Dialektlər Əsasında Fərdlərin İdentifikasiyası Problemlərinin Araşdırılması və Təhlili

Əfruz Qurbanova[1], Səriyə Qasımova[2]
[1]İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan
[2]Naxçıvan Dövlət Universiteti, Naxçıvan, Azərbaycan
[1]*afruz1961@gmail.com*,[2]*qasimovaseriyye@ndu.edu.az*

***Xülasə*— Məqalədə dialektlər əsasında fərdlərin identifikasiyası problemi analitik ədəbiyyat icmalı əsasında təhlil edilir. Göstərilir ki, bu problem dialektologiya, məhkəmə dilçiliyi, nitq texnologiyaları, korpus dilçiliyi və təbii dilin emalı sahələrinin kəsişməsində formalaşan mürəkkəb tədqiqat istiqamətidir. Məqalədə nitq siqnalları əsasında danışanın profilləşdirilməsi, regional mənsubiyyətin müəyyənləşdirilməsi, aksentin tanınması və məhkəmə-fonetik müqayisə yanaşmaları, həmçinin mətnlər əsasında müəllifin profilləşdirilməsi, regional müəlliflik markerlərinin çıxarılması və idiolektin təhlili müqayisəli şəkildə nəzərdən keçirilir. Təhlil göstərir ki, dialekt və aksent göstəriciləri birbaşa və qəti şəxs identifikasiyasından daha çox, fərdin regional, sosial və kommunikativ profilinin ehtimal əsaslı müəyyənləşdirilməsi üçün elmi baxımdan daha əsaslı imkanlar yaradır. Məqalədə Azərbaycan dili üzrə mövcud dialektoloji və dialektometrik resursların bu istiqamətdə tətbiq imkanları dəyərləndirilir, başlıca metodoloji və praktik çətinliklər müəyyənləşdirilir və mərhələli inkişaf istiqamətləri təklif olunur.**

Açar sözlər*— *dialekt; fərdin identifikasiyası; profilləmə; idiolekt; məhkəmə dilçiliyi; təbii dilin emalı.

I. GİRİŞ

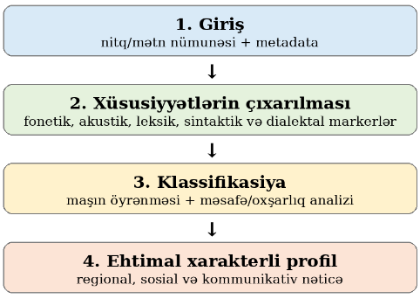
Rəqəmsal kommunikasiya mühitinin genişlənməsi, sosial media platformalarının sürətli inkişafı və nitq emalı texnologiyalarının təkmilləşməsi dil məlumatlarının həm həcmini, həm də tətbiq sahələrini əhəmiyyətli dərəcədə artırmışdır. Belə şəraitdə dil təkcə ünsiyyət vasitəsi deyil, həm də fərd haqqında dolayı məlumat daşıyan identifikasiya mənbəyi kimi aktuallaşmışdır. Xüsusilə dialekt, aksent, regional dil xüsusiyyətləri və fərdi üslub göstəriciləri danışanın və ya müəllifin regional mənsubiyyəti, sosial mühit və kommunikativ davranışı haqqında mühüm məlumatlar verə bilir [1, 2].

Məhkəmə dilçiliyi və məhkəmə fonetikası sahəsində bu problem daha çox danışanın profilləşdirilməsi və nitq nümunələrinin müqayisəsi prizmasından araşdırılır. Burada məqsəd şəxsiyyətin birbaşa və qəti müəyyənləşdirilməsi deyil, anonim nitq və ya mətn nümunələri əsasında onun mümkün regional, sosial və qrup mənsubiyyəti haqqında ehtimal əsaslı nəticələr əldə etməkdir [1]. Son illərdə müəllifin profilləşdirilməsi, idiolektin modelləşdirilməsi və regional müəlliflik xüsusiyyətlərinin çıxarılması ilə bağlı mətnyönlü tədqiqatlar da bu sahəyə yeni məzmun qazandırmışdır [3, 4].

Azərbaycan dili üçün bu problem xüsusi aktuallıq daşıyır. Azərbaycan dilinin zəngin dialekt sistemi, dialektoloji atlasları və dialektometrik yanaşmaları regional variasiyanın sistemli təhlili üçün əlverişli baza yaradır. Bununla belə, bu imkanların korpuslar, hesablama alətləri və tətbiqi modellərlə əlaqələndirilməsi hələ kifayət qədər inkişaf etməmişdir [5].

Bu məqalə analitik ədəbiyyat icmalı xarakteri daşıyır. Məqsəd dialektlər əsasında fərdlərin identifikasiyası problemini nitq və mətnəsaslı tədqiqatlar əsasında müqayisəli şəkildə təhlil etmək, əsas elmi-metodoloji çətinlikləri müəyyənləşdirmək və Azərbaycan dili üçün tətbiq perspektivlərini göstərməkdir. Bu məqalədə "identifikasiya" anlayışı dar kriminalistik mənada şəxsiyyətin qəti müəyyənləşdirilməsi kimi deyil, dialekt və aksent göstəriciləri əsasında fərdin regional, sosial və kommunikativ profilinin ehtimal əsaslı müəyyənləşdirilməsi kimi başa düşülür. Bu baxımdan məqalənin əsas tezisi ondan ibarətdir ki, dialektlər əsasında birbaşa fərd identifikasiyasından daha çox regional və sosial profilləmə elmi baxımdan daha real və daha əsaslandırılmış yanaşmadır.

Dialekt əsaslı identifikasiya və profilləmə prosesinin ümumi metodoloji ardıcıllığı aşağıdakı blok-sxemdə göstərilmişdir. Bu ardıcıllıq məqalədə araşdırılan nitq və mətn əsaslı yanaşmaları vahid profilləmə məntiqində birləşdirir.



Şəkil 1. Dialekt əsaslı identifikasiya və profilləmə prosesinin blok-sxemi

Blok-sxemdən göründüyü kimi, dialekt əsaslı identifikasiya birbaşa və qəti şəxs tanıma prosesi kimi deyil, dil göstəricilərinin çıxarılması, təsnif edilməsi və nəticədə ehtimal əsaslı regional-sosial profilin qurulması kimi başa düşülür.

II. ƏDƏBİYYAT İCMALI

Müasir tədqiqatlarda dialektlər əsasında fərdlərin identifikasiyası iki əsas istiqamətdə inkişaf edir. Birinci istiqamət nitq siqnalları əsasında danışanın profilləşdirilməsi, regional mənsubiyyətin müəyyənləşdirilməsi və aksentin tanınması ilə bağlıdır. İkinci istiqamət isə mətnlər əsasında

DOI: 10.25045/forensic.gov.2026.46

müəllifin profilləşdirilməsi, regional müəlliflik markerlərinin çıxarılması və idiolektin təhlili ilə əlaqədardır.

Nitq əsaslı tədqiqatlar göstərir ki, dialekt və aksent fərqləri səs siqnalında müşahidə olunan fonetik və akustik xüsusiyyətlər vasitəsilə aşkarlana bilir. Gujarati dialektləri üzrə aparılmış araşdırmada sait keyfiyyəti, intonasiya, ton və digər supraseqmental əlamətlərin danışanın regional aksentinin müəyyənləşdirilməsində və məhkəmə məqsədli profilin təhlilində faydalı olduğu göstərilmişdir [6]. Alman dilində telefon danışıqları əsasında aparılmış araşdırmada regional aksentin məhkəmə-fonetik təhlildə mühüm rol oynadığı qeyd edilir [7]. Türk dili materialı üzərində qurulmuş tədqiqatda isə formant göstəriciləri vasitəsilə regional aksent və mənşəyin müəyyənləşdirilməsi imkanları nümayiş etdirilmişdir [8]. Bu işlər göstərir ki, fonetik və akustik markerlər danışanın regional profili barədə etibarlı ehtimallar formalaşdırmağa imkan verir.

Bu istiqamətdə müasir hesablama üsullarının tətbiqi nəticələri daha da gücləndirmişdir. Aksent və dil məlumatlarının qısa nitq parçalarından çıxarılmasına əsaslanan yanaşmaların məhkəmə məqsədli danışanın tanınmasının dəqiqliyini artırdığı göstərilmişdir [9]. Eyni zamanda, aksentin avtomatik tanınmasında fonetik-seqmental məlumata əsaslanan yanaşmaların daha izahlı və praktik baxımdan daha münasib olduğu vurğulanmışdır [10]. Qısa zaman ərzində çıxarılan sait əsaslı akustik göstəricilərin də fərqləndirici potensialı ayrıca qeyd edilmişdir [11].

Mətn əsaslı tədqiqatlar isə fərdin dil davranışının başqa bir ölçüsünü üzə çıxarır. Müəllifin profilləşdirilməsi və müəllifliyin müəyyənləşdirilməsi sahəsində aparılmış araşdırmalarda göstərilir ki, leksik göstəricilərlə yanaşı, sintaktik və diskurs xüsusiyyətləri də yüksək ayırdetmə gücünə malikdir [12]. İdiolektin öyrənilməsi ilə bağlı müasir işlərdə fərdi yazı üslubunun həm fərqləndirici, həm də müəyyən qədər sabit olduğu əsaslandırılır [4]. Regional müəlliflik xüsusiyyətlərinin müəyyənləşdirilməsi sahəsində geoişarələnmiş sosial media məlumatları əsasında aparılan araşdırmalar göstərir ki, böyük həcmli korpuslar və məkan statistikası vasitəsilə regional markerlər daha obyektiv üsullarla aşkarlana bilər [3]. Yazılı Amerika ingiliscəsində regional variasiya üzrə aparılmış korpus əsaslı tədqiqat da bu yanaşmanın nəzəri və metodoloji baxımdan səmərəli olduğunu təsdiq edir [13].

Azərbaycan dili kontekstində dialektometriya ilə bağlı aparılmış tədqiqatlar bu problem üçün mühüm metodoloji baza yaradır. Dialektoloji atlas məlumatlarının rəqəmsallaşdırılması, verilənlər bazasına daxil edilməsi və statistik-hesablama üsulları ilə təhlili regional variasiyanın sistemli öyrənilməsi üçün vacib istiqamət kimi qiymətləndirilmişdir [5].

Beləliklə, ədəbiyyatın təhlili göstərir ki, nitqəsaslı tədqiqatlar bu sahədə daha yetkin metodoloji baza formalaşdırmış, mətnəsaslı tədqiqatlar isə son illərdə idiolekt və regional müəlliflik markerləri üzərində daha sürətlə inkişaf etməyə başlamışdır. Bu isə problemin kompleks təhlili üçün hər iki istiqamətin birlikdə nəzərdən keçirilməsini zəruri edir.

III. Dialekt, Aksent və İdiolektin İdentifikasiya Baxımından Nəzəri Əsasları

Dialekt müəyyən coğrafi və ya sosial qrupa məxsus olan dil variasiyasıdır. Aksent əsasən tələffüz, intonasiya və fonetik xüsusiyyətlərlə bağlı olub danışanın regional və ya sosial mənsubiyyətinə işarə edir. İdiolekt isə konkret fərdin dil istifadəsinə xas olan fərdi xüsusiyyətlər sistemidir. Bu anlayışlar birlikdə identifikasiya probleminin nəzəri çərçivəsini təşkil edir.

Dialekt və aksent əsasən qrup səviyyəsində fərqləndirici əlamətlər yaratsa da, idiolekt fərdi səviyyədə ayırdetmə imkanı verir. Lakin praktik tədqiqatlarda bu səviyyələr həmişə aydın şəkildə ayrılmır. Danışanın fərdi dili regional fondan, regional fond isə sosial amillərdən təsirlənir. Buna görə də müasir araşdırmalarda bu problem birbaşa identifikasiya kimi deyil, profilləmə və ehtimal yönümlü xarakteristika çıxarılması kimi izah olunur [1, 2, 4]. Məhz bu səbəbdən dialekt, aksent və idiolekt səviyyələrinin qarşılıqlı təsiri bu sahədə həm elmi imkanları, həm də metodoloji məhdudiyyətləri müəyyən edən əsas nəzəri problem hesab olunmalıdır.

IV. Nitq Əsasında Dialekt Yönümlü İdentifikasiya Yanaşmaları

4.1. Fonetik və akustik markerlər

Nitq əsaslı yanaşmalar dialektlər əsasında fərdlərin identifikasiyası sahəsində ən geniş işlənmiş istiqamətdir. Bu yanaşmalarda məqsəd səs siqnalından istifadə etməklə danışanın aksentini, regional mənsubiyyətini və müəyyən hallarda qrup mənsubiyyəti profilini çıxarmaqdır.

Aparılmış tədqiqatlar göstərir ki, regional aksent və dialekt fərqləri səsin seqmental və supraseqmental xüsusiyyətlərində aydın şəkildə əks olunur. Sait keyfiyyəti, formant göstəriciləri, intonasiya, ton, səs siqnalının müddəti və intensivliyi kimi əlamətlər danışanın mənsub olduğu regional dil mühiti haqqında informativ nəticələr verir [6, 7, 8].

4.2. Hesablama və təsnifat yanaşmaları

Nitq əsaslı identifikasiya yanaşmalarında ən çox istifadə olunan göstəricilərə Mel tezlik kepstral əmsalları, formantlar, spektral göstəricilər, ton, enerji və müddət daxildir [6, 9, 11]. Bunlar həm ənənəvi statistik təsnifat üsullarında, həm də müasir maşın öyrənməsi modellərində tətbiq olunur.

Saleem və həmmüəlliflərin tədqiqatı göstərir ki, aksent təsnifatı və dilin müəyyənləşdirilməsi haqqında məlumatların eyni modeldə birləşdirilməsi danışanın tanınması nəticələrini əhəmiyyətli şəkildə yaxşılaşdıra bilər [9]. Brown və həmmüəlliflər aksentin avtomatik tanınmasında fonetik-seqmental məlumatın istifadəsinin daha izahlı və tətbiq baxımından daha etibarlı olduğunu vurğulamışlar [10]. Humayun və həmmüəlliflər isə qısa zaman pəncərələrində çıxarılan sait əsaslı akustik xüsusiyyətlərin fərqləndirici potensialını xüsusi qeyd etmişlər [11].

4.3. Məhkəmə kontekstində tətbiq imkanları

Məhkəmə tətbiqləri baxımından bu yanaşmaların əsas üstünlüyü ondan ibarətdir ki, anonim nitq nümunəsindən danışanın regional və sosial mənsubiyyətinə dair ilkin xarakteristika çıxarmaq mümkündür. Bununla belə, alınan nəticələrin qəti və dəyişməz identifikasiya kimi deyil, ehtimal yönümlü mühakimə kimi qiymətləndirilməsi metodoloji baxımdan daha doğrudur [1, 2]. Başqa sözlə, nitq əsaslı yanaşmalar birbaşa şəxsiyyətin müəyyənləşdirilməsindən daha çox, şübhəli dairənin daraldılması və danışanın profilinin müəyyən olunması üçün səmərəli hesab edilməlidir.

V. Mətn əsasında identifikasiya və regional müəlliflik xüsusiyyətlərinin müəyyənləşdirilməsi

5.1. Müəllifin profilləşdirilməsi

Mətn əsaslı yanaşmalar dialektlər əsasında fərdlərin identifikasiyası probleminin ikinci mühüm istiqamətidir. Burada müəllifin leksik seçimi, sintaktik quruluşu, diskurs xüsusiyyətləri və fərdi üslub xüsusiyyətləri təhlil olunur. Tədqiqatlar göstərir ki, sintaktik və diskurs əlamətləri yalnız leksik göstəricilərlə müqayisədə daha sabit və bəzi hallarda daha informativ ola bilir [12].

5.2. İdiolekt və fərdi üslub

İdiolektin modelləşdirilməsinə dair işlər fərdi yazı üslubunun təsadüfi deyil, müəyyən dərəcədə sabit və fərqləndirici olduğunu göstərir [4]. Bu, fərdin dil davranışını yalnız qrup səviyyəsində deyil, fərdi üslub səviyyəsində də təhlil etməyə imkan verir.

5.3. Regional müəlliflik markerləri

Regional müəlliflik xüsusiyyətlərinin müəyyənləşdirilməsi sahəsində geoişarələnmiş sosial media korpusları və məkan statistikası xüsusi rol oynayır. Bu yanaşmalar regional markerlərin intuisiya əsaslı üsullarla deyil, verilənlərə əsaslanan obyektiv prosedurlarla *çıxarılmasına* şərait yaradır [3]. Yazılı dildə regional variasiyanın geniş yayılması və statistik cəhətdən izlənə bilməsi də bu istiqamətin perspektivliyini artırır [13].

Beləliklə, mətn əsaslı yanaşmalar fərdin birbaşa identifikasiyasından daha çox onun fərdi üslub sabitliyinin, regional yazı davranışının və sosial-kommunikativ profilinin müəyyənləşdirilməsi baxımından əhəmiyyətlidir. Azərbaycan dili üçün də bu xətt gələcəkdə praktik cəhətdən mühüm ola bilər, çünki yazılı korpusların qurulması bəzi hallarda geniş həcmli səs bazalarının yaradılmasından daha əlçatan görünür.

VI. Dialekt əsaslı identifikasiyanın əsas problemləri

6.1. Elmi-metodoloji problemlər

Dialektlər əsasında fərdlərin identifikasiyası bir sıra mühüm elmi çətinliklərlə müşayiət olunur. İlk problem dialekt, aksent və idiolekt səviyyələrinin həmişə aydın şəkildə ayrılmamasıdır. Eyni dialekt qrupuna mənsub olan şəxslər arasında ciddi fərdi fərqlər ola bilər; bununla yanaşı, bir şəxsin dil davranışı situasiya, auditoriya və kommunikativ məqsəd dəyişdikcə fərqlənə bilər [4].

İkinci mühüm problem resurs çatışmazlığıdır. Bir çox dillərdə, xüsusilə aşağı resurslu dillərdə, balanslaşdırılmış dialekt korpusları, etibarlı səs bazaları, standart annotasiya prinsipləri və regional marker siyahıları kifayət qədər formalaşmamışdır [3, 5]. Bu problem Azərbaycan dili üçün də aktualdır.

Üçüncü problem nitq siqnalının dəyişkənliyi ilə bağlıdır. Səs keyfiyyəti, rabitə kanalı, kənar səs-küy, qısa nitq parçaları, emosional vəziyyət, maskalanmış səs və ya qəsdən dəyişdirilmiş aksent nəticələrin etibarlılığına mənfi təsir göstərə bilər [2, 9].

6.2. Tətbiqi və etik problemlər

Mətn mühitində də ciddi çətinliklər mövcuddur. Sosial media və qeyri-rəsmi yazışmalarda orfoqrafik variasiya, qısa cümlələr, kod dəyişməsi və qeyri-standart yazı formaları regional və fərdi markerlərin aşkar edilməsini çətinləşdirir [3, 12].

Nəhayət, etik və hüquqi risklər də nəzərə alınmalıdır. Yanlış profilin qurulması, region və ya sosial qrup haqqında səhv nəticələrin çıxarılması, stereotipləşdirmə və ekspert nəticələrinin həddindən artıq qəti təqdim olunması məhkəmə və təhlükəsizlik kontekstində xüsusi həssaslıq yaradır [1, 2]. Buna görə də bu sahədə əldə olunan nəticələr həmişə ehtiyatlı və mövcud metodoloji məhdudiyyətlər nəzərə alınmaqla şərh olunmalıdır.

VII. Azərbaycan dili üçün elmi-praktik perspektivlər

Azərbaycan dili üçün dialektlər əsasında fərdin identifikasiyası məsələsi həm nəzəri, həm də tətbiqi baxımdan perspektivlidir. Burada əsas məqsəd ilkin mərhələdə şəxsin tam identifikasiyası deyil, regional profilləmə və dialekt markerlərinin çıxarılması olmalıdır.

7.1. Birinci mərhələ: resurs bazasının qurulması

İlk mərhələdə mövcud dialektoloji materialların rəqəmsallaşdırılması və sistemləşdirilməsi vacibdir. Bu baxımdan Naxçıvan və Qarabağ dialektoloji atlasları, mövcud lüğətlər və səs yazıları ilkin baza kimi çıxış edə bilər [5]. Regional markerlərin ilkin siyahısının hazırlanması, səs və mətn materiallarının toplanması, coğrafi və sosial metaverilənlərin əlavə olunması bu mərhələnin əsas vəzifələri olmalıdır.

7.2. İkinci mərhələ: analitik modelin qurulması

Növbəti mərhələdə dialekt marker bazası yaradılmalı, akustik, leksik və sintaktik xüsusiyyətlər seçilməli, onların ölçülməsi və qruplaşdırılması üçün uyğun hesablama üsulları tətbiq olunmalıdır. Dialektometrik yanaşmalarda istifadə olunan oxşarlıq və məsafə matrisləri, klasterləşdirmə və vizuallaşdırma üsulları burada faydalı ola bilər [5]. Pilot mərhələdə əvvəlcə bir və ya iki regional qrup əsasında pilot modelin hazırlanması daha məqsədəuyğun görünür.

7.3. Üçüncü mərhələ: tətbiqi sınaqlar

Son mərhələdə regional profilləmə modellərinin sınağı aparıla bilər. Buraya danışanın mənsubiyyətinin ilkin proqnozlaşdırılması, anonim mətnlərdə regional müəlliflik markerlərinin yoxlanması və nitq-mətn birgə modellərinin sınaqdan keçirilməsi daxildir. Bu istiqamət gələcəkdə həm elmi tədqiqatlar, həm də tətbiqi sistemlər üçün əhəmiyyətli nəticələr verə bilər. Beləliklə, Azərbaycan dili üçün bu sahənin inkişafı mərhələli yanaşma tələb edir və ilkin prioritet dialekt markerlərinin çıxarılması, korpusların yaradılması və regional profilləmə modelinin qurulması olmalıdır.

Nəticə

Aparılan təhlil göstərir ki, dialektlər əsasında fərdlərin identifikasiyası problemi klassik dialektologiyanın sərhədlərini aşaraq məhkəmə dilçiliyi, nitq texnologiyaları, korpus dilçiliyi və təbii dilin emalı ilə kəsişən kompleks elmi istiqamətə çevrilmişdir. Dialekt və aksent göstəriciləri birbaşa şəxsin identifikasiyasından daha çox regional və sosial profilləmə üçün uyğundur. Nitq əsaslı yanaşmalar bu sahədə daha yetkin metodoloji baza formalaşdırmışdır, lakin mətn əsaslı idiolekt və regional müəlliflik təhlili də sürətlə inkişaf edən istiqamətlərdəndir. Azərbaycan dili üçün prioritet istiqamət dialekt markerlərinin çıxarılması, rəqəmsal korpusların

yaradılması və regional profilləmə modelinin qurulması olmalıdır. Bu səbəbdən Azərbaycan dili üçün ilkin mərhələdə birbaşa şəxsin identifikasiyasını hədəfləyən modellərdən əvvəl, dialekt markerlərinin sistemləşdirilməsi və regional profilləmə sistemlərinin yaradılması elmi və praktik baxımdan daha əsaslandırılmış istiqamət hesab oluna bilər.

ƏDƏBİYYAT

[1] P. Foulkes and P. French, "Forensic speaker comparison: A linguistic-acoustic perspective," in *The Oxford Handbook of Language and Law*. Oxford, U.K.: Oxford University Press, 2012.

[2] N. Schilling and A. Marsters, "Unmasking identity: Speaker profiling for forensic linguistic purposes," *Annual Review of Applied Linguistics*, vol. 35, pp. 195–214, 2015.

[3] D. Roemling, "Forensic authorship profiling using geolocated social media data: A corpus linguistic and cartographic approach," *Applied Corpus Linguistics*, vol. 5, no. 3, Art. no. 100146, 2025.

[4] J. Zhu and D. Jurgens, "Idiosyncratic but not arbitrary: Learning idiolects in online registers reveals distinctive yet consistent individual styles," in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2021, pp. 279–297.

[5] A. Gurbanoba and M. Baghirova, "Dialectometry and the Azerbaijani language: Problems, solutions and prospects," *e-Journal of Linguistics*, vol. 18, no. 1, pp. 1–10, 2023. https://library.ict.az/elektron/2023/Dialectometry_and_the_azerbaijani_language_problems_solutions_and_perspectives.pdf

[6] P. Ahuja and J. M. Vyas, "Forensic speaker profiling: The study of supra-segmental features of Gujarati dialects for text-independent speaker identification," *Australian Journal of Forensic Sciences*, vol. 50, no. 2, pp. 152–165, 2018.

[7] O. Köster, R. Kehrein, K. Masthoff, and Y. H. Boubaker, "The tell-tale accent: Identification of regionally marked speech in German telephone conversations by forensic phoneticians," *International Journal of Speech, Language and the Law*, vol. 19, no. 1, pp. 51–71, 2012.

[8] Y. Korkmaz and A. Boyacı, "A comprehensive Turkish accent/dialect recognition system using acoustic perceptual formants," *Applied Acoustics*, vol. 193, Art. no. 108761, 2022.

[9] S. Saleem, F. Subhan, N. Naseer, A. Bais, and A. Imtiaz, "Forensic speaker recognition: A new method based on extracting accent and language information from short utterances," *Forensic Science International: Digital Investigation*, vol. 34, Art. no. 300982, 2020.

[10] G. Brown, J. Franco-Pedroso, and J. González-Rodríguez, "A segmentally informed solution to automatic accent classification and its advantages to forensic applications," *International Journal of Speech, Language and the Law*, vol. 28, no. 2, pp. 201–232, 2021.

[11] M. A. Humayun, J. Shuja, and P. E. Abas, "Speaker profiling based on the short-term acoustic features of vowels," *Technologies*, vol. 11, no. 5, Art. no. 119, 2023.

[12] J. Soler-Company and L. Wanner, "On the relevance of syntactic and discourse features for author profiling and identification," in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2017, pp. 681–687.

[13] J. Grieve, *Regional Variation in Written American English*. Cambridge, U.K.: Cambridge University Press, 2016.

Investigation and Analysis of the Problems of Individual Identification Based on Dialects

Afruz Gurbanova[1], Sariya Gasimova[2]

[1] Institute of Information Technology, Baku, Azerbaijan
[2] Nakhchivan State University, Nakhchivan, Azerbaijan
[1]afruz1961@gmail.com,[2]qasimovaseriyye@ndu.edu.az

Abstract - This paper investigates the problem of individual identification based on dialects through an analytical literature review. It is shown that this problem represents a complex research direction formed at the intersection of dialectology, forensic linguistics, speech technologies, corpus linguistics, and natural language processing. The paper comparatively examines speaker profiling based on speech signals, regional affiliation identification, accent recognition, and forensic-phonetic comparison approaches, as well as author profiling based on texts, the extraction of regional authorship markers, and the analysis of idiolect. The analysis shows that dialect and accent indicators provide more scientifically grounded opportunities for the probabilistic determination of an individual's regional, social, and communicative profile rather than for direct and definitive personal identification. The paper evaluates the application potential of existing dialectological and dialectometric resources for the Azerbaijani language in this direction, identifies the main methodological and practical challenges, and proposes phased development directions.

Keywords - dialect; individual identification; profiling; idiolect; forensic linguistics; natural language processing.

III SEKSİYA
Kritik informasiya infrastrukturu, kiberfiziki və idarəetmə sistemlərinin kibertəhlükəsizliyi problemləri