



Mersin Üniversitesi Dil ve Edebiyat Dergisi, MEÜDED, 2024; 20 (2), 105-126.

DİL BİLİMDE KULLANILAN SES BİLGİSİ MESAFE ARAÇLARI

Phonetic Distance Tools Used in Linguistics

Cemile UZUN¹

Fırat Üniversitesi

ORCID ID: 0000-0002-0102-3306

Abstract: Studies of phonetic distance between languages are important for understanding language diversity and analysing how language structures vary. These studies reveal how certain words or language structures vary across different regions and provide quantitative data for the study of language diversity. The goal of the study is to determine which tools are used to measure the phonetic distance of languages and to classify these tools according to the methods they employ. For this purpose, research on phonetic distance from the past to the present has been examined. In this respect, the study will guide research that aims to analyze phonetic distances between languages and develop phonetic distance measurement tools. The most cited articles, books and theses on phonological distance in languages or dialects were identified. The methodological sections of the identified studies were classified and

¹ Öğr. Gör. Dr., Fırat Üniversitesi, TÖMER, cemileuzun000@gmail.com

Gönderim tarihi: 13 Mayıs 2024; Kabul tarihi: 5 Temmuz 2024

Kaynak gösterimi: Uzun, C. (2024). Dil biliminde kullanılan ses bilgisi mesafe araçları. *Mersin Üniversitesi Dil ve Edebiyat Dergisi*, 20(2), 105-126.

compared. Since the findings of many studies were classified, a systematic review was used in this study. Studies with the same data collection and analysis methods were combined. It was seen that three types of methods, namely traditional, perceptual and computational methods, were used to calculate and classify the differences between languages and dialects.

Keywords: *Language, Phonetic, Distance, Phonetic distance, Phonetic distance tools*

Öz: Diller arasındaki ses bilgisi mesafesi arařtırmaları, dil çeřitliliđini anlamak ve dil yapılarının nasıl deđiřtiđini incelemek için oldukça önemlidir. Bu alıřmalar, belirli sözcüklerin veya dil yapılarının farklı bölgeler arasında nasıl deđiřtiđini ortaya koyar ve dil çeřitliliđinin arařtırılmasına yönelik niceliksel veri sunar. alıřmanın amacı, dillerin ses bilgisi mesafesini ölçmek için hangi araçların kullanıldığını belirlemek ve bu araçları kullandıkları yöntemlere göre sınıflandırmaktır. Bu amaçla, geçmişten günümüze kadar ses bilgisi mesafe ile ilgili alıřmalar incelenmiştir. Bu yönüyle alıřma, diller arasındaki ses bilgisel mesafeyi çözümlmek, ses bilgisel mesafe ölçüm araçları geliřtirmek isteyen arařtırmalara rehberlik edecektir. Diller veya lehçeler hakkında yapılmış ses bilgisel mesafe ile ilgili en çok atıf alan makaleler, kitaplar ve tezler belirlenmiştir. Belirlenen alıřmaların yöntem bölümleri sınıflandırılmış ve karşılaştırılmıştır. Çok sayıda alıřmanın bulguları sınıflandırıldığı için alıřmada sistematik inceleme kullanılmıştır. Aynı veri toplama ve çözümlme yöntemine sahip alıřmalar birleştirilmiştir. Diller ve lehçeler arasındaki farklılıkları hesaplamak ve sınıflamak için geleneksel, algısal ve hesaplamalı yöntem olmak üzere üç çeřit yöntemin kullanıldığı görülmüřtür.

Anahtar sözcükler: *Dil, Ses bilim, Mesafe, Ses bilgisi mesafesi, Ses bilgisel mesafe araçları*

1. GİRİŐ

Dil bilim, dillerin yapısını, kullanımını ve gelişimini inceleyen bir bilim dalıdır. Diller arasındaki ses farklılıklarını belirlemek dil bilimin konularından bir tanesidir. Dil bilimde, dillerin birbirine olan ses bilgisi mesafesini belirlemek için birçok yöntem kullanılmaktadır. Ses bilgisi mesafesi, farklı konuşma çeřitleri arasındaki ses bilgisel farklılıkların ölçülmesini ifade eder. Ses bilgisi mesafeleri, dildeki ses bilgisel çeřitliliđi incelemek için istatistiksel yöntemleri ve dil bilimsel verileri kullanır. Bu yöntemlerin hedeflerinden biri, belirli sözcüklerin veya dil

yapılarının farklı bölgeler veya topluluklar arasında nasıl dağıldığını ve bu dağılımın nasıl değişiklik gösterdiğini incelemektir.

Ses bilgisi mesafe yöntemleri, diller arasındaki ses farklılıklarını belirlemek için kullanılır. Ses bilgisi mesafe yöntemleri, dillerin ses sistemlerindeki benzerlikleri ve farklılıkları keşfetmenin ve anlamının niceliksel bir yoludur. Özellikle dillerin birbirleriyle ilişkisini anlamak için kullanılan bir kavramdır. Ayrıca ses bilimsel mesafe coğrafi alanlar arasındaki dil çeşitliliğini ve değişimini anlamaya katkıda bulunur. İnsanlar yeni yerlere göç edince dilsel özelliklerini de beraberlerinde taşırlar. Zamanla bu özellikler yerel dille karışır, yeni konuşma farklılıkları oluşur. Ses bilgisi mesafe yöntemleri, dillerin nasıl yayıldığını ve birbirlerini nasıl etkilediğini anlamaya yardımcı olur. Ses bilgisi mesafesini ölçmek için kullanılan araçlar, çeşitli alanlarda dil bilimsel sınıflandırmalar için kullanılan nesnel verilerdir. Ses bilgisi mesafesini doğru bir şekilde ölçecek araçların geliştirilmesi, dil bilim için oldukça önemlidir. Ses bilimsel mesafeyi ölçmek için standartlaştırılmış bir yöntem belirlemek, dil bilim araştırmalarında kullanılan ses bilgisi mesafe ölçme yöntemlerini birleştirmeye ve farklı dil grupları ile ilgili yapılan çalışmalarda güvenilir sonuçlar elde etmeye yarar. Ses bilimsel mesafe ölçüm yöntemlerinin çeşitli dillere ve dil birimlerine uygulanabilirliğini test etmek, dil bilim araştırmaların kapsamını genişletecektir.

Ses bilgisi mesafesi, ses bilimsel çözümlemenin çok yönlü doğasını vurgulayan çeşitli kavramları kapsar (Kessler, 2005, s. 247). Ses bilimsel mesafenin incelenmesi, dil tipolojisi, lehçe çeşitliliği, dil teması, ikinci dil edinimi, toplum dil bilim ve küreselleşmenin dil üzerindeki etkisi konularında önemli bir rol oynamaktadır. Ses bilgisi mesafe yöntemleri, diller arasındaki ses bilimsel farklılıkları belirlemek için kullanılan sistematik bir yaklaşımdır, araştırmacıların ses bilimsel farklılıkları inceleyerek dilin sosyal ve kültürel boyutlarını ölçmesine ve incelemesine olanak tanır. Aynı zamanda, bir dildeki ses bilimsel farklılıkların ayrıntılı bir şekilde araştırılmasını sağlar.

Ses bilgisi mesafe yöntemleri, dillerin birbirleriyle olan etkileşimlerine ışık tutar. Dil değişimlerinin sosyo-kültürel bağlantılarını aydınlatır. Ses bilgisi mesafe çalışmaları, araştırmacılara farklı bakış açıları sunar. Araştırmacılar, dil içindeki ve diller arasındaki ses farklılıklarını inceleyerek birçok alanda bilgi edinebilirler. Bu durum, dillerin bölgesel ve sosyal olarak nasıl değiştiğini anlamaya yardımcı olur. Ses bilgisi mesafe yöntemleri, diller arasındaki ilişkileri ve dillerin ortak atalardan nasıl farklılaştıklarını anlamak için bir araçtır. Dillerin

birbirlerinden ödünç aldığı sesleri ortaya çıkarmaya yardımcı olur. Bir dil içindeki çeşitliliği ortaya çıkarabilir.

Bu çalışmanın amacı, farklı diller arasındaki ses bilgisi mesafelerini hesaplamak için kullanılan araçları belirlemektir. Bu amaçla, çalışmalarda kullanılan ses bilgisi mesafe araçlarının sınıflandırılması, bu araçların etkinliğinin değerlendirilmesi hedeflenmiştir. Bu çalışmanın özgün değeri, ses bilgisel mesafe ölçüm araçlarını sistematik bir şekilde sınıflandırarak dil bilimi ve dil teknolojisi alanlarındaki araştırmalara katkıda bulunmasıdır.

2. LİTERATÜR TARAMASI

İlgili alan yazını tarandığında ses bilgisel mesafeyi ölçmeye yönelik yöntemlerin yıllar içerisinde teknolojiye bağlı olarak geliştiği ve ilerlediği görülmektedir. Séguy'un (1973) çalışması ses bilgisel mesafe yaklaşımlarının temelini atmıştır. Séguy'un tekniği, ses bilgisel mesafe yöntemlerinin gelişmesini sağlamıştır. Wagner ve Fischer (1974), en az işlem yaparak iki dize arasındaki ses bilgisel mesafeyi bulmayı amaçlamışlardır. Bunun için "string to string" tekniğini kullanmışlardır. Nerbonne ve diğerleri (1996), Hollanda lehçeleri arasındaki ses bilgisel uzaklığı ölçmek için Levenstein mesafesini kullanmışlardır. Nerbonne ve Heeringa (1997) Öklid mesafesi ve Pearson korelasyonunu kullanarak lehçeler arasındaki ses bilgisel mesafeyi ölçmenin etkisini araştırmışlardır. Chambers ve Trudgill (1998), diyalektometri alanında yapılan çalışmalarını sorgulamışlar, hemen kabul etmemişlerdir. Goebel gibi araştırmacıların çalışmalarını inceledikten sonra diyalektler arasındaki sınırları çizmek için aynı kökene sahip sözcüklerin kullanılabilirliğini belirtmişlerdir (Chambers ve Trudgill, 1998). Covington (1998), birden fazla dil arasındaki sözcükleri hizalayan bir yöntem geliştirmiştir. Tiedemann (1999), iki dil arasında aynı kökten gelen sözcükleri belirlemek için dize benzerlik metrislerini kullanmıştır. Melamed (1999), çalışmasında iki farklı dilde yazılmış metinlerin (bitextlerin) karşılıklı olarak eşleştirilmesi ve bu metinler arasındaki benzerliklerin belirlenmesi için yeni bir yöntem geliştirmiştir.

2000'li yılların başında itibaren diller arasında ses bilgisel benzerliği ve sözcük hizalamasını hesaplamak için yeni algoritmalar geliştirmeye başlanmıştır. Kondrak (2000), çalışmasında ses dizilerinin hizalanma yöntemlerini değerlendirdikten sonra ses bilgisel benzerliği

hesaplamak için yeni bir algoritma sunmuştur. Ribeiro ve diğerleri (2001), farklı dillerdeki sözcükler arasındaki benzerlikleri ölçmek için bir yöntem geliştirmişlerdir. Sooful ve Botha (2002), sesler arasındaki akustik uzaklığı ölçmek için Kullback-Leibler, Bhattacharyya, Öklid, L2 metriği ve Jeffreys-Matusita ölçülerini kullanmışlar ve bu ölçüleri karşılaştırmışlardır. Kondrak (2003), farklı dillere ait sözcükler arasındaki ses bilimsel benzerliği ve hizalamayı hesaplamak için yeni bir yöntem geliştirmiştir. Gooskens ve Heeringa (2004), 15 Norveç lehçesi arasındaki mesafeleri hem Levenshtein mesafesiyle hem de Norveçli dinleyicilerin algılarına dayanarak ölçmüşler ve Levenshtein mesafesinin lehçe konuşurlarının psikoakustik algılarıyla ne derece uyum sağladığını incelemişlerdir. Gooskens (2005), Norveç lehçelerinde, dil bilimsel mesafeler ile coğrafi mesafeler arasındaki bağlantıyı hesaplamıştır. Kondrak ve Dorr (2006), aynı yazıma ve aynı telaffuza sahip ilaç adlarının karışıklığını önlemek için, yazım ve ses benzerliklerine dayalı iki yöntem önermiş. Ardından, bu iki yöntemin birleştirilmesinin (yazım ve ses benzerliği üzerine kurulu bir sistem) ilaç adları arasındaki karışıklığı en iyi şekilde azalttığını göstermişlerdir. Goebel'in çalışmaları ses bilimsel mesafe alanında önemli bir ilerleme sağlamıştır. Goebel aynı sözcüklerin farklı diyalektlerde nasıl seslendirildiği gibi özellikleri içine alan bir sınıflandırma geliştirmiştir (Goebel, 2006). Mulloni ve Pekar (2006), benzer kökenli sözcükleri otomatik olarak tespit etmek için yeni bir yöntem önermişlerdir. Ellison ve Kirby (2006), PHILOLOGICON adındaki algoritmayı kullanarak aynı dile ait sözcükler arasındaki ses bilgisi benzerliğini ölçmüşlerdir. Daha sonra bu matrisleri kullanılarak diller arasındaki ses bilgisi mesafelerini hesaplamışlardır. Pucher ve diğerleri (2007), Sözcük-Hata-Oranı (Word-error-rate) (WER) ölçümlerinden elde edilen matrisler ile ses bilgisi mesafe ölçümleri arasındaki ilişkiyi araştırmışlardır. Ses bilgisi mesafesini, çeviri yazılar arasındaki minimum-düzenleme mesafeleri ve Gizli-Markov Modelleri (HMM) kullanarak değerlendirmişlerdir. Heeringa ve diğerleri (2009), çeviri yazı kullanmadan sadece akustik sinyallerle Norveç lehçeleri arasındaki ses bilgisi mesafesini ölçmüşlerdir.

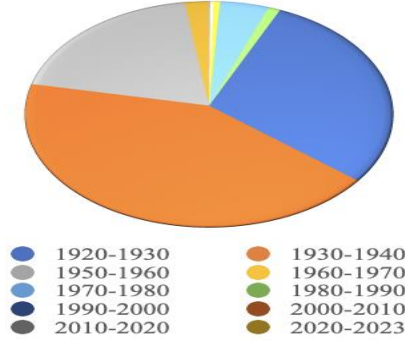
2010 yılında itibaren yapılan çalışmalarda, önceki yıllarda ortaya konmuş yöntemlerin daha da ileriye taşındığı görülmektedir. Bu yıllarda yapılan çalışmalarda, yeni yöntemler geliştirilmiş ve bu yöntemler farklı dillere, konuşma tanıma sistemlerine ve teknolojiye uyarlanmıştır. Kisler ve Reichel (2013), Levenshtein mesafesinin bir dizi uzantısını tanıtmışlardır. Kang (2015), Korecenin ses bilgisi mesafesini ölçmek için hangi yöntemin en uygun olduğunu belirlemeye

çalışmıştır. Babu ve diğerleri (2015), Dinamik Fonem Çarpıtma (DPW) tekniğini kullanarak iki ses arasındaki uzaklığı ölçmeyi amaçlamışlardır. Eden (2018), sesler arasındaki benzerliği ölçmek için kullanılan yöntemleri araştırmıştır. Droppo ve Acero (2010), ses dize düzenleme mesafesinin konuşma tanıma sistemlerine uyarlanması üzerinde durmuşlardır. Kudera ve diğerleri (2021), Slav dilleri arasındaki ses bilgisel uzaklığı ölçmek için “Word adaptation surprisal” yöntemini kullanmışlardır. MacLeod (2021), çalışmasında Difference-in-Distance (DID) ve linear kombinasyon tekniklerini kullanarak hangi yöntemin ses bilgisi mesafesini ölçmede daha iyi olduğunu ortaya çıkarmaya çalışmıştır. Ahmed ve diğerleri (2022), farklı diller arasındaki sözcük benzerliğini ölçmek için geleneksel mesafe yöntemlerinden farklı olarak IPA çeviri yazı sistemini kullanmışlar, Phonetic Edit Distance yöntemi ile karşılaştırmışlardır.

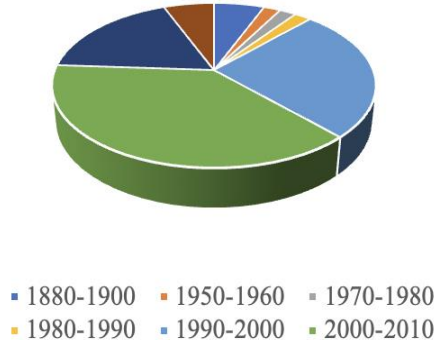
3. YÖNTEM

Çalışmanın yöntemi, “Dil arasında ses bilgisi mesafesini ölçmek için hangi yöntemler kullanılır?” araştırma sorusu üzerine kurulmuştur. Araştırma sorusundan hareketle günümüze kadar yapılmış ses bilgisi mesafe araçları ile ilgili araştırmaların yöntem kısımları araştırılmıştır. Çok sayıda çalışmanın bulguları sınıflandırıldığı için çalışmada sistematik inceleme kullanılmıştır. Sistematik inceleme, önceden tanımlanmış bir soruya yanıt aramak için kullanılan, rapor olarak da adlandırılan akademik bir makale türüdür. Sistematik bir çalışma, araştırmaya dahil edilen ve hariç tutulan çalışmaların ölçütlerini, ilgili çalışmalara ulaşmak için kullanılan arama stratejisini ve araştırmaya dahil edilen çalışmaların sistematik kodlanmasını ve çözümlemesini içermelidir (Campbell Collaboration, 2024).

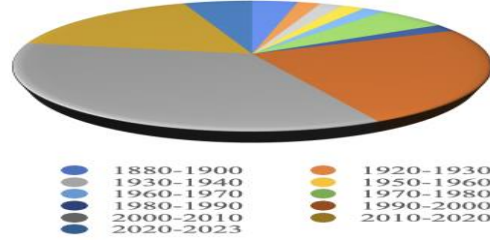
Dokümanlara ulaşmak için Scopus, Scopus h-index, Pubmed, ScienceDirect, Google Scholar, Google Scholar Citations ve Research Rabbit veri tabanları kullanılmıştır. Toplamda 2378 dokümana ulaşılmıştır. Bu çalışmaların 1556 çalışma makale, 402 bildiri, 215 kitap ve 205 tez türündedir. Tespit edilen çalışmaların yayımlandığı yıllar aşağıdaki tabloda gösterilmiştir.

Tablo 1. Tespit edilen çalışmaların yayım yılına göre dağılımı

Araştırmaya dahil edilen çalışmalar yukarıdaki tabloda gösterilen yıllar arasında yazılmış, yayım dili İngilizce ve Almanca olan, en çok atıf alan çalışmalardır. Diller arasındaki köken akrabalığını araştıran ses bilgisi mesafe araçları çalışmaya dahil edilmemiştir.

Tablo 2. En çok atıf alan araştırmaların yıllara göre oranı

Örneklem seçerken olasılığa dayalı örneklem tekniklerinden olan sistematik örneklem ve tabakalı (katmanlı) örneklem tespiti kullanılmıştır. Sistematik örneklem tespiti, bir topluluktan belirli aralıklarla örneklem seçmek için kullanılan olasılıklı örnekleme yöntemidir. Tabakalı örneklem tespiti, örnekleme hatalarını azaltarak daha yüksek bir temsil yeteneğine sahip örneklem oluşturmak için kullanılmaktadır (Sencer ve Sencer 1978, s. 466). Örneklem aralığı 50 olarak belirlenmiştir. Bu veriler doğrultusunda 31 tane makale, 8 tane bildiri, 4 tane kitap ve 4 tane tez çalışma kapsamında incelenmiştir. Bu eserler ise en çok atıf almış çalışmalardan seçilmiştir. Seçilen çalışmaların yıllara göre sayısal dağılımı aşağıdaki tabloda gösterilmiştir.

Tablo 3. Seçilen çalışmaların yıllara göre sayıları

4. İNCELEME

İncelenen çalışmaların yöntemlerinde kullanılan teknikler belirlenmiş, daha sonra bu yöntemler benzerliklerine ve farklılıklarına göre karşılaştırılmıştır.

Diller ve lehçeler arasındaki ses bilgisel farklılıkları hesaplamak için geleneksel, algısal ve hesaplamalı yöntem olmak üzere üç çeşit yöntemin kullanıldığı görülmüştür. Bu çalışmaların sayısal olarak dağılımı aşağıdaki tabloda gösterilmiştir.

Tablo 4. Araştırmalarda ses bilgisi mesafe aracı olarak kullanılan yöntemler

4.1. GELENEKSEL YÖNTEMLER

Geleneksel ses bilgisi mesafe ölçüm araçları tamamen sezgi veya bilgiye dayalı sınıflandırmalar kullanılarak yapılmıştır. Araştırmacıların kişisel deneyimleri ve lehçelere olan aşinalıkları, veri toplama ve sınıflandırma süreçlerinde önemli bir yer tutmuştur. Geleneksel yöntemler, genellikle saha çalışması ile ilgili çalışmalardan oluşmaktadır. Araştırmacılar saha çalışması sırasında veri toplarken anket ve röportaj tekniklerini kullanmışlardır. Winkler (1874) çalışmasında Hollanda, Almanya ve Belçika'nın bazı bölgelerinde

konusulan ağızların kapsamlı bir araştırmasını yapmıştır. Bu ağızların çok sayıda örneğini toplayıp çözümleyerek her lehçenin dil bilgisi özelliklerine ilişkin ayrıntılı açıklamalar yapmıştır. Jellinghaus (1892), çalışmasında farklı ağızları konuşan kişilere anketler göndererek veri toplamıştır. Bu anketlerde katılımcılardan belirli sözcükleri ve ifadeleri kendi ağızlarında nasıl ifade ettiklerini belirtmelerini istemiştir. Toplanan verileri çözümleyerek farklı ağızlar arasındaki ses ve biçim bilgisi özelliklerini karşılaştırmıştır. Elde edilen bulgular, Hollanda ağızlarının sınıflandırılmasında kullanılmış ve bu ağızların tarihi ve coğrafi bağlantılarına dair bilgiler sağlamıştır. Hoppenbrouwers ve Hoppenbrouwers (2001)'in çalışması 156 farklı şehir ve köyde konuşulan Hollanda lehçelerinin sınıflandırmasını içermektedir. Kessler (1995), çalışmasında İrlanda lehçelerini karşılaştırmak için izogloss matrisini kullanmıştır. İzogloss matrisi, coğrafi bölgeler arasındaki dil farklılıklarını ve dil özelliklerini incelemek için kullanılan bir tekniktir. Kessler tarafından 1995 yılında İrlanda lehçelerini sınıflandırmak için kullanılan izogloss matrisi kavramı, ses bilgisi, söz dizimi gibi farklı dil özelliklerine sahip alanlar arasındaki coğrafi sınırı belirtmek için kullanılmıştır.

Geleneksel yöntemde, başlangıçta bilgi ve sezgiye güvenilmiş ve yöntem araştırmacıların bilgisi ve sezgisi üzerine kurulmuştur. Ancak bu yöntem nesnellik, tekrarlanabilirlik ve kapsam açısından yetersiz kalmıştır. Geleneksel yöntem, diyalektolojiyi bir disiplin olarak kullanma, saha çalışmaları yapma, ses bilgisini çeşitli açılardan çözümleme ve ağız bölgelerini görsel olarak haritalandırma gibi aşamalarla daha sistematik yöntemlere geçiş için hazırlık sürecinin başlangıcı olmuştur. Bu gelişmeler, deneysel verilere dayalı kapsamlı diyalekt sınıflandırmalarına olanak sağlamıştır. Geleneksel yöntemlere dayanan lehçe sınıflandırması, dil, kültür ve tarih arasındaki karmaşık ilişkileri vurgulayan ve dil çeşitliliği çalışmalarında önemli bir rol oynayan, zengin ve çok yönlü sınıflandırmalara zemin hazırlamıştır.

4.2. ALGISAL YÖNTEMLER

Algısal yöntemler nicel ölçümlere dayanmak yerine, dili insan algısına dayalı olarak değerlendirmeyi ve sınıflandırmayı içerir. Algısal yöntem, geleneksel yöntemler ile algı tekniklerini birleştirir. Dil biliminde algısal yöntemler, farklı lehçelerin konuşmacılar ve dinleyiciler tarafından nasıl algılandığını ve ayırt edildiğini ortaya çıkarmak için kullanılır. Bu yöntemde, araştırmacılar çeşitli ağızları konuşanlara bazı anketler yapar. Bu anketlerde kişilere kendi konuşma

özelliklerine benzeyen ve kendi konuşma özelliklerinden farklılık gösteren yerleri belirtmeleri istenir. Kişilerin yanıtlarına göre haritalar oluşturulur. Haritalarda kişilerin kendi konuşma özelliklerine benzeyen yer olarak işaretledikleri alanlar oklar ile birleştirilir. Kişilerin kendi konuşma özelliklerinden farklılık gösteren yerler ise beyaz çizgi ile birbirine bağlanır (Eden, 2018). Bu tür algısal mesafe ölçen yaklaşımlar farklı bölgelerden konuşmacıların diğer lehçeleri kendi lehçeleriyle ilişkili olarak nasıl algıladıklarına ve sınıflandırdıklarına dair içgörü sağlar. Algısal yöntemlerin en temel faydası basit olmalarıdır (Och ve Ney, 2003, s. 25). Algısal yöntemde iki farklı teknik kullanılmıştır: ok tekniği ve derecelendirme ölçek teknikleri (Heeringa, 2004, ss. 13-14).

4.2.1. OK TEKNİĞİ

Hollanda'daki lehçe sınırlarını çözümlmek için bir dizi dil bilimsel anket ve haritalama tekniği geliştirilmiştir. Ok tekniğinde katılımcılar, söz konusu dillerin konuşulduğu bölgeleri kendi dillerine benzerliği ya da kendi dillerinden farklılığı açısından değerlendirirler. Weijnen (1946), Hollanda'nın Kuzey Brabant eyaletinin doğu ve kuzey bölgelerinde konuşulan lehçeler üzerine bir çalışma yapmıştır. Çalışmasında 1500 katılımcıya iki soruluk bir anket uygulanmıştır. Anket sonuçlarına göre ağız haritası oluşturmuştur. Anket sorularına verilen yanıtlara dayanarak benzer lehçeleri birbirine bağlamak ve lehçe sınırlarını belirtmek için beyaz şeritler kullanmıştır. Rensink'i de Weijnen'in yönteminden yararlanarak Hollanda'nın ağızları için algısal ağız haritası oluşturmuştur. Rensink (1955), Hollanda lehçelerinin benzerliklerini ortaya koymak için oklu harita yöntemini kullanmıştır. Daan ve Blok (1969), Standart Hollandacanın dil özelliklerini gösteren harita çizmişlerdir. Haritalarında yeşil, beyaz, sarı, turuncu, kırmızı renklerini kullanarak Hollanda lehçelerini benzerliklerine göre sınırlandırmışlardır.

4.2.2. DERECELENDİRME ÖLÇEK TEKNİĞİ

Bu ölçek, katılımcıların bir dilin kendi dillerine ne kadar benzer ya da kendi dillerinden ne kadar farklı olduğunu sayısal olarak belirtmeleri üzerine kuruludur.

Gooskens (1997), Hollandacanın farklı lehçelerinin Standart Hollandacaya ne kadar yakın olduğunu belirlemek için algı tekniklerini kullanmıştır. Katılımcılar bir parçayı dinlemiş ve Standart

Hollandacaya göre benzerliği 1 ile 10 arasında bir sayı ile derecelendirmişlerdir.

4.3. HESAPLAMALI YÖNTEMLER

Hesaplamalı yöntemler, veri kümelerini çözümlmek için kullanılan bir yöntemdir. Büyük veri setlerini çözümlmek için hesaplamalı yöntemler gün geçtikçe daha fazla kullanılmaya başlanmıştır. Teknoloji gelişmeye devam ettikçe ses bilgisi mesafesini ölçen yeni yazılımlar ve analitik yöntemler ortaya çıkmıştır. Hesaplamalı yöntemlerin dil bilime uyarlanması ses bilgisi mesafesini çözümlmede çığır açmış ve araştırmacılara dil çeşitliliği ve değişiminin karmaşık kalıplarını çözümlmek ve anlamak için güçlü araçlar sağlamıştır. Ses bilgisi mesafesini çözümlmede hesaplamalı yöntemler iki şekilde kullanılmaktadır: Ağız haritalama ve derlem analizi.

Ağız haritalama, araştırmacıların dil verilerini coğrafi haritalar üzerinde görsel olarak temsil etmesine olanak tanır ve lehçe özelliklerinin farklı bölgeler arasındaki dağılımını gösterir. Araştırmacılar, dil verilerini coğrafi haritalar üzerine yerleştirerek dil çeşitlilik kalıplarını ve bunların coğrafi sınırlarla nasıl ilişkilendirildiğini kolayca belirlemişlerdir. Derlem çözümlmesi için tasarlanan yazılımlar, dil bilimcilerin dilin büyük veri kümelerini işlemesine ve sınıflandırmasına olanak tanımaktadır. Bu araçlar, sıklık sayımlarından daha karmaşık istatistiksel çözümlmelere kadar çeşitli teknikler gerçekleştirerek dil içi ve dil dışı eğilimlerin belirlenmesine yardımcı olurlar.

Dil bilimde hesaplamalı yöntemler Séguay ile başlamıştır. Séguay çalışmasında lehçe çeşitliliğini sınıflandırmak için daha sistematik ve nesnel bir yöntem geliştirmeyi amaçlamıştır. Séguay'un yöntemi diyalektoloji alanına niceliksel bir boyut kazandırmış, bölgeler arasındaki diyalekt farklılıklarını ölçmek ve karşılaştırmak için daha nesnel bir yöntem sunmuştur. Séguay, Gascony bölgesindeki ağızları benzerliklerine göre sınıflandırmıştır (Chambers ve Trudgill, 1998, s. 138). Goebel'in çalışmaları Séguay tarafından kullanılan metodolojiyle benzerlik taşısa da farklı özelliklere sahiptir. Séguay, lehçeler arasındaki farkları belirlemiş, Goebel benzerlikleri hesaplamaya odaklanmıştır.

Goebel çalışmasında;

#equal nominal values x 100

#equal nominal values + #different nominal values

formülünü kullanmıştır. Bu formül iki diyalekt arasındaki benzerliği ölçmek için kullanılır. Formül iki lehçe arasındaki benzerliği niceliksel olarak çözümlmek için tasarlanmıştır. Teknolojinin gelişmesi ile birlikte ses bilgisi mesafesini ölçmek için farklı yazılımlar geliştirilmiştir. Günümüze kadar çeşitli teknikler, yazılımlar vs. Goebel'in hesaplamalı yöntemi üzerine kurulmuştur. Diller arasındaki ses bilgisi mesafesini çözümlmeye yönelik bu hareket, lehçe çeşitliliğini daha kesin bir şekilde haritalamak için tasarlanmış, çeşitli yazılım araçlarının geliştirilmesine yol açan teknolojik gelişmelerle kolaylaştırılmıştır.

Günümüze kadar yapılan ses bilgisi mesafesi ile ilgili çalışmalarda kullanılan hesaplamalı yöntemler aşağıdaki gibi sınıflandırılabilir:

4.3.1. ALINE

ALINE, 2000 yılında Kondrak tarafından geliştirilmiştir (Kondrak, 2000). Aynı kökenli sözcük çiftleri arasındaki ses bilgisel uzaklığı, eşleştirme yaparak belirleyen bir algoritmadır. İlaç isimleri de dahil olmak üzere herhangi bir sözcük çiftinin ses bilgisi özelliklerini karşılaştırarak benzerliğini tahmin eden bir algoritmadır. ALINE, farklı dillerdeki sözcük çiftlerinin ses bilgisel benzerliğini değerlendirir. Başlangıçta, kognatları (benzer kökenli sözcükler) tanımlamak ve hizalamak için tasarlanmış olmasına rağmen evrensel ses bilgisi kurallarına dayandığı için ilaç isimleri dahil herhangi bir sözcük çiftinin benzerliğini tahmin etmede kullanılabilir (Kondrak, 2001; Kondrak ve Dorr, 2006).

4.3.2. BHATTACHARYYA MESAFESİ

İki grup veya kategori arasındaki farkı ölçmek için kullanılan bir yöntemdir (Mak ve Barnard, 1996, Sooful ve Botha, 2002). Farklı seslerin veya ses birimlerinin dağılımlarını karşılaştırmak için kullanılır. İki farklı dildeki benzer seslerin veya ses birimlerinin özellikleri, Bhattacharyya mesafesi kullanılarak nicel olarak hesaplanabilir.

4.3.3. CORDI

2002 yılında Kondrak tarafından geliştirilmiş bir programdır. Bu program, iki farklı dildeki sözcükler arasındaki ses benzerliklerini saptamayı ve bu benzerlikleri ya ortak bir kökene ya da diller

arasındaki sistematik bir ses değişikliğine kadar izlemeyi amaçlamaktadır (Kondrak ve Sherif, 2006).

4.3.4. COVINGTON

Sözcükler arasında hizalama yapmak için kullanılan yöntemlerden biridir. Bu yöntem, ünlüleri, ünsüzleri ve yarı ünlüleri ayırt eder. İki farklı dildeki sözcük çiftlerinin ses bilgisel olarak hizalanmasını sağlayan bir algoritmadır. Bu algoritma, iki sözcük arasındaki benzerlikleri ve farklılıkları belirleyerek, her iki sözcüğün ses bilgisel olarak en iyi şekilde hizalanmasını sağlar (Covington, 1998). Covington algoritması, ses bilgisel hizalama için dikkate değer bir yöntemdir ve özellikle hesaplamalı dil bilim ve diyalektolojide farklı dillerdeki sözcükleri karşılaştırmak için kullanışlıdır.

4.3.5. ÇİFT GİZLİ MARKOV MODELİ (PAIR HMM)

Mackay ve Kondrak (2005) çalışmasında, sözcük çiftleri arasındaki benzerliği ölçmek için bioinformatik alanında kullanılan tekniği kullanmışlardır. Çalışmalarında, genetik dizilerin çözümlenmesinde kullanılan bir yöntem olan Çift Gizli Markov Modelini (Pair HMM) dil bilime uyarlamışlardır. Bu model, sözcük çiftlerini yan yana hizalayarak ve sözcük benzerliklerini çözümlenmeye olanak tanımaktadır.

4.3.6. DIATECH

Diyalektoloji araştırmaları için tasarlanmış bir web uygulamasıdır. Bu yöntem, coğrafi alanlar arasındaki lehçe çeşitliliğinin grafiksel temsilini vurgulayarak dilsel verilerin sezgisel olarak anlaşılmasına olanak tanır. Diatech'in birden fazla yanıt çözümlenme yeteneği, çeşitli kaynaklardan veya anketlerden gelen verileri aynı anda işleme ve yorumlama yeteneğini öne çıkarır. Bu uygulama Gabmap ile benzer işleve sahiptir. Uygulama, Visual Dialectometry'de (VDM) kullanılan yöntemlere ve çoklu yanıtların çözümlenmesine odaklanmıştır (Wieling ve Nerbonne, 2015). Diatech karmaşık veri kümelerini işlemesi açısından kullanışlıdır.

4.3.7. DİZE BENZERLİĞİ METRİKLERİ (STRING SIMILARITY METRICS)

Bu metrikler, iki dizinin ne kadar benzer olduğunu değerlendirmek için kullanılan araçlar veya algoritmalarıdır (Tiedemann, 1999).

4.3.8. DİNAMİK FONEM ŞEKİLLENDİRME (DYNAMIC PHONE WARPING) (DPW)

Dinamik Fonem Çarpıtma, ses bilgisi mesafesini ölçmek için kullanılan gelişmiş bir tekniktir (Babu ve diğerleri 2015). Bu algoritma, her bir sesin boğumlanma şekli, boğumlanma yeri ve seslendirme gibi özelliklerin karşılaştırmasını içermektedir.

4.3.9. GABMAP

Gabmap, 2011 yılında Nerbonne ve arkadaşları tarafından geliştirilen bir web uygulamasıdır. Diyalektologlar için diyalektometrik çözümlemeyi basitleştirmek amacıyla oluşturulmuştur. Gabmap çok yönlü bir programdır. Sözcükler arasındaki benzerliğin bir ölçüsü olan düzenleme mesafesini kullanarak transkripsiyonlu telaffuzları karşılaştırma yeteneğine sahiptir. Ayrıca Gabmap, verilerin coğrafi dağılımını belirten bir harita sunar. (Nerbonne ve diğerleri 2011). Gabmap'in ücretsiz olarak erişilebilen, açık kaynaklı bir web uygulamasıdır. Araştırmacılar için dijital dil kaynakları ve araçları sağlayan büyük bir altyapısı sunmaktadır. Gabmap, dijital araçların ve otomatik çözümlemelerin dil araştırmalarını nasıl dönüştürdüğünü örnekleyerek dil çeşitliliği hakkındaki karmaşık soruların daha büyük bir hassasiyet ve anlayışla ele alınmasını mümkün kılmaktadır.

4.3.10. JEFFREYS-MATUSITA MESAFESİ

İki istatistiksel dağılım arasındaki ayrımı ölçen bir metriktir ve özellikle Bhattacharyya mesafesiyle yakın bir ilişkiye sahiptir (Sooful ve Botha, 2002).

4.3.11. KULLBACK-LEİBLER (KL)

İki olasılık arasındaki farkı ölçen bir algoritmadır. İki olasılık arasındaki farklılığın miktarını sayısal olarak belirlemek için kullanılır (Sooful ve Botha, 2002).

4.3.12. LEVENSHTAIN MESAFESİ

Adını 1965 yılında bu kavramı ortaya atan Sovyet matematikçi Vladimir Levenshtein'dan almıştır. Levenshtein mesafesi, bir sözcüğü diğerine dönüştürmek için gereken tek karakterlik düzenlemelerin minimum sayısıdır. İki dizi arasındaki farkın ölçüsüdür. Bu

düzenlemeler tek bir karakterin eklenmesi, çıkarılması veya değiştirilmesinden oluşur. Bu ölçü özellikle yazım denetimi, doğal dil işleme görevleri gibi metin çözümlemesi içeren uygulamalarda kullanışlıdır. İki dizge arasındaki benzerliğin ölçülmesine olanak tanıyan metinsel verilerin karşılaştırılması için bir temel sağlamaktadır. Levenshtein mesafesi hem teorik hem de pratik uygulamalarda çeşitli çalışmalarda kullanılmaktadır. Levenshtein tarafından geliştirilmesi, araştırmacılara ve uygulayıcılara diziler arasındaki benzerliği ölçmek için sağlam bir araç sağlayarak dil bilim, bilgisayar bilimi ve genetik alanlarındaki ilerlemeleri kolaylaştırmıştır (Heeringa, 2004; Kruskal, 1999). Kessler (1995) İrlanda Galcesi'nin lehçeleri arasındaki ses bilgisi mesafesini ölçmek için Levenshtein mesafesini kullanmıştır.

4.3.13. MAHALANOBIS

Bu algoritma, Mahalanobis tarafından geliştirilmiştir. İki ses arasındaki mesafeyi ölçmek için kullanılan bir algoritmadır (Sooful ve Botha, 2002). Mahalanobis mesafesi, özellikle verilerin birbirine ne kadar benzediğini veya bir grup içindeki verilerin merkeze ne kadar yakın olduğunu göstermek için kullanılmaktadır.

4.3.14. RUG/L04

Kleiweg tarafından lehçe farklılıklarının nicel olarak incelenmesi için geliştirilmiş bir yazılım aracıdır (Nerbonne ve diğerleri 2011). Bu yazılım özellikle Levenshtein mesafe algoritmasını kullanmasıyla dikkat çekmektedir. Bu algoritma diyalektometride dilsel mesafeyi ölçmenin yaygın bir yolu olan, bir sözcüğü diğerine dönüştürmek için gereken minimum tek karakterli düzenleme (ekleme, silme veya değiştirme) sayısını belirler. Bu yazılımın güçlü yönlerinden biri çevrim içi olarak kullanılmasıdır. Veriler uyumlu formatta olursa herhangi bir dil için kullanılabilen bir uygulamadır.

4.3.15. PHILOLOGICON

İki dil arasındaki farklılıkları ve benzerlikleri ölçmek için tasarlanmış bir algoritmadır. Bu yöntem, dillerin kendi içindeki sözcük benzerliklerine dayalı matrisler oluşturularak ve bu matrisleri karşılaştırarak diller arası mesafeleri hesaplar (Ellison ve Kirby, 2006).

4.3.16. FONETİK DÜZENLEME MESAFESİ (PHONETIC EDIT DISTANCE) (PED)

Bu yöntem, harflerin artikülasyon özelliklerine dayanarak bir sözcüğün diğerine ne kadar benzediğini belirler. Ses bilgisi mesafesini seslerin benzerliğini dikkate alarak hesaplar (Ahmed ve diğerleri 2022).

4.3.17. METİNSEL İLİŞKİLERİN BELİRLENMESİ YAZILIMI (SOFTWARE FOR THE EXTRACTION OF NARY TEXTUAL ASSOCIATIONS) (SENTA)

Bu yöntem, sözcüklerin sadece yazılış ve anlam benzerliklerine değil, aynı zamanda harflerin bir araya gelme şekillerine de bakarak sözcükler arasındaki benzerlikleri hesaplamaktadır (Ribeiro ve diğerleri 2001).

4.3.18. PÜRÜZSÜZ BİREBİR EŞLEME TANIMLAYICI (THE SMOOTH INJECTIVE MAP RECOGNIZER) (SIMR)

İki farklı dilde yazılmış metinler arasındaki benzerlikleri belirlemek için kullanılan bir algoritmadır. Bu algoritma, bitextlerdeki (iki farklı dildeki metinler arasındaki) karşılıklı noktaları tanımlamak için kullanılan bir algoritmadır. SIMR, bitextler arasında benzerlikleri belirlemek için desen tanıma yaklaşımını kullanır. Bu algoritma, bitextlerin dil bilimsel çözümlenmelerin ve çeviri uygulamalarında kullanılmaktadır (Melamed, 1999).

4.3.19. STOKASTİK DÖNÜŞTÜRÜCÜ (STOCHASTIC TRANSDUCER)

Ristad ve Yianilos tarafından 1998 yılında geliştirilmiş bir programdır. Bu programda, bir sözcüğün diğerine nasıl dönüştürülebileceği (örneğin bir harfin değiştirilmesi, eklenmesi ya da çıkarılması gibi) ve bu dönüşümlerin olasılığı hesaplanmaya çalışılmıştır. Her bir dönüşüm işlemi (ekleme, çıkarma, değiştirme) için bir olasılık değeri belirlenmiştir. Daha sonra, bu olasılıklar kullanılarak iki sözcük karşılaştırıp aralarındaki benzerliği veya farklılığı puanlayan bir sistem (stokastik transdüser) oluşturulmuştur. Stokastik Dönüştürücü, öncelikle hesaplamalı dil bilim, doğal işleme ve konuşma tanıma alanlarında kullanılan bir modelidir. Bu araç, bir dildeki sözcükleri veya sesleri alıp bunları başka bir dildeki sözcükler veya seslere dönüştürebilir (Kondrak ve Sherif, 2006).

4.3.20. GÖRSEL DİYALEKTOMETRİ (VISUAL DIALECTOMETRY) (VDM)

Salzburg Üniversitesi'nden Hans Goebel tarafından geliştirilen bir araçtır. Dil bilimsel verilerin coğrafi dağılımını çözümlmek ve görselleştirmek için kullanılır. VDM'nin temel amacı farklı coğrafi bölgelerdeki lehçe çeşitliliğinin görsel bir temsilini sağlamaktır. VDM, belirli dilsel özelliklerin coğrafi bölgelerdeki dağılımını görsel olarak temsil eden haritalar oluşturmayı içerir. Bu özellikler telaffuz, söz dizimi ve dilin bölgesel olarak değişebilen diğer yönlerini içerebilir. Yazılım, farklı lehçelerden toplanan ses, sözcük veya söz dizim bilgilerini içeren dilsel verileri çözümlmek için istatistiksel yöntemler kullanır. VDM yazılımının temel amacı, lehçelerin farklı bölgeler arasında nasıl değiştiğini gösteren haritalar ve grafikler oluşturmaktır. Karmaşık dil bilimsel verileri grafikler ve haritalar gibi görsel formatlara dönüştürerek araştırmacıların verilerdeki örüntüleri ve ilişkileri görmelerine yardımcı olur. VDM, coğrafi sınıflandırma kullanarak farklı coğrafi bölgelerdeki dilsel çeşitlilikteki kalıpları, kümeleri veya eğimleri belirleyebilir. Bu, lehçelerin zaman içinde nasıl geliştiğini ve yayıldığını anlamaya yardımcı olur (Wieling ve Nerbonne, 2015).

4.3.21. ALGISAL YÖNTEM İLE HESAPLAMALI YÖNTEMLERİ KARŞILAŞTIRAN TEKNİK

Klatt (1982), bilgisayarda ürettikleri yapay sesleri insanlara dinletmiştir. Katılımcılar seslerin kendilerine ne kadar farklı göründüğünü değerlendirmişlerdir. Araştırmacı daha sonra istatistiksel işlemler ile iki ses arasındaki mesafeyi ölçmüştür. İnsanların algıladıkları mesafe ile hesaplamalı işlemler sonucunda elde ettikleri verileri karşılaştırmıştır.

5. SONUÇ

19. yy. dan beri diller arasındaki ses bilgisi mesafesini ölçmek için birçok çalışmanın yapıldığı görülmektedir. Bu çalışmalar conyöntemlerine göre sınıflandırıldığında, diller arasındaki ses bilgisi mesafesini ölçmek için üç farklı yöntemin kullanıldığı belirlenmiştir: Geleneksel yöntemler, algısal yöntemler ve hesaplamalı yöntemler. Geleneksel yöntemlerin büyük ölçüde araştırmacıların kişisel bilgilerine, dillerle ilgili deneyimlerine ve tarihi ya da edebi kaynaklara dayanır. Geleneksel yöntemlerde veri toplamak için anketlerden ve

mülakatları içeren saha çalışmalarından yararlanılmıştır. Ayrıca, farklı dil özelliklerinin hâkim olduğu alanları tanımlamak ve dil sınırlarını belirtmek için izogloslar kullanılmıştır.

Algısal yöntemler, geleneksel yöntem analizi ile konuşmacıların kendi dillerini ve diğer dilleri nasıl algıladıklarına dair düşüncelerini birleştiren teknikleri içermektedir. Algısal yöntemlerde, ok yöntemi ve derecelendirme ölçeklerinden yararlanılmıştır. Ok yöntemi, katılımcıların dillerdeki benzerlikleri ve farklılıkları nasıl algıladıklarına dayalı olarak lehçe sınırlarını haritalandırmaktadır. Derecelendirme ölçeklerinde, katılımcılar lehçelerin bir standarda veya birbirlerine benzerliğini derecelendirerek algısal mesafenin nicel bir ölçüsünü sağlamaktadır.

Hesaplamalı yöntemler, ses bilgisi ve söz dizimi verilerini içeren büyük veri kümelerini çözümlen yöntemlerdir. Hesaplamalı yöntemlerde, haritalamadan, derlem analizinden ve teknolojiyen yararlanılmıştır. Haritalama, verileri görsel olarak temsil etmek ve bölgeler arasındaki dil farklılıklarını vurgulamak için coğrafi haritalama araçlarını kullanmaktadır. Derlem analizi, kapsamlı veri kümelerini işlemek ve çözümlmek, dil kullanımındaki kalıpları ve eğilimleri belirlemek için tasarlanmış yazılım araçlarını kullanmaktadır. Ses bilgisi mesafesini ölçümünü geliştirmek için yeni yazılımlar ve araçlar geliştirilerek bu çözümlmeler daha hassas hale getirilmiştir.

Hesaplamalı yöntemlerde, ALINE, Bhattacharyya mesafesi, CORDI, Covington, Pair HMM, Diatech, String similarity metrics, Dynamic phone warping, Gabmap, Jeffreys-Matusita mesafesi, Kullback-Leibler, Levenshtein mesafesi, Mahalanobis, RUG/L04, PHILOLOGICON, Phonetic edit distance, SENTA, SIMR, Stochastic Transducer ve Visual Dialectometry programlarının kullanıldığı tespit edilmiştir.

Geleneksel yöntemler insan sezgisine dayanırken algısal yöntemler bilişsel algılar ile geleneksel yöntemler arasındaki boşluğu doldurmuştur. Hesaplamalı yöntemlerde ise veri işlemeye odaklanılmış ve teknolojik gelişmelerden yararlanılmıştır.

Ses bilgisi mesafesini ölçen üç yöntemin de sınırlılıkları olduğu görülmüştür. Geleneksel yöntem, veri toplama ve çözümlme aşamalarında öznellikten yanadır. Bu durum, daha modern yöntemlere kıyasla geleneksel yöntemi tekrarlanabilirlik açısından sınırlı kılmaktadır. Algısal yöntemlerde, veri toplama ve çözümlme aşamaları öznel algılara dayandığı için ses bilgisi mesafesini etkileyen

tüm unsurları çözümlene açısından sınırlıdır. Hesaplamalı yöntemler, verilerin yorumlanması aşamasında yetersiz kalabilir. Diller arasındaki ses bilgisi mesafesini etkileyen ve dil kullanımını şekillendiren sosyo-kültürel bağlamın yorumlanması aşamasında hesaplamalı yöntemler sınırlı kalacaktır.

6. GENEL SONUÇ

Ses bilgisi mesafe araçları çok yönlü bir metodolojiyi içermektedir. Bu çok yönlü yaklaşım, yalnızca dilsel çeşitlilik konusundaki anlayışı geliştirmekle kalmaz, aynı zamanda dil çalışmalarının teknik, sosyal ve tarihsel boyutları arasındaki boşlukları da kapatır. Bu yöntemlerin birleştirilmesi ses bilgisi mesafelerinin daha bütünsel bir görünümünü sağlayacaktır. Örneğin, büyük veri kümelerini çözümlenmek için hesaplamalı yöntemler kullanmak, ardından bu kalıpların konuşmacılar tarafından nasıl algılandığını belirlemek için algısal yöntemler uygulamak, sosyo-kültürel bağlamın yorumlanması için geleneksel yöntemler kullanmak diller arasındaki ses bilimsel mesafeleri bir bütün olarak ortaya koyacaktır.

Teknolojideki ilerlemeler hesaplamalı yöntemlerin hassasiyetini ve kapsamını artırmaya devam edecektir. Araştırmacılar, her bir yöntemin güçlü yönlerinden yararlanarak daha sağlam ve kapsamlı ses bilgisi mesafe çalışmaları yapabilirler, bu da dil çeşitliliği hakkında daha zengin veriler elde edilmesini sağlayabilir. Hesaplamalı yöntemlerin yapay zekâ ile birleştirilmesi, elde edilen ses bilgisi mesafe verilerinin daha kapsamlı işlenmesine imkân sağlayacaktır.

KAYNAKÇA

- Ahmed, T., Suffian, M., Khan, M. Y. & Bogliolo, A. (2022). Discovering lexical similarity using articulatory feature-based phonetic edit distance. *IEEE Access*, 10, 1533-1544.
- Babu, A. A., Yellasiri, R. & Rao, A. A. (2015). Phonetic distance based accent classifier to identify pronunciation variants and oov words. *Signal & Image Processing: An International Journal (SIPIJ)*, 6(4), 33-46.
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27-40.
- Campbell Collaboration, (2024). What is a systematic review? <https://www.campbellcollaboration.org/what-is-a-systematic-review.html>.
- Chambers, J. K. & Trudgill, P. (1998). *Dialectology*. Cambridge, United Kingdom: Cambridge University Press.

- Covington, M. A. (1998). Alignment of multiple languages for historical comparison. *In 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*, 1, 275–279.
- Daan, J. & Blok, D. P. (1969). Van randstad tot landrand; toelichting bij de kaart: dialecten en naamkunde, volume XXXVII of bijdragen en mededelingen der dialectencommissie van de koninklijke nederlandse akademie van wetenschappen te amsterdam. Noord-Hollandsche Uitgevers Maatschappij, Amsterdam.
- Droppo, J. & Acero, A. (2010). Context dependent phonetic string edit distance for automatic speech recognition. *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, Dallas, TX, USA, 2010, 4358-4361.
- Eden, S. E. (2018). *Measuring phonological distance between languages*. PhD thesis. Department of Linguistics.
- Ellison, T.M., & Kirby, S. (2006). Measuring language divergence by intra-lexical comparison. *Annual Meeting of the Association for Computational Linguistics*. 273-278.
- Goebel, H. (2006). Recent advances in salzburg dialectometry. *Literary and Linguistic Computing*, 21(4), 411–436.
- Jellinghaus, H. (1892). *Die Niederländischen Volksmundarten; nach den Aufzeichnungen der Niederländer*. D. Soltau's Verlag, Norden and Leipzig.
- Kang, S. S. (2015). Word similarity calculation by using the edit distance metrics with consonant normalization. *Journal of Information Processing Systems*, 11(4), 573-582.
- Karasar, N. (2005). *Bilimsel araştırma yöntemi*. Nobel Yayın Dağıtım.
- Kessler, B. (1995). Computational dialectology in irish gaelic. *In Proceedings of the 7th Conference of the European Chapter of the Association for computational Linguistics*, Dublin. EACL, 60–67.
- Kessler, B. (2005). Phonetic comparison algorithms. *Transactions of the philological society*, 103(2), 243-260.
- Kisler, T. & Reichel, U. D. (2013). A dialect distance metric based on string and temporal alignment. *Elektronische Sprachsignalverarbeitung ESSV*, 58-165.
- Klatt, D. (1982). Prediction of perceived phonetic distance from critical-band spectra: A first step. *ICASSP 82. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Paris, France, 1278-1281.
- Gooskens, C. (1997). *On the role of prosodic and verbal information in the perception of dutch and english varieties*. PhD thesis, Nijmegen: Katholieke Universiteit Nijmegen.
- Gooskens, C. (2005). Travel time as a predictor of linguistic distance. *DiG 13*, 38-62.
- Gooskens, C., & Heeringa, W. (2004). Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data. *Language Variation and Change*, 16, 189-207.
- Heeringa, W. J. (2004). *Measuring dialect pronunciation differences using levenshtein distance*. Thesis fully internal (DIV), University of Groningen.
- Heeringa, W., Johnson, K. & Gooskens, C. (2009). Measuring norwegian dialect distances using acoustic features. *Speech Communication*, 51, 167-183.
- Hoppenbrouwers, C. & Hoppenbrouwers, G. (2001). *De indeling van de nederlandse streektalen. Dialecten van 156 steden en dorpen geklasseerd volgens de FFM*. Koninklijke Van Gorcum B.V., Assen.
- Kondrak, G. (2000). A new algorithm for the alignment of phonetic sequences. *Applied Natural Language Processing Conference*.
- Kondrak, G. (2001). Identifying cognates by phonetic and semantic similarity. *North American Chapter of the Association for Computational Linguistics*, 1-8.

- Kondrak, G. (2002). Determining recurrent sound correspondences by inducing translation models. *In Proceedings of COLING*, 488–494.
- Kondrak, G. (2003). Phonetic alignment and similarity. *Computers and the Humanities* 37, 273–291
- Kondrak, G., ve Dorr, B. (2006). Automatic identification of confusable drug names. *Artificial Intelligence in Medicine*, 36(1), 29–42.
- Kondrak, G. ve Sherif, T. (2006). Evaluation of several phonetic similarity algorithms on the task of cognate identification. *In Proceedings of the Workshop on Linguistic Distances*, Sydney, Australia: Association for Computational Linguistics, 43–50.
- Kruskal, J. B. (1999). An overview of sequence comparison. *Time Warps, String edits, and Macromolecules. The Theory and Practice of Sequence Comparison*, CSLI, Stanford, 1–44.
- Kudare, J., Georgis, P., Möbius, Avgustinova, T. & Klakow, D. (2021). Phonetic distance and surprisal in multilingual priming: Evidence from slavic. *Interspeech*, 3944–3948.
- Levenshtein, V. (1965). Binary codes capable of correcting deletions, insertions and reversals. *Doklady Akademii Nauk SSSR*, 163(4):845–848.
- Mackay, W. & Kondrak, G. 2005. Computing word similarity and identifying cognates with Pair Hidden Markov Models. *In Proceedings of the 9th Conference on Computational Natural Language Learning (CoNLL)*, 40–47.
- MacLeod, B. (2021). Problems in the difference-in-distance measure of phonetic imitation. *Journal of Phonetics*, 87, 1–21.
- Mak, B. K., & Barnard, E. (1996). Phone clustering using the Bhattacharyya distance. *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP '96*, 4, 2005–2008, 4.
- Melamed, I. D. (1999). Bitext Maps and alignment via pattern recognition. *Association for Computational Linguistics*, 25(1), 107–130.
- Mulloni, A. & Pekar, V. (2006). Automatic detection of orthographic cues for cognate recognition. *Proceedings of the Fifth International Conference on Language Resources and Evaluation*, 2387–2391.
- Nerbonne, J., Colen, R., Gooskens, C., Kleiweg, P. ve Leinonen, T. (2011). Gabmap- A Web Application for Dialectology. *Dialectologia*, 1–23.
- Nerbonne, J. & Heeringa, W. (1997). Measuring dialect distance phonetically. *In Computational Phonology: Third Meeting of the ACL Special Interest Group in Computational Phonology*, 11–19.
- Nerbonne, J., Heeringa, W., Hout, E.V., Kooi, P.V., Otten, S., & Vis, W.V. (1996). Phonetic distance between dutch dialects. 1–15.
- Och, F. J. & Ney, H. (2003). A systematic comparison of various statistical alignment models. *Computational Linguistics*, 29(1), 19–51.
- Pucher, M., Türk, A., Ajmera, J., & Fecher, N. (2007). *Phonetic distance measures for speech recognition vocabulary and grammar optimization*. Computer Science, Linguistics.
- Rensink, W. G. (1955). *Dialectindeling naar opgaven van medewerkers*. Mededelingen der Centrale Commissie voor Onderzoek van het Nederlandse Volkseigen, 7:20–23.
- Ribeiro, A., Dias, G., Lopes, G. & Mexia, J. (2001). Cognates alignment. *In Proceedings of Machine Translation Summit VIII*, Spain, 18–22 Sem.
- Ristad, E. S. & Yianilos, P. N. (1998). Learning string-edit distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(5), 522–532.
- Sé guy, J. (1973). La dialectometrie dans l'atlas linguistique de la gascogne. *Revue de Linguistique Romane* 37, 1–24.

- Sencer, M. & Sencer, Y. (1978). *Toplumsal arařtırmalarda yöntem bilim*. Ankara: Türkiye ve Orta Doęu Anne İdaresi Enstitüsü Yayını.
- Sooful, J. J., & Botha, E. C. (2002). An acoustic distance measure for automatic cross-language phoneme mapping. *7 th International Conference on Spoken Language Processing Denver, Colorado*, 16-20.
- Tiedemann, J. (1999). Automatic construction of weighted string similarity measures. *Conference on Empirical Methods in Natural Language Processing*. 213-220.
- Wagner, R. & Fischer, J. (1974). The string-to-string correction problem. *Journal of the Association for Computing Machinery*, 21(1), 168-173.
- Weijnen, A. (1946). De grenzen tussen de oost-noord-Brabantse dialecten on-derling. *Bijdragen en Mededeelingen der Dialectcommissie van de Koninklijke Adademie van Wetenschappen te Amsterdam*, 8:1-15.
- Wieling, M. & Nerbonne, J. (2015). Advances in dialectometry. *Annu. Rev. Linguist*, 1, 243-264.
- Winkler, J. (1874). *Algemeen Nederduitsch en Friesch Dialecticon*. Martinus Nijhoff, 's-Gravenhage.