialectology as Living Tradition adlı Bulgar derlemi (2.054.705 kelime ve 3400 metinden oluşur), Gürcü derlemi, Arnavut derlemi ve Çin dilinin çeşitliliğini ve değişkenliğini yansıtan Mandarin derlemi dikkat çekicidir. Rusça ağzılara dair derlemeler arasında Saratov diyalekt derlemi, Arhangelsk bölgesinin Ustyia ilçesi ağız derlemi, Rus Ulusal Ağzılarının Elektronik Veritabanı, Rus Ulusal Ağzılarının Elektronik Kütüphanesi gibi birçok derlem sayılabilir. Günümüzde Özbek bilgisayarlı dilbilimi alanında da diyalekt derlem üzerine araştırmalara önem verilmektedir. Çünkü diyalekt derlemelerinin doğal dil işleme (NLP) alanındaki birçok sorunu çözmede etkili olduğu kanıtlanmıştır. Bu nedenle diyalekt derlemeleri oluşturma üzerine bilimsel araştırmalar yapılmaktadır.

Bu konuda özellikle M. Xolova [4], N. Abdurahmonova, S. Muhammedova [5], B. Elov [6] gibi araştırmacılar çalışmalar yürütmektedir. Bu araştırmacılar diyalekt derlemelerini oluşturma bilimsel ve uygulamalı yöntemlerini ortaya koymuşlardır. Örneğin, M. Xolova, Özbekçenin diyalektolojik zenginliğini korumak ve sistematikleştirmek amacıyla Boysun bölgesindeki "j" ünsüzüyle

başlayan ağızlara dayalı ulusal ağızlar derlemi oluşturmada teorik ve uygulamalı-dilbilimsel temellerini geliştirmiştir. Araştırmada halk ağızlarının dijital ortamda saklanması, transkripsiyon ve transliterasyon işaretleriyle derleme dahil edilmesi, tür, konu ve metin biçimlerine göre meta-etiketleme yapılması öngörülmüştür. Ayrıca diyalektolojik haritalar ve interaktif çevrim içi platform (dialect.uz) temelinde modern bilgisayarlı dilbilim yöntemleriyle derlem oluşturma metodolojisi geliştirilmiştir. Bunun yanı sıra, ulusal ağızların sözlükbilimsel, fonetik, morfolojik, anlamsal ve dilbilgisel birimleri derinlemesine analiz edilerek, bunların bilgisayarlı dilbilimi, otomatik çeviri, elektronik sözlükler ve eğitim kaynaklarının geliştirilmesinde kullanılabilirliği bilimsel olarak temellendirilmiştir. B. Elov ise araştırmasında ağızlara dair metinlerin yeterli miktarda toplanması ve doğru transkripsiyon yapılması konularına odaklanmıştır. Ayrıca Özbekçedeki üç büyük lehçeyi analiz etmiştir.

III. ANA BÖLÜM

Milli dil, belirli bir millete ait bireyler için tarihsel olarak ortak olan, milli kültürün aynası niteliğinde onları manevi olarak birleştiren bir araçtır. Bu dil, edebi dil ile birlikte halk ağızlarını da kapsar. Sovyetler döneminde dilimize geniş haklar tanınmamış olsa da, Rus şarkiyatçı bilim insanlarının çabalarıyla Özbek halk ağızları bilimsel temelde incelenmeye başlanmıştır. Bu bilim insanlarının ağızlarla ilgili çalışmaları sonucunda, Özbek edebi dilinden söz edebilmek için mutlaka halk ağızlarının da bilinmesi ve araştırılması gerektiği konusu gündeme gelmiştir.

Rus bilim insanları, 19. yüzyılın 1880'lerinden 20. yüzyılın 1920'li yıllarına kadar Özbek dilini, özellikle halk ağızlarını incelemek için Rusçayı bir araç olarak kullanmışlardır [7]. Ancak tüm bunlara rağmen, Özbek ağızları bağımsızlık döneminden günümüze kadar edebi Özbekçe ile karşılaştırmalı olarak incelenmiş ve bu çalışmalarda tanımlayıcı (deskriptif) ve karşılaştırmalı (komparatif) yöntemler kullanılmıştır. Bu yöntemlerde ağızların fonetik, leksik ve morfolojik özellikleri geniş kapsamda ele alınmış ve zengin bir veri birikimi oluşturulmuştur.

Özbek ağızlarını araştırırken, 19. yüzyılın sonlarında ortaya çıkan karşılaştırmalı-tarihi yöntemden de yararlanılmıştır. Bu yönetime göre Özbek ağızlarının verileri, Divanü Lügati't-Türk, Kutadgu Bilig, Ali Şir Nevaî'nin eserlerinin dili ve diğer tarihî dil kaynaklarıyla karşılaştırılarak incelenmiştir. Ancak bu yöntem, temel yöntem olarak değil, sadece gerekli görülen durumlarda ve bazı dil tarihiyle ilgili makale ve eserlerde kullanılmıştır. Özbek ağızlarını ilk kez tasnif etme görevini Prof. K. K. Yudahin gerçekleştirmiştir. Yudahin, Özbek ağızlarını Tacikçe ile olan ilişkileri ve ses uyumunu (singarmonizm) koruma derecesine göre sınıflandırarak dört gruba ayırır:

Gerçek Türk ses sistemini ve ünlü uyumunu koruyan Özbek ağızları;

Ses yapısını kısmen değiştirmiş, ünlü uyumunu kaybetmiş ağızlar;

İran dillerindeki ünlü sistemini koruyan Özbek ağızları;

Özbek ve Tacik dillerinde konuşulan Tacik ağızları [8].

Özbek halk ağızları üç büyük lehçeye ayrılır: Karluk lehçesi, Kıpçak lehçesi ve Oğuz lehçesi. Bu nedenle Özbek ağızlarının son derece çeşitli ve karmaşık bir yapıya sahip

olduğu uzun zamandır bilinmektedir. Yukarıda verilen sınıflandırmanın bazı edebî kaynaklarda küçük farklılıklarla verildiği görülmektedir. Ağızlardaki kelime birimleri, dilbilgisel ekler ve bazı durumlarda cümle yapısı hem edebi dilden hem de birbirlerinden farklılık gösterebilir. Bunun örneğini günlük hayatta kullanılan basit kelime biçimlerinde dahi görmek mümkündür. Örneğin, "ona" (anne) kelimesi farklı ağızlarda çeşitli şekillerde kullanılabilir.

ÇİZELGE I. EDEBİ DİLDEKİ SÖZCÜKLERİN FARKLI AĞIZLARDAKİ KULLANIMI

Adabiy til/sheva	Leksik birlik
Adabiy til	Ona
Buxoro	Nana
Navoiy/Qashqadaryo/Farg'ona	Aya
Navoiy	Ena
Toshkent	Oyi/opa
Xorazm	Opo/opa
Surxondaryo	Buva
Namangan	Aba

Kıpçak lehçesi, diğer lehçelere kıyasla daha geniş coğrafyalara yayılmıştır. Özbekistan'ın tüm vilayetlerinde bu lehçeye mensup konuşur toplulukların bulunması da bu görüşün doğruluğunu kanıtlamaktadır. Kıpçak şivesi, tarihî Kıpçak dilinin etkisiyle şekillenmiştir. Kıpçak dili ise Kıpçak boylarının dili olarak kabul edilir ve Türk dilleri grubuna dâhildir. Kıpçak şivesi yalnızca diyalektolog dilbilimciler için değil, aynı zamanda halkbilim (folklor) araştırmacıları için de son derece önemli bir şive olarak kabul edilmektedir. Çünkü tam da bu şivede halkın sözlü kültürüne ait çok değerli destanlar yaratılmıştır.

Dünya dilbiliminde diyalekt (ağız) derlemeleri oluşturulurken, bu derlemelerin gramatik, leksik ve fonetik yönlerine de özel önem verilmektedir. Diyalektal derlemelerin hazırlanmasında transkripsiyon konusuna özellikle dikkat edilmesi gerekir. Transkripsiyon, fonemlerin konuşma akışındaki çeşitli nüanslarını; bölgesel ağız ve lehçelerin fonetik özelliklerini (seslerin kısalması veya uzaması, kök ve eklerin uyumlu değişimi gibi) yansıtan özel bir yazı sistemidir. Bu nedenle, transkripsiyon işaretlerinin sayıca alfabenin seslerinden fazla olması doğaldır. Günümüzde transkripsiyon örnekleri hem Kiril hem de Latin alfabesiyle yazılmış biçimlerde dilbilimsel çalışmalarda yer almaktadır. Özbek dilbiliminin şekillenmeye başladığı ilk dönemlerden itibaren transkripsiyona olan ihtiyaç hissedilmiş, hatta bu meseleye 1920'li yılların devlet adamları da dikkat çekmişlerdir. Özellikle Prof. Y.D. Polivanov, Özbek ağızlarını yazıya geçirirken Latin alfabesine dayalı uluslararası transkripsiyon sistemini kullanmaya özen göstermiş, ancak zaman zaman Kiril sistemine de başvurmuştur.

ÇİZELGE II. POLIVANOV TRANSKRİPSİYONUNDA KULLANILAN TEMEL FONETİK İŞARETLER

	Türkistan Ağızı	İkan Ağızı	Ses Sırası Fonem	Türkistan Ağızı	İkan Ağızı
Ön sıra	ı	i	Ön sıra	ь	u

Ön sıra	y	y	Ön sıra	u	u
Ön sıra	e	e	Ön sıra	o	o
Ön sıra	ø	ø	Ön sıra	a	a
Ön sıra	æ	ε			

G'ozı Olim Yunus da transkripsiyon konusunda bir dizi bilimsel araştırma yürütmüştür. Bilim insanı ünlü ve ünsüz sesleri üç farklı alfabe sistemi kullanarak açıklamaktadır. Transkripsiyondaki bu çeşitlilik, diğer diyalektoloji çalışmalarında da gözlemlenmektedir. Transkripsiyondaki bu durum, her yönüyle bilimsel temellere dayanan bir transkripsiyon sisteminin oluşturulmasını zorunlu hale getirmiştir. N. Rajabov da bilimsel çalışmalarında şu görüşleri öne sürmüştür:

1. Transkripsiyonda yazım kurallarına değil, kelimelerin telaffuz özelliklerine dikkat edilmelidir. Ancak bu kural henüz gereğince uygulanmamaktadır.

2. Ona göre, transkripsiyonda büyük harfler kullanılmaz. Bu görüşte doğruluk payı vardır, çünkü transkripsiyon işaretleri genellikle küçük harflerle gösterilir; ancak metin yazımında cümle başında ve özel isimlerde bu işaretlerin büyük harfleri kullanılması geleneği de vardır.

3. Transkripsiyonda her harf bir sesi ifade etmelidir. Bu görüş yerindedir, ancak bu kural zaman zaman ihlal edilmiştir. Burada biz bu kuralların olumlu karşılanan yönlerini dikkate alıyoruz.

Bu tür bir transkripsiyon oluşturma çalışmalarını V. Reshetov başlatmıştır. Bilim insanı, Kiril alfabesine dayanarak Özbek ağızları için kullanılan bir transkripsiyon geliştirmiştir. V. V. Reshetov'un transkripsiyonunda geçen f1, r1, k1, l1, h1, h1k, тш, x1, kx gibi işaretler ise araştırmalarda kullanılmamıştır. Çünkü bu işaretler sertlik ve yumuşaklığı ifade etmek için kullanılsa da, bu durum tek yerli ünsüzler dışında tüm ünsüzlerde mevcut olduğundan aslında bu tür işaretlere gerek kalmamıştır [9].

Bağımsızlıktan sonra Latin alfabesine geçilmesiyle birlikte, yeni Latin alfabesi temelinde bir transkripsiyon sistemi oluşturma ihtiyacı doğmuştur. Ancak bu görev biraz gecikmiştir, çünkü ağızlarımızla ilgili oluşturulan büyük miras Kiril alfabesine dayalı transkripsiyonla yazılmış ve yayımlanmıştır. Latin ve Kiril alfabesine dayalı transkripsiyon sistemi 2016 yılında S. Ashirboyev tarafından öğrencilere sunulmuştur.

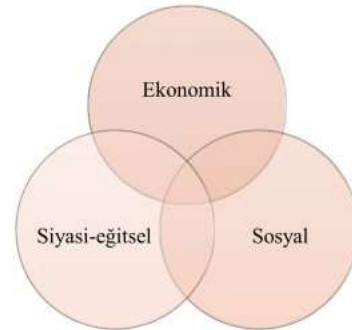
Diyalektal korpusun pasaport metaetiketi bu korpusta yer alan bilgi topluluğunu tanımlayan bir belgedir. Bu belge, dilbilimsel kaynakların doğru yönetilmesine hizmet eder. Herhangi bir ağıza ait metinler üç yönlü meta-metinli tanımlayıcı bir yapıya sahip olur. İlk olarak şu bilgiler yazılır: kaydedildiği yer (şehir, ilçe, köy), kaydedildiği zaman (yıl, ay, gün), bilgi veren kişinin eğitimi (yükseköğrenim görmüş, eğitimsiz), bilgi verenin mesleği, veriyi toplayan kişi (soyadı, adı, babasının adı), konu (tarım, el sanatları), metin biçimi (sözlü) gibi bilgiler kaydedilir. Yukarıdaki parametrelere göre bir arama yapıldığında şu tür sonuçlara ulaşılır.

ÇİZELGE III. AĞIZLARA AİT PASAPORT METAETİKETİ

T/R	Anket Bilgisi	Cevaplar
1.	Kaydedildiği yer (şehir, ilçe, köy)	Yangihayot Mahallesi, Nurota İlçesi, Navoi Vilayeti

2.	Kaydedildiği tarih (gün, ay, yıl)	Temmuz 2023
3.	Bilgi verenin adı, soyadı, baba adı	Hakimova Qontigul Polvonovna
4.	Bilgi verenin cinsiyeti	Kadın
5.	Bilgi verenin doğum yeri ve yılı	Nurota İlçesi, Yangihayot Mahallesi
6.	Bilgi verenin eğitimi	Ortaöğrenim
7.	Bilgi verenin mesleği	Ev hanımı
8.	Konu	Çocukluk anıları
9.	Metin türü	Sözlü anlatım

Diyalektal korpuslar için toplanan metinlerde konusal farklılıkların olduğu gözlemlenebilir. Ağızlardaki konusal ayırım, halkın sosyal yaşamda tuttuğu yerle ilgilidir; çünkü hayvancılık, konut, meslek ve el sanatları, inanç, din gibi kavramlar ağız temsilcisinin dünya görüşünde aynı derecede yer tutmaz. Belirli bir kavramın ön planda olması, yerel halkın yaşam tarzı ve uğraştığı alanla bağlantılı bir durumdur. Diyalektal birimlerde genel konular olarak şunlar gösterilebilir: Törenler, kişisel yaşamla ilgili kavramlar, inanç, el sanatları (el emeği) gibi. Bu makro konular kendi içinde aşağıdaki mikro konulara ayrılabilir: Bireyin kişisel hayatı: anne-baba, çocuk, akrabalar, köylüler, El sanatları: halıcılık, el sanatlarına ait nesneler, vb. Konuşmalarda belirlenen konulara göre hazırlanacak metaetiket, ağız konuşmasının kendine has özelliklerini, köy halkının kavramsal alanını araştırmak açısından önemlidir; ayrıca ağız sözlükleri için sözcük seçiminde de faydalı olabilir. Diyalektal metinlerin dilbilgisel analizi ise belli zorluklar yaratmaktadır. Günümüzde Türk dilleri için bir morfolojik standart geliştirilmiştir [10]. Ağızlar korpusu konusal olarak üçe ayrılır:



Şekil-1. Ağızlara Ait Konu Metaetiketi

Ağız korpusu için toplanan metinlerde en sık rastlanan konuşma türleri şunlardır:

Bilgi veren kişinin yaşam öyküsünü içeren, resmi olmayan biçimde anlatılan otobiyografik hikâye,

Hatıra – geçmişte yaşanmış olayların anlatıldığı monolog şeklindeki hikâye,

Başka kişiler hakkında anlatılan hikâyeler,

Yakınlarının bizzat yaşadığı ilginç ve sıra dışı olaylar hakkında bilgi veren olay anlatımı.

Milli dili ve onun ağızlarını bilgi dünyasına taşımak, çağdaş araştırmalar için önemli bir temel oluşturur. Bu süreç yalnızca dilbilim açısından değil, aynı zamanda kültürel çalışmalar ve antropoloji alanlarında da yeni imkânlar sunar. İlk ağız veri tabanının oluşturulması süreci, bir dizi önemli gereksinim ve hedefler doğrultusunda yürütülmektedir. Bu veri tabanının temel amacı, Özbek milli ağızlarının çeşitliliğini ve onların kendine özgü leksik ve dilbilgisel özelliklerini derinlemesine incelemektir. Bu amaç doğrultusunda şu temel ilkeler dikkate alınır:

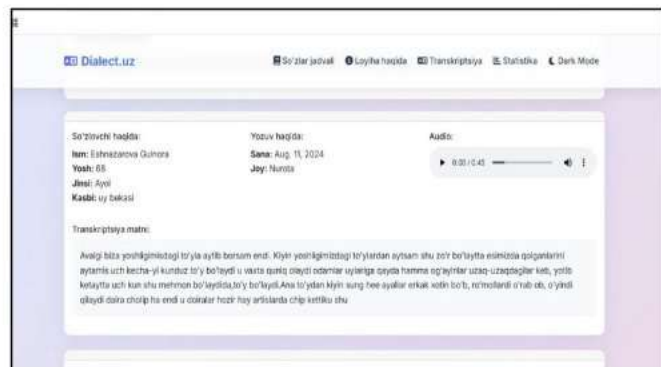
ÇİZELGE IV. AĞIZ VERİ TABANINDA AĞIZA AIT KELİMELERİN SUNUMU

Ağıza Ait Kelime	Shipirgi
Edebî Dildeki Karşılığı	Supurgi
Türü	İsim
Kelimenin Açıklaması	Süpürme, temizleme amacıyla kullanılan; çalı, pelin otu, mısır süpürgesi gibi bitkilerden demetlenerek yapılan bir ev eşyasıdır.

Bu metodolojik yaklaşımlar sayesinde anket sonuçları daha net ve doğru hale gelir. Farklı diyalekt kelimelerinin kullanılması, katılımcıların farklı sosyal ve kültürel bağlamlarını yansıtmaya açısından büyük önem taşır. Bu da araştırmacılara daha güvenilir ve isabetli sonuçlar çıkarma imkânı sunar. Bir kelimenin ankette farklı şekillerde sunulması, kaynağın daha geniş kapsamda kullanılmasına olanak tanır. Nurota ağızlarının elektronik sözlüğü, internette çevrimiçi çalışan bir veri tabanıdır. Bu veri tabanı, içeriğini genişletme özelliğine sahip, kapsamlı ve kullanıcı dostu diyalekt veri tabanlarından biridir. Bu elektronik veri tabanı geniş olanaklara sahiptir ve kullanıcı kitlesi için aşağıdaki imkânları sunar:

1. Ana Sayfa
2. Kelime Tablosu
3. Proje Hakkında
4. Transkripsiyon
5. İstatistikler
6. İletişim <https://nurota-dialect.uz/transkript/>

Diyalekt konuşma kayıtlarını saklamak ve analiz etmek amacıyla geliştirilen “Recording” modeli ve “DialectWord” modeli incelenmiştir. Bu modeller aracılığıyla ses kayıtları, konuşmacı bilgileri, kelimelerin bölgesel varyantları, transkripsiyonları ve edebî Özbekçedeki karşılıkları gibi bilgiler saklanmaktadır. Recording modeli, ses dosyalarını ve onların transkripsiyonlarını kaydetme imkânı sunar. Aynı zamanda konuşmacının yaş, cinsiyet ve meslek gibi kişisel bilgilerini de içerir. DialectWord modeli ise kelimelerin sözcük türlerini, bölgesel varyasyonlarını ve Özbekçedeki eşdeğerlerini kayıt altına almak için kullanılır. Bu modellerin birlikte çalışması, diyalekt konuşmalarının sistematik biçimde düzenlenmesi ve bilimsel analiz edilmesi açısından büyük kolaylık sağlar. Recording modeli, diyalektal ses kayıtlarını arşivlemek, analiz etmek ve kullanıcı bilgilerini toplamak için geliştirilmiştir.



Şekil-2. Korpusta Sesli Metinlerin Sunumu

```
class Recording(models.Model):
    dialect_word = models.ForeignKey(
        DialectWord,
        on_delete=models.CASCADE,
        related_name='recordings',
        null=True,
        blank=True
    )
    audio_file = models.FileField(upload_to='recordings/')
    speaker_name = models.CharField(max_length=100,
        default='Unknown')
    speaker_age = models.IntegerField(default=0)
    speaker_gender = models.CharField(
        max_length=10,
        choices=[
            ('male', 'Erkak'),
            ('female', 'Ayol')
        ],
        default='male'
    )
    speaker_occupation =
models.CharField(max_length=100, default='Unknown')
    transcription_text = models.TextField(
        help_text="Audio matni transkripsiyasi",
        default=""
    )
    recorded_date = models.DateField(default=date.today)
    location = models.CharField(
        max_length=200,
        help_text="Yozib olingan joy",
        default='Nurota'
    )
    notes = models.TextField(
        blank=True,
        help_text="Qo'shimcha ma'lumotlar"
    )

    def __str__(self):
        if self.dialect_word:
            return f'Recording of {self.dialect_word.word} by {self.speaker_name}'
        return f'Recording by {self.speaker_name}'
```

```
class Meta:
    ordering = ['-recorded_date']
```

Recording Modeli, diyalekt arařtırmaları için toplanan sesli verilerin düzenlenmesi, saklanması ve daha sonra analiz

edilmesi imkânını sağlar. Bu model, ağızlara ait ses kayıtlarının sistematik şekilde arřıvlenmesi için yararlı bir yapıdır. Ařağıdaki tabloda, Recording modeli içindeki her alan (field) ve onun korpustaki işlevi (rolü) tanımlanmıştır:

ÇİZELGE V. RECORDING MODELİNDEKİ ALANLARIN AÇIKLAMALARI

Alan Adı	Veri Türü	Açıklama
dialect_word	ForeignKey	Yazıya ait diyalekt kelimesini belirlemek için kullanılır
audio_file	FileField	Ses kaydını saklama ve yükleme imkanı sağlar
speaker_name	CharField	Konuşmacının adını saklar
speaker_age	IntegerField	Konuşmacının yaşını belirtir
speaker_gender	CharField (choices)	onuşmacının cinsiyetini (erkek veya kadın) belirtir
speaker_occupation	CharField	Konuşmacının mesleğı hakkında bilgi verir
transcription_text	TextField	Ses kaydının yazılı hali (transkripsiyon)
recorded_date	DateField (default=date.today)	Kayıt tarihini belirtir
location	CharField	Kayıt yapılan yer hakkında bilgi verir
notes	TextField (blank=True)	Ek açıklamalar için alan

Bu model, dilbilimsel arařtırmalar ve diyalektal analizler için önemli bir değere sahiptir. Geniş kullanım olanaklarıyla farklı alanlarda etkili şekilde uygulanabilir. Özellikle Özbek dilinin bölgesel özelliklerini derinlemesine inceleme, sesli materyalleri toplama ve bilimsel verileri sistematik olarak saklama imkanı sunar. Veri tabanında diyalekt kelimeleri ve onların ses kayıtları muhafaza edilir. Recording modelindeki audio_file alanı, ses dosyalarının saklanması ve analiz edilmesini sağlar. Django kullanılarak medya dosyalarının yüklenmesi ve işlenmesi ise ařağıdaki şekilde gerçekleştirilir.

```
recording = Recording.objects.create(
    speaker_name="Ali",
    speaker_age=30,
    speaker_gender="male",
    audio_file="recordings/sample.mp3"
)
```

Bu model, diyalektal konuşmayı dijitalleştirme, saklama ve bilimsel analiz yapma imkanı sunar. Django modelleri kullanılarak etkin ve sistematik bir veri tabanı oluşturmak mümkündür. Sonuç olarak, yukarıda belirtilen aşamalar doğrultusunda kapsamlı ve işlevsel bir şive korusu oluşturma algoritması geliştirilebilir.



Şekil-3. Diyalekt (şive) korusu oluşturma algoritması

IV. SONUÇ VE ÖNERİLER

Sonuç olarak şunu belirtmek gerekir ki, diyalektal korusu oluşturmak yalnızca bilgisayar dilbilimi için değil, aynı zamanda diyalektoloji, folklor bilimi ve sosyodilbilim gibi alanlar için de gerekli veri ve kaynak sağlar. Diyalektal korusu için veri toplanması sayesinde ağıza ait birimler daha derinlemesine incelenmiş ve analiz edilmiştir. Özellikle bazı kelimelerin sadece yaşlıların konuşmasında korunduğı gözlemlenmiştir. Bazı kelimelerin yalnızca belirli bölgelerde kullanılması Özbekçenin çeşitliliğini gösterir. Korpustaki kelimelerin transkripsiyonu, kelimelerin daha anlaşılır olmasını sağlar. Nurota ağızlar korusunda kadınlar ve erkeklerin konuşmalarında en aktif kullanılan birimler belirlenmiştir. Ayrıca, en aktif 10 kelimenin istatistikleri çıkarılmıştır. Ses dosyalarının türlere göre ayrılması, halk arasında hangi tür metinlerin sözlü anlatımda kullanıldığını göstermiştir. Özellikle yaşlıların dilinde daha çok bilgi verme, hikaye anlatma, açıklama ve analiz gibi anılarla ilgili sohbetlerin hakim olduğu tespit edilmiştir.

KAYNAKÇA

- [1] A. ПўлатовКомпьютер лингвистикаси. – Тошкент: Академнашр, 2011. – 175 Б.
- [2] B. Elov, O. Abdullayeva Dialektal korpusni qurish va uning ahamiyati // Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti o'zbek shevashunosligi markazi. 2024. № 4. B. 56-57.
- [3] R. Amery Phoenix or Relic? Documentation of Languages with Revitalization in Mind. Language Documentation and Conservation, 2009. 3(2): 138-148.
- [4] M. A. Xolova Transcription in the corpus of the uzbek national Boysun dialect (on The example of baysun district “j” dialects). // Journal of Critical Reviews. 2020. Vol.№ 7. Is.№ 5. DOI:<http://dx.doi.org/10.31838/jcr.07>. –Pg. 844-849.
- [5] C. Мухаммедова Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. Методик қўлланма. – Тошкент, 2006. – 156 Б.
- [6] B. Elov O'zbek tili shevalari korpusi arxitekturasini o'zbek folklori va shevalari tadqiqotlari: amaliyot, metodologiya, yangicha yondashuv. 2023. № 4. B. 41-46.
- [7] S. Ashirboyev, Z.R. Xidraliyeva O'zbek dialektologiyasi ta'limi: Yangilanish tamoyillari (xorij tajribasini milliy lashtirish) // Pedagogika. 2017, 1-son. – B. 109-114.
- [8] S. Ashirboyev O'zbek dialektologiyasi. – Toshkent: Nodirabegim, 2021. – B. 34.
- [9] S. Ashirboyev O'zbek milliy transkripsiyasini yaratish yo'lida. // Oltin bitiglar 2024. № 1. – B 12-13.
- [10] S.Khakimov, F.Salimov, D. Ramzanova Building dialectological corpora for Turkic languages: Mishar dialect of Tatar, 2015. – P. 221.