

# Türkmence Morfolik Analiz Motoru

Esma AYDIN

## Giriş

Türkmen Türkçesi, eklemeli (agglutinative) yapısı ve fonetik uyum yasalarıyla, bilgisayar bilimlerindeki "Finite State Automata" (Sonlu Durum Makineleri) mantığına doğal bir yatkınlık sergiler. Bir kökten düzinelere sözcük üretilirken ses uyumuna göre anlamın evrilmesi, rastlantısal bir süreç değil; hiyerarşik bir algoritmanın sonucudur.

Bu çalışma, Türkmen Türkçesinin standart yazı dilindeki isim çekim süreçlerini dijital bir ortama aktaran bir morfoloji motorunun geliştirilmesini amaçlamaktadır. Çalışmanın metodolojisi, statik bir veri listesi kullanmak yerine, dilin fonetik kurallarını dinamik fonksiyonlara dönüştürmeye dayanmaktadır. Özellikle "San" (Sayı), "Degişlilik" (İyelik) ve "Düşüm" (Hal) kategorileri arasındaki hiyerarşik ilişki ve bu ekleşme sırasında meydana gelen "Orta Hece Yuvarlaklaşması" (ogry > ogrular) ile "Ünlü Düşmesi" (garyn > garnyň) gibi olaylar simüle edilmiştir.

Türkmen Türkçesi, bünyesinde barındırdığı asli uzun ünlüler, orta hece yuvarlaklaşması ve vokal uzaması gibi olaylar nedeniyle Türk dilleri arasında morfolojik açıdan en karmaşık yapılardan birine sahiptir. Proje kapsamında, bu dil bilgisi kuralları Python programlama dili kullanılarak algoritmik bir düzlemde modellenmiştir.

Bu çalışma, Türkmen Türkçesinin morfolojik yapısını statik bir veri tabanı yerine, Kural Tabanlı Doğal Dil İşleme (Rule-Based NLP) prensipleriyle modellemektedir. Sistem bu mantıksal hiyerarşiyi temel alarak, her girdi üzerinde ardışık fonotaktik denetimler gerçekleştiren bir Sonlu Durum Dönüşürücüsü (Finite State Transducer) mantığıyla çalışır.

## Metodoloji

### • Veri Yapısı ve Fonem Envanteri

Motorun temelini oluşturan fonetik sınıflandırma, Türkmen Türkçesinin 9 ünlülü sisteminin ayırt edici özelliklerine göre düzenlenmiştir:

özellik küme Dilbilimsel karşılığı Kalın – ince Yogyn = {"a", "o", "u", "y"} ince = {"e", "a", "a", "ö", "u", "ü"} Palatal harmony Düz – yuvarlak Yuvarlak = {"o", "ö", "u", "ü"}

düz = {"a", "e", "y", "i"} Labial Harmony Dar – geniş Dar = {"y", "i", "u", "ü"} geniş = {"a", "a", "e", "o", "ö"} Vowel height · Kalın-İnce Ayrımı: yogyn = {"a", "o", "u", "y"} ve ince = {"e", "ä", "ö", "i", "ü"} kümeleri ile tanımlanmıştır. Bu ayrılmış, ünlü uyumunun belkemiğini oluşturur.

Özellik	Küme	Dilbilimsel karşılığı
Kalın - ince	Yogyn = {"a", "o", "u", "y"} Ince = {"e", "ä", "ö", "i", "ü"}	Palatal harmony
Düz - yuvarlak	Yuvarlak = {"o", "ö", "u", "ü"} düz = {"a", "e", "y", "i"}	Labial Harmony
Dar - geniş	Dar = {"y", "i", "u", "ü"} geniş = {"a", "ä", "e", "o", "ö"}	Vowel height

#### • Algoritmik Süreç ve Dilbilimsel İşlevler

Geliştirilen motor, isim çekimini San (Sayı) → Degişlilik (İyelik) → Düşüm (Hal) hiyerarşisi içinde işletir. Her aşama, kendinden önceki aşamanın çıktısını girdi olarak alır ve dilbilgisel kuralları uygular.

#### • Ünlü Niteliği Tespit Fonksiyonu (unlu\_niteligi) Dilbilimsel Karşılığı: Kalınlık-İncelik Uyumu (Palatal Harmony)

Fonksiyon, bir sözcüğün son ünlüsünü bularak sözcüğün kalın veya ince ünlülü olduğunu belirler. Bu işlem, eklerin hangi ünlü varyantını alacağını tayin eder. Türk dillerinde ekler, kökün son ünlüsünün niteliğine göre şekillenir; bu fonksiyon, işte bu dilbilimsel ilkeyi algoritmik düzlemden karşılar.

#### • Yuvarlaklık Tespit Fonksiyonu (yuvarlak\_mi) Dilbilimsel Karşılığı: Düzlük-Yuvarlaklık Uyumu (Labial Harmony)

Sözcükte yuvarlak ünlü bulunup bulunmadığını denetler. Özellikle iyelik eklerinin oluşumunda (ym/im, um/üm ayrımı) ve "Orta Hece Yuvarlaklaşması" sürecinde belirleyici rol oynar. Bu fonksiyon, dildeki yuvarlaklaşma eğiliminin matematiksel ifadesidir.

#### • Ünsüz Yumuşaması (tam\_yumusama) Dilbilimsel Karşılığı: Ötümlüleşme (Lenition) / Ünsüz Benzeşmesi

Türkmen Türkçesinde sözcük sonundaki p, ç, t, k ünsüzleri, ünlü ile başlayan bir ek alındıklarında sırasıyla b, j, d, g ünsüzlerine dönüşür. Bu fonksiyon, söz konusu dönüşümü gerçekleştirir. Örneğin:

kitap + y → kitaby. Fonksiyon, bu ses olayını otomatikleştirir.

- **Ünlü Düşmesi (dusme\_ugula) Dilbilimsel Karşılığı: Syncope (Orta Hece Ünlü Düşmesi)**

İki heceli ve ikinci hecesi dar ünlü (y, i, u, ü) ile biten sözcükler, ünlü ile başlayan bir ek alındıklarında ortadaki dar ünlüyü kaybeder. Bu fonksiyon:

1. Öncelikle sözcüğün istisnalar listesinde olup olmadığını kontrol eder.
2. Ardından ünlü düşmesi adayları arasında olup olmadığını denetler.
3. Koşullar sağlanıyorsa, sözcüğün son iki harfinden ilkini atarak (örneğin: garyn → garn) ünlü düşmesini uygular.

Bu işlem, Türk dillerinde yaygın görülen bir fonetik sürecin algoritmik modellemesidir.

- **Dudak Uyumu (Labial Harmony) Kuralları**

Türkmen yazı dilinin dudak uyumuyla ilgili özel kuralları algoritmaşa şu şekilde yansıtılmıştır:

Kural 1: Geniş-yuvarlak ünlüler (o, ö) sadece ilk hecede bulunur. Motor, bu ünlülerin sonraki hecelerde türemesini engelleyecek şekilde ek seçimi yapar.

Kural 2: Dar-yuvarlak ünlüler (u, ü) ikinci hecede bulunabilir, ancak üçüncü heceden itibaren düz ünlülere (y, i) dönüşür. Bu kural, özellikle çok heceli iyelik eklerinin seçiminde belirleyicidir göz + üm gözüm gözüm 2. hecede yuvarlak kalır bozguç + umyz bozguçumyz bozguçumyz 3. hecede düzleşir

Kural 3: Kelime sonunda u, ü ünlüler bulunmaz. Bu kural, özellikle 3. tekil şahıs iyelik ekinin ve bazı isim köklerinin yazımında kendini gösterir:

Kural 4: Sonu y, i ile biten kelimeye çokluk, iyelik, bulunma ve ayrılma ekleri geldiğinde, son ünlü yuvarlaklaşır:

- **Orta Hece Yuvarlaklaşması Dilbilimsel Karşılığı: Labial Assimilation in Closed Syllables**

Bu ses olayı iki aşamada modellenmiştir:

1. Çokluk eki öncesi: Kök yuvarlak ünlü içeriyorsa ve son hecede dar düz ünlü (y veya i) varsa, bu ünlü yuvarlaklaşır (ogry → ogru + lar → ogrular). Bu dönüşüm, cokluk koşulunda gerçekleştirilir.

2. İyelik eki sonrası hal eklerinde: 3. şahıs iyelik eki almış sözcüklerde, kök yuvarlak ünlü içeriyorsa ve son ünlü dar düz ise, yuvarlaklaşma uygulanır (burny → burnu + nyň → burnunyň).

Bu modelleme, dildeki yuvarlaklaşma eğiliminin hem morfolojik hem de fonetik bağlamını yansıtır.

- Ek Çekim Sistemi ve Hiyerarşik Yapı

Motorun temel işlevi isim\_cekimle(kok, cokluk, iyelik, i\_tip, hal) parametreleriyle çalışır ve aşağıdaki hiperarşik adımları izler:

## **1- San (Sayı) Kategorisi**

Türkmen Türkçesinde sayı kategorisi, kelime köküne hal ve iyelik eklerinden önce eklenen birincil morfolojik katmandır. Geliştirilen motor, bu aşamada sadece ek eklemekle kalmayıp, Praktikum'da belirtilen şu fonetik kuralları da işletmektedir:

- Büyük Ünlü Uyumu: Algoritma, kelimenin son hecesini analiz ederek -lar (kalın) veya -ler (ince) seçimini otomatik olarak yapar.

Adam+ köplük > adamlar El + köplük > eller

- Orta Hece Dudaklaşması (Yuvarlaklaşma): İlk hecesi yuvarlak ünlü (o, ö, u, ü) barındıran ve dar-düz ünlüyle (y, i) biten iki heceli köklerde, çokluk eki eklendiğinde kök sonundaki ünlü yuvarlaklaşır.

guzy + lar guzular (y > u dönüşümü) süri + ler sürüler (i > ü dönüşümü)

## **2- Degişlilik (İyelik) Kategorisi**

İyelik ekleri tablodaki şekliyle tanımlanmıştır:

D1b: 1. Birlik Degişlilik	D1k: 1. Köplük Degişlilik
D2b: 2. Birlik Degişlilik	D2k: 2. Köplük Degişlilik
D3b: 3. Birlik Degişlilik	D3k: 3. Köplük Degişlilik

Ek seçiminde dört temel parametre rol oynar:

- Kökün son harfinin ünlü/ünsüz oluşu
- Kökün ünlü niteliği (kalın/ince)
- Kökte yuvarlak ünlü bulunup bulunmadığı
- İyelik ekinin şahıs ve sayısı

Bu çok değişkenli seçim sistemi, Türkmen Türkçesinin iyelik paradigmاسının karmaşıklığını yansıtır. Örneğin, 1. şahıs tekil iyelik eki (D1b):

- Ünlüyle biten köklerde: -m · Ünsüzle biten köklerde: yuvarlaklık durumuna göre -um/-üm (yuvarlak) veya -ym/-im (düz)

### 3 - Düşüm (Hal) Kategorisi

Hal ekleri (A1: Yalın, A2: İlgi, A3: Yönelme, A4: Belirtme, A5: Bulunma, A6: Ayrılma) uygulanırken en kritik belirleyici, "n-kaynaştırması" koşuludur. Bu, 3. şahıs iyelik ekinden sonra gelen hal eklerinin önüne "n" ünsüzünün gelmesi kuralıdır (n\_kay = iyelik == "A3").

Bu koşul, dilbilgisel birincil ve ikincil ekler arasındaki ayrimı yansıtır: İyelik ekleri (birincil) ile hal ekleri (ikincil) arasında kaynaştırma ünsüzü türemesi, Türk dillerinin morfolojik yapısının ayırt edici bir özelliğiidir.

İlgi Hali (A<sup>2</sup>) Özel Durumları İlgi hali eki seçiminde dört parametre rol oynar:

1. n-kaynaştırması varsa: -nyň / -niň
2. Ünlü ile biten kökler: -nyň / -niň
3. Ünsüz ile biten kısa kökler ( $\leq 4$  harf): Yuvarlak ünlülü ise -uň / -üň, değilse -yň / -iň
4. Ünsüz ile biten uzun kökler: -yň / -iň

Örnekler:

- at + A<sup>2</sup> → atyň · göz + A<sup>2</sup> → gözüň · defter + A<sup>2</sup> → defteriň · alma + A<sup>2</sup> → almanyň
- Yönelme Hali (A<sup>3</sup>) Ses Değişimleri Yönelme hali, Türkmen Türkçesinde

önemli ses değişimlerine yol açar:

### Kök Sonu Değişim

#### Örnek

a - a (uzun okunur, değişmez) alma + A<sup>3</sup> → alma

e - ä şecere + A<sup>3</sup> → şecerä y - a gapy + A<sup>3</sup> → gapa

i - ä üzümçi + A<sup>3</sup> → üzümçä

### Belirtme Hali (A<sup>4</sup>) Özel Durumları

· Ünlü ile biten kökler: -ny / -ni · Ünsüz ile biten kökler: -y / -i Örnekler:

- at + A<sup>4</sup> → aty
- defter + A<sup>4</sup> → defteri
- alma + A<sup>4</sup> → almany

### Bulunma (A<sup>5</sup>) ve Ayrılma (A<sup>6</sup>) Halleri

Bu hallerde ek seçimi basittir: n-kaynaştırması varsa -nda / -nde (bulunma) veya -ndan / -nden (ayırılma); yoksa -da / -de veya -dan / -den.

Hal eklerinin seçiminde ayrıca:

- Kökün ünlü/ünsüz ile bitisi
- Kökün ünlü niteliği
- Kısa köklerde ( $\leq 4$  harf) yuvarlaklaşma eğilimi

gibi parametreler de dikkate alınır.

### Programlama Dili ve Algoritmik Yapı

Motor, Python 3 programlama diliyle geliştirilmiştir. Python'un dinamik tip sistemi, string işlemefonksiyonları ve küme (set) veri yapısı, dilbilimsel kategorilerin modellenmesinde esneklik sağlamıştır. Algoritmanın işleyışı şu prensiplere dayanır:

1. Kuralların Öncelik Sırası: İstisnalar > Düzenli Kurallar > Varsayılan Durum 2. Bilgi Akışı:  
Her morfolojik işlem, bir sonraki işlem için girdi oluşturur.

3. Geri İzlenebilirlik: yol listesi, her ekleme adımını kaydederek şecere çıktısı üretir. Bu, hem dilbilimsel analiz hem de hata ayıklama için kullanışlıdır.

Bu çalışma, Türkmen Türkçesinin algoritmik yapısını dijital bir prototip ile somutlaştırmıştır. Mevcut motorun başarısı, projenin daha kapsamlı bir hesaplamalı analiz sistemine dönüşmesi için bir temel teşkil etmektedir.

Gelecek aşamalarda projenin şu üç ana yönde geliştirilmesi hedeflenmektedir:

#### Morfolojik Analizatör (Parser) Modülü

Mevcut "kelime üreten" mantığın tersine çevrilmesiyle oluşturulacak bu modül, bir "çözümleyici" olarak görev yapacaktır. Kullanıcı tarafından girilen karmaşık kelime yapılarını (örn: burnuňyzdan) saniyeler içinde kök ve eklerine (burun + y + ſyz + dan) ayırtılabilecektir. Bu süreçte sistem, kelimenin uğradığı ünlü düşmesi veya yuvarlaklaşma gibi fonetik değişimleri geriye dönük olarak takip ederek kullanıcısına kelimenin dil bilgisi haritasını sunacaktır.

#### Fiil Çekimi ve Cümle Yapısı Entegrasyonu

İsim çekimlerinde elde edilen fonetik başarı, projenin Fiil (İşlik) Çekimi modülüne aktarılacaktır. Türkmen Türkçesindeki zaman, şahıs ve kip ekleri, isim çekimlerindeki hiyerarşik mantıkla sisteme dahil edilecektir.

Projenin temel vizyonu, Türkmen Türkçesinin bu algoritmik yapısını herkesin kullanımına açmaktadır. Hazırlanan algoritmaların açık kaynaklı bir kütüphane haline getirilmesiyle; dil araştırmacıları, öğrenciler ve yapay zeka geliştiricileri için standart bir Hesaplamalı Türkoloji Kaynağı oluşturulması hedeflenmektedir.