

Kitap Tanıtımı / Book Review

A Frequency Dictionary of Turkish: Core Vocabulary for Learners.
Yeşim Aksan, Mustafa Aksan, Ümit Mersinli ve Umut Ufuk
Demirhan, 2017. Londra/New York: Routledge. 349 sayfa. ISBN 978-
1-138-83965-6

Özgür Aydın

Ankara Üniversitesi Dilbilim Bölümü 06100 Sıhhiye, Ankara
ozguraydin66@gmail.com

(Gönderilme tarihi 10 Aralık 2017; Kabul edilme tarihi 25 Aralık 2017)

1 Giriş

Sözlüksel anlam ve bağlam ilişkisi, dilbilgisi, derlem ve *berim* (computation) gibi konularında, dilbilim alanında en önde gelen bilim insanlarından biri olan Charles J. Fillmore, “derlem dilbilimi” ile “masa başı dilbilimini” karşılaştırdığı yazısında (Fillmore, 2008), gözleri kapalı biçimde ellerini başının arkasında kenetleyip dilsel gerçekleri keşfetmeyi bekleyen masa başı dilbilimcisiyle, milyonlarca sözcüğün içinde kaybolan derlem dilbilimcisinin aslında birbirlerine ne kadar da gereksinimi olduğunu ortaya koymaktadır. Fillmore’un vurguladığı gibi, ne kadar geniş olursa olsun, keşfetmek istediğimiz sözlüksel ya da dilbilgisel *tüm* konular hakkında bilgi sunabilecek bir derlem olmasa bile, aslında ne kadar küçük olursa olsun, bir derlemin tasavvur edemeyeceğimiz gerçekleri su yüzüne çıkarma olasılığı bulunmaktadır. İşte bu nedenle, özellikle son yirmi yıldır, sadece uygulama amaçlı çalışmalarda değil, kuramsal çalışmalarda da derlem dilbilimi önem kazanmaya başlamıştır.

Derlem dilbilimi kapsamında sürdürülen çalışmaların bir türünü de sözcük sayımı ve sıklığı çalışmaları oluşturmaktadır. Sözcüğün kullanım sıklığı, doğrudan sözcüğün içsel bir özelliği olmasa bile, pek çok dilbilimsel çalışmada ya da uygulamada, karar alma süreçlerini doğrudan etkileyebilmektedir. Söz gelimi, ikinci ya da yabancı dil öğretimine/öğrenimine yönelik uygulamalarda, doğal dil işleme uygulamalarında, sözlükbilim, sözcük anlambilim gibi alanlarında, sıklık verisi kullanılabilir. Sıklık verisi ayrıca, ruhdilbilim, sinirdilbilim gibi alanlarda yapılacak deneysel çalışmalar için de önemli bir bilgi kaynağı sunmaktadır.

Bu yazıda, Yeşim Aksan, Mustafa Aksan, Ümit Mersinli ve Umut Ufuk Demirhan tarafından yayımlanan *A Frequency Dictionary of Turkish* adlı

<http://dx.doi.org/10.18492/dad.364401>

Dilbilim Araştırmaları Dergisi, 2017/2, 93-102.

© 2017 Boğaziçi Üniversitesi Yayınevi, İstanbul.

sözlük çalıřması, genel çizgileriyle tanıtılmaktadır. *Routledge Frequency Dictionaries* serisi içinde yayımlanan Türkçe sıklık sözlüğü çalıřması, aynı seride daha önce yayımlanmış olan Almanca, Arapça, Çekçe, Çince, Fransızca, Hollandaca, İngilizce, İspanyolca, Japonca, Korece, Portekizce, Rusça ve çok yakın bir tarihte yayımlanan Farsça sıklık sözlüklerinin son halkasını oluşturmaktadır.

2 Türkçede Sıklık Çalıřmaları

Türkçede sözcüklerin görünüm sıklıkları konusundaki çalıřmalar 1930'ların sonlarına kadar gitmektedir. İlk olarak Aksoy (1936), Türkçe eğitimi alanında kullanılmak üzere, Thorndike'nin (1921) İngilizce eğitimi için hazırlamış olduđu sözcük sıklığı listesinin¹ ilk 2,500 sözcüğüyle, Türkçe yazılmış üç farklı Türkçe yapıta dayanan 15,000 sözcüklük sıklık listesini karşılařtırmıştır. Ancak daha ayrıntılı çalıřmalar hemen hemen otuz yıl sonra Joe E. Pierce tarafından gerçekleştirilmiştir. Pierce'in 1957 Eylül'ünde başlayıp 1960 Kasım'ında tamamladıđı derleme dayanarak Türkçede eklerin ve sözcüklerin sözlü ve yazılı dildeki görünüm sıklıklarına ilişkin söz konusu çalıřmaları (Pierce, 1960, 1961, 1962, 1963), 1950'lerin sonlarındaki Türkçe sözvarlığına ilişkin önemli ipuçları sunmaktadır. 1961'deki çalıřmasında Pierce, sözlü dile ilişkin 140,000 sözcükten oluşan derlemde elde edilen bulgulara dayanarak Türkçedeki 110 ekin (52 çekimsel, 58 türetimsel) sıklığını ortaya koymuştur. 1962'deki çalıřmasında ise yazılı dile ilişkin 2,000,000 sözcükten oluşan derlemde örneklem olarak seçilen 100,000 sözcüklük veri tabanına dayanan Türkçedeki 139 farklı ekin (36 çekimsel, 103 türetimsel) görünüm sıklıkları sunulmuştur. 1963'teki çalıřmasında ise aynı derleme dayanarak eklerle sözlükbirimler arasındaki ilişki ele alınmıştır. Tarihsel olarak önem taşısa da Pierce'in görünüm sıklığına ilişkin bulgularının dayandıđı derlemin sistematik bir biçimde oluşturulmadığı görülmektedir.²

Gerçek anlamda bir sıklık sözlüğü olarak değerlendirilmese de oluşturduđu sözcük dizini, sözcüklerin görünüm sıklıklarına dayandıđı için Tezcan'ın (1988) temel sözvarlığı çalıřmasına da değinmek gerekir. Bu çalıřma, görünüm sıklığına ilişkin iki çalıřmaya dayanmaktadır: Bunlardan biri Pierce'in 1960

¹ Thorndike'nin (1921) 10,000 sözcükten oluşan sıklık listesi, çocuk yazını (625,000 sözcük); İngilizce klasikler ve İncil (3,000,000 sözcük); ilkokul ders kitapları (300,000 sözcük); mutfak, dikiř, çiftçilik, ticaret gibi konulardaki kitaplar (50,000 sözcük), gazeteler (90,000 sözcük) ve yazışmalardan oluşan (500,000 sözcük) 41 farklı kaynađa dayanmaktadır.

² Pierce'in sözlü derleminde fabrika işçileriyle yapılan görüşmeler ve otobiyografiler bulunurken, yazılı derleminde askeri alanda kılavuzlar, ders kitapları, şiirler, dini metinlerle, gazete ve dergilerde derlenmiş makaleler yer almaktadır.

yılındaki çalışması, diğeri ise Milli Eğitim Bakanlığı tarafından 1975 yılında stenografi dersleri için hazırlanan yayımlanmamış bir çalışmadır (Tezcan, 1988). Sözlükte sözcük listesi sunulmuş, ancak sözcüklere ilişkin herhangi bir sıklık bilgisi verilmemiştir.

Göz'ün (2003) çalışması daha önceki sıklık çalışmalarının tersine, daha sistematik bir biçimde oluşturulmuş çok daha geniş bir derleme (1,006,306 sözcük) dayanan sözlük formatında hazırlanmış ilk çalışmadır. 22,693 madde başından oluşan *Yazılı Türkçenin Sözcük Sıklığı Sözlüğü* adlı bu sözlük, 1995-2000 tarihleri arasında yayımlanmış 470 metinden oluşan derlem temelinde hazırlanmıştır. Göz'ün (2003) sıklık sözlüğünün dayandığı derlemdeki gereçlerin dağılımı, Brown derlemindeki (Kučera ve Francis, 1967) dağılıma uygun olarak oluşturulmuştur. Sözlükte Pierce'in çalışmalarının tersine, eklere ilişkin sıklık bilgisi bulmamakla birlikte, sözlüksel birimler hem abecesel hem de sıklıklarına göre sıralanmıştır. Sözlükte sıklık bilgisi sadece derleme dayalı ham veri olarak sunulmuş, herhangi bir *yayılım ölçümü* (dispersion measure) ya da *düzeltilmiş sıklık* (adjusted frequency) bilgisi verilmemiştir. Bununla birlikte, sözlükteki sıklık bilgisinin geçerliliği bağımsız deney ve sormacalarla ortaya konmuştur (Tekcan ve diğ., 2002).

Sözlük formatında hazırlanmış Türkçe sözcük sıklığı çalışmalarından diğeri ikisi Memoğlu-Süleymanoğlu'nun (2006) ve Ölker'in (2011) sıklık sözlükleridir. 500,000 sözcükten oluşan derlem temelinde hazırlanan Memoğlu-Süleymanoğlu'nun (2006) çalışmasında, sözcükler sadece abecesel olarak ve sıklıklarına göre dizilmemiş, aynı zamanda tersinden abecesel olarak da dizilmiştir. Ölker (2011) ise Göz'ün (2003) derlemindeki metin dağılımına sadık kalarak, 1945-1950 yılları arasındaki döneme ait Türkçe sözlük sıklığını ortaya koymuştur. Bu sözlükte 929,015 sözcükten oluşan derlemden elde edilen 26,425 sözcüğün sıklık bilgisi yer almaktadır.

Göz (2003), Memoğlu-Süleymanoğlu (2006), Ölkel (2011) ve tanıtımını yapmakta olduğumuz Aksan ve diğ. (2017) dışında, bildiğimiz kadarıyla, Türkçe için hazırlanmış sözlük formatında, sözcüklerin görünüm sıklıklarını sunan başkaca bir çalışma bulunmamaktadır. Bununla birlikte, sözcük sıklığına ilişkin çeşitli çözümleme çalışmaları bulunmaktadır. Söz gelimi, İlgen ve Karaoğlan (2007), sözcük sıklıkları ile sözcüklerin anlam sayısı arasında korelasyon olduğunu, söz konusu korelasyonun da Türkçeye özgü olmadığını ortaya koymuşlardır. Bunlar dışında, Türkçe Ulusal Derlemi'ne (TUD) dayalı sıklık çalışmaları yapılmıştır. Bunlardan biri, kurgusal metinlerle gazete metinlerindeki kök sözcük çeşitlerinin sayısal sıralı sözcük sıklığı bakımından karşılaştırıldığı çalışmadır (Aksan ve Yaldır, 2012). Bir diğeri de ilköğretim Türkçe ders kitaplarında yer alan sözcük sıklıklarıyla TUD örneklemindeki sözcük sıklıklarının karşılaştırıldığı çalışmadır (Aksan, Mersinli ve Yaldır, 2011).

3 Sıklık Sözlüğünün Özellikleri

3.1 Derlem

Sözcük sıklığı konusunda yapılan çalışmalarının dile ilişkin gerçek anlamda örneklem sunabilmesi için, söz konusu sıklık çalışmalarının dayandığı derlemlerin ölçünlüleştirilmiş ve normalize edilmiş olması gerekmektedir. Diğer bir deyişle, derlemi oluşturan türlerin kapsamının dağılımsal olarak dengeli olması, her tür için metin seçiminin uygun ve yeterli olması, metinlerin zamana uygun olarak seçilmesi, derlemdeki sözcük sayısının kapsamlı olması, sonuç olarak derlemin temsil gücünün yüksek olması beklenmektedir. Yeşim Aksan, Mustafa Aksan, Ümit Mersinli ve Umut Ufuk Demirhan tarafından hazırlanan *A Frequency Dictionary of Turkish* adlı çalışma, (Aksan ve diğ., 2017), 23 yılı kapsayan (1990-2013) çeşitli metin kategorileri içinde yer alan 49,983,288 sözcük biçiminden oluşan TUD'a dayanmaktadır (Aksan ve diğ., 2012). TUD web arayüzü aracılığıyla kullanım ve sorgulama olanağı sunmaktadır (bkz. <http://www.tnc.org.tr/>).

Tablo 1. TUD'da metinlerin konu alanına göre dağılımı

Konu alanı	Sözcük sayısı	Sözcük oranı	Metin sayısı	Metin oranı
Kurgusal: düz yazı	9,365,775	18.74	674	13.54
Bilgilendirici: doğa ve temel bilimler	1,367,213	2.74	253	5.08
Bilgilendirici: uygulamalı bilimler	3,464,557	6.93	461	9.26
Bilgilendirici: toplum bilimleri	7,151,622	14.31	671	13.48
Bilgilendirici: dünya sorunları	9,840,241	19.69	757	15.21
Bilgilendirici: ticaret ve finans	4,513,233	9.03	429	8.62
Bilgilendirici: sanat	3,659,025	7.32	347	6.97
Bilgilendirici: inanç ve düşünce	2,200,019	4.4	297	5.97
Bilgilendirici: boş zaman	8,421,603	16.85	1,089	21.88
Toplam	49,983,288	100	4,978	100

Tasarım ilkeleri *British National Corpus* (Aston ve Burnard, 1998) örnek alınarak geliştirilmiş olan TUD'un, güncel dil kullanımına yönelik örneklem sunması ve metin türlerinin dengeli dağılımı bakımlarından temsil gücünün yüksek olduğu görülmektedir. Derlemin %98'ini oluşturan yazılı metinler, kurgusal (yazınsal metinler) ve bilgilendirici (toplumbilim, sanat, ticaret-finans, düşünce-inanç, dünya sorunları, uygulamalı bilimler, doğa-temel bilimler, hobi, yemek tarifi, vb.) metinlerden oluşmaktadır (bkz. Tablo 1). Yayın ortamı olarak kitaplar, süreli yayınlar, yayımlanmış ve yayımlanmamış çeşitli metinler ve konuşma için hazırlanan metinler (haber metinleri, senaryo) kullanılmıştır. Derlemin %2'sini oluşturan sözlü metinler ise, günlük konuşmalar, toplantı,

ders ve söyleşi gibi alanlardan toplanan sözlü verinin çeviriyazıyla geçirilmiş metinlerini içermektedir. Davies ve Preto-Bay'i (2008) izleyerek, sıklık sözlüğüne %10 oranında sözlü metin verisi eklenmiştir. Özetle, yukarıda sıklık konusundaki alanyazındaki çalışmalardan farklı olarak *A Frequency Dictionary of Turkish* adlı çalışmanın, örneklemin kapsamı ve dengesi, sözlü metin verisinin eklenmiş olması bakımlarından, temsil gücü çok daha yüksek olan bir derleme dayandığı görülmektedir.

3.2 Ham Sıklık Bilgisi ve Yayılım Değerleri

Daha önceki Türkçe sıklık çalışmalarında, sıklık bilgisinin, ham sıklık değerleri üzerinden verildiği görülmektedir. Diğer bir deyişle, şimdiye kadarki sıklık sözlüklerinde tüm dilsel değişkelerde –herhangi bir yayılımsal etki gözetmeksizin– birimlerin yinelenen görünüm sıklıkları bir bütün olarak hesaplanmıştır. Ancak derlem dilbiliminde, bu tür bir hesaplamanın önemli ölçüde yanıltıcı olabileceği belirtilmektedir (Gries, 2008). Tablo 2’de sıklık değerleri aynı, ancak yayılım değerleri farklı çiftlerde ham sıklık bilgisinin ne derece yanıltıcı olabileceği açıkça görülebilmektedir. Tablo 2’deki her bir çift aynı ham sıklık değerine sahip olmasına karşın, çiftlerden ilkinin, ikincisine göre yayılım değeri daha yüksektir, yani sözcük derlemde çok daha fazla metinde yer almaktadır. Söz gelimi, sözlükte *elektron*, *ezan* ve *içsel* sözcüklerinin tümünün ham sıklığı 1287 olarak belirtilmektedir. Terim özelliği sergileyen *elektron*, sözcüğünün diğerleriyle aynı sıklık oranına sahip olması beklendik bir durum değildir. Ancak sözlükte sözcüklerin yayılımına bakıldığında Juilland’ın yayılım ölçümüne göre (Juilland’s D), *ezan* ve *içsel* sözcüklerinin sırasıyla 0.87 ve 0.82 yayılım değerlerine sahip olduğu, oysa *elektron* sözcüğünün sadece 0.51 değerinde olduğu görülmektedir.³ Benzer olarak, sözlükte *enzim*, *kızıl*, *tarihçi*, *düşmanlık* sözcükleri aynı ham sıklık değerine (1813) sahip olmalarına karşın, yine bir terim olan *enzim* sözcüğünün yayılım değerleri 0.52 iken, *kızıl* ve *tarihçi* sözcüklerinin 0.89, *düşmanlık* sözcüğünün ise 0.88’dir.

Sözlükte sözcüklerin derlemdeki yayılımsal etkisini belirleyebilmek için Juilland’ın yayılım ölçümü kullanılmıştır. Bu ölçüm ilk olarak İspanyolca için hazırlanmış sıklık sözlüğünde (Juilland ve Chang-Rodriguez, 1964), daha sonra da Fransızca için hazırlanmış sıklık sözlüğünde (Juilland, Brodin ve Davidovitch, 1970) kullanılmıştır. Juilland’ın ölçümünün, sözlüksel yayılım için oldukça güvenilir ölçüm sağladığı ileri sürülmüş (Lyne, 1985; Leech, Raysonve ve Wilson, 2001: 18), pek çok sıklık çalışmasında da kullanılmıştır (örneğin bkz. Davies ve Gardner, 2010; Gardner ve Davies, 2013). Bununla

³ Sözlükte kullanılan Juilland’ın ölçümüne dayanan yayılım değerleri, 0 ile 1 arasındadır, sayının artması sözcüğün derlemde yayılımının arttığını göstermektedir.

birlikte, Gries (2008) alternatif bir ölçüm yöntemi (Deviation of Proportions, DP) önermiştir. *British National Corpus*'u kullanarak Juilland'ın yayılım ölçümünü değerlendiren Biber ve diğ. (2016) de Juilland'ın ölçümünün verimliliğini tartışarak Gries'in söz konusu alternatif ölçümünün katkılarını ortaya koymuştur. Ancak tüm bunlara karşın, Juilland'ın ölçümünün alanyazında geçerliliği sürmektedir.

Sözlükte sıklık sıralanışı, ham sıklık değeri ile yayılım değerinin çarpımından elde edilen skora göre gerçekleştirilmiştir, ancak sözlükte skor değerleri yer almamaktadır. Söz gelimi, Tablo 2'de görüldüğü gibi, *haz* ve *iyon* sözcüklerinin ham sıklık değerleri 1977'dir, ancak Juilland'ın ölçümüne göre yayılım değerleri farklı olduğundan (sırasıyla, 0.87 ve 0.48), *haz* sözcüğüne ilişkin skor, *iyon* sözcüğüne ilişkin skordan fazladır. Bu da *haz* sözcüğünün, *iyon* sözcüğüne göre sıralamada daha önde yer almasına neden olmaktadır.

Tablo 2. Ham sıklık, yayılım, skor ve sıralama değerlerinin karşılaştırılması*

Başsözcük	Sıklık	Yayılım	Skor	Sıralama
katılımcı	2712	0.82	2223.84	2150
çözelti	2712	0.40	1084.80	3278
haz	1977	0.87	1719.99	2552
iyon	1977	0.48	948.96	3521
düşmanlık	1813	0.88	1595.44	2652
enzim	1813	0.52	942.76	3543
olanaklı	1471	0.85	1250.35	3046
mimarlık	1471	0.63	926.73	3585
kılık	1463	0.89	1302.07	2986
hadis	1463	0.61	892.43	3644
aşağıla- ihracatçı	1424	0.90	1281.60	3000
	1424	0.62	882.88	3667
ezan	1287	0.87	1119.69	3226
elektron	1287	0.51	656.37	4277
seks	1269	0.83	1053.27	3338
ülkücü	1269	0.53	672.57	4224
gereç	1054	0.90	948.60	3537
antibiyotik	1054	0.46	484.84	4948

*Ham sıklık, yayılım ve sıralama değerleri Aksan ve diğ. (2017)'e dayanmaktadır. Skor= Sıklık × Yayılım

3.3 Sözlükte Sıklık Bilgilerinin Sunumu

Sözlükte sözcük ve eklerin sıklıklarına ilişkin bilgiler, sırasıyla sıklık dizininde, abecesel dizinde, sözcük kategorilerine göre oluşturulmuş dizinde ve eklere göre oluşturulmuş dizinde sunulmaktadır.

Sıklık dizininde 5,000 başsözcük, yukarıda değindiğimiz sıralama ölçütlerine göre, en sık sözcükten en az sık sözcüğe göre sıralanmıştır. Şekil 1’de görüldüğü gibi, her bir maddede (örneğin, *devrim*), başsözcüğün sıralama numarası (966), sözcüğün kategorisi (*n*), İngilizce karşılığı (*revolution*), derlemeden elde edilen örnek (*Devrim ne zaman sona erecek?*) ve İngilizce çevirisi (*When will the revolution end?*), ham sıklık ve yayılım değerleri (8197 | 0.86) yer almaktadır. Sözlükte özel adlar, yabancı dile ait sözcükler ve diğer ilgisiz sözcükler sıklık sayımına dâhil edilmemiştir. Sözlükte eşsesli sözcükler sıklık sayımı bakımından ayrı olarak ele alınmamış olmasına karşın, sözcük kategorisi, İngilizce karşılıkları ve örnekler bakımından ayrı ayrı sunulmuştur. Söz gelimi, 2242. sırada yer alan *kalkış* başsözcüğüne ilişkin tek bir sıklık bilgisi verilmesine karşın (2301|0.91), sözcük kategorisi ve İngilizce karşılıkları ayrı ayrı sunulmuştur: (*v*) *to try to* (*n*) *departure, take off*. Sıklık dizininde ayrıca kutu içinde yirmi konu listesinde (hayvanlar, yiyecekler, durum ekleri vb.) sözcük ve eklere ilişkin sıklıklar sunulmuştur (bkz. Şekil 2).

Şekil 1. Sıklık dizinine ilişkin örnek üç maddebaşı

<p>966 devrim <i>n</i> revolution</p> <ul style="list-style-type: none"> • Devrim ne zaman sona erecek? — When will the revolution end? 8197 0.86
<p>967 takdir <i>n</i> appreciation</p> <ul style="list-style-type: none"> • Çalışmalarının takdir görmesi genç bilgini heyecanlandırıyordu. — Appreciation of his work excited the young scientist. 7566 0.93
<p>968 gelenek <i>n</i> tradition, custom</p> <ul style="list-style-type: none"> • Gelenek yeni bir aydın hareketi yaratma mücadelemizde daha etkili olmalıdır. — The tradition must be more effective in our struggle to create a new intellectual movement. 8045 0.87

Ana sıklık dizini dışında, sözlükte 5,000 sözcük abecesel olarak da sunulmuştur. Abecesel dizinde, ana sıklık dizininden farklı olarak örnek tümce, ham sıklık değeri ve yayılım değeri yer almamaktadır. Sözlükte üçüncü bir dizin olarak sözcük kategorilerine göre sözcüklerin dizildiği sözcük türü dizini bulunmaktadır. Bu dizinde her bir maddebaşında, ilgili kategori içinde sıralama değeri, tüm sözcükler içinde sıralama değeri, başsözcük ve sözcüğün İngilizce

karřılıđı sunulmaktadır. Sözlükte son olarak, adcıl ve eylemcil çekim eklerinin diziliřine iliřkin sıklık dizini yer almaktadır. Diđer Türkçe sıklık sözlüklerinde ve dil öğretim gereçlerinde bulunmayan söz konusu sıklık dizininde, en sık 50 biçimleniř sunulmaktadır. Her bir maddebařında, sıralama deđer, ek(ler)e iliřkin bilgi⁴, ham sıklık deđer ve örnek sözcük biçimi yer almaktadır (bkz. Őekil 3).

Őekil 2. Örnek konu listesi

15 Case suffixes		
Accusative	4503459	defteri the notebook
Locative	3253653	defterde in the notebook
Dative	2896450	deftere to the notebook
Genitive	2560613	defterin (rengi) (color) of the notebook
Ablative	1358407	defterden from the notebook
Instrumental	765498	defterle with the notebook

Őekil 3. Ekler dizinine iliřkin örnek maddebařları

21	futr+3s	125281	olacak
22	pcck	111406	olacak
23	pasv+nzma+p3s	107434	fırçalanması
24	imprf+1s	102849	istiyorum
25	imprf+vi+past+3s	102740	acıyordu
26	pasv+perf+3s	96454	kazanılmıř
27	cont+cop+3s	89068	etmektedir
28	perf+cop+3s	85934	konuşmuřtur
29	perf+vi+past+3s	85516	önemsemiřti
30	pasv+aor+3s	85403	edilir

4 Sonuç

A Frequency Dictionary of Turkish adlı çalıřma, ařađı yukarı 50 milyon sözcükten oluřan, 23 yıllık bir dönemi kapsayan, günümüz Türkçesinin çok sayıda farklı konu alanını ve metin türünü içeren, geniř kapsamlı, dengeli ve dili temsil yeterliliđine sahip bir derleme (TUD) dayandıđı için güvenilir bir sıklık verisi sunmaktadır. Sözlükte bařsözcüklerin sıklık sıralaması, ham sıklık deđerlerine göre deđil, yayılım deđerlerinin hesaplanmasıyla elde edilen skorlara göre oluřturulduđundan gerçek anlamda bir sıklık bilgisi sunmaktadır.

⁴ Sözlükte eklerin gösterimine iliřkin açıklamalar, sözlüğün kısaltmalar bölümünde bulunmaktadır.

Ayrıca her başşözcük için GDEX ölçütleri (bkz. Kilgarriff ve diğ., 2008) temel alınarak derlemeden otantik örnek tümcelerinin çıkartılmış olması, Türkçe sözlükçülük tarihinde bir ilktir. Yine eklere ilişkin sıklık değerlerinin sunulmuş olması, sözcük kategorilerine ve kimi konulara göre sıklık bilgilerinin sunulmuş olması sözlüğün önemli özelliklerindedir.

Sözlük temel olarak ikinci ya da yabancı dil olarak Türkçe öğrenenler için sözcük sıklık bilgisi veriyor olsa da aslında sözlükbilim, sözcük anlambilim gibi alanlarda çalışanlar için, ruhdilbilim, sinirdilbilim gibi alanlarda yapılacak deneysel çalışmalar için önemli bir kaynak niteliğindedir. Ayrıca sözlükle birlikte CD içinde sekmeye-ayrılmış formatta sıklık bilgilerine ilişkin bir veri tabanı sunulmaktadır. Sayısal ortamda sunulan bu veri tabanı derlem dilbilimi ya da berimsel dilbilim alanında çalışanlar için güvenilir bir veri tabanı niteliğindedir.

Kaynaklar

- Aksan, Y., Aksan, M., Koltuksuz, A., Sezer, T., Mersinli, Ü., Demirhan, U.U., ... Yıldız, İ. (2012). Construction of the Turkish National Corpus (TNC). *Proceedings of the eight international conference on language resources and evaluation (LREC 2012)* (ss. 3223–3237) içinde. İstanbul, Türkiye. <http://www.lrec-conf.org/proceedings/lrec2012/papers.html>
- Aksan, Y. ve Yaldir, Y. (2012). A corpus-based word frequency list of Turkish: Evidence from the Turkish National Corpus. E. Kincses-Nagy ve M. Biacsi (yay.), *The Szged Conference: Proceedings of the 15th international conference on Turkish linguistics* (ss. 47–58) içinde. Szged, Hungary: Studia Uralo-Altaica.
- Aksan, Y., Aksan, M., Mersinli, Ü., ve Demirhan, U.U. (2017). *A frequency dictionary of Turkish: Core vocabulary for learners*. Londra/New York: Routledge.
- Aksan, Y., Mersinli, Ü. ve Yaldir, Y. (2011). İlköğretim Türkçe ders kitapları derlemi ve Türkçe ulusal dil derlemi örneklemindeki sözcük sıklıkları. D. Günay ve diğ. (yay.), *Türkçe öğretimi üzerine çalışmalar* (ss. 397–408) içinde. İzmir: Dokuz Eylül Üniversitesi Yayınları.
- Aksoy, Ö.A. (1936). *Bir dili öğrenmek için en lüzumlu kelimeler ve bu kelimelerin belirtme usulü*. Gaziantep: Gaziantep Halkevi Dil, Edebiyat, Tarih Şubesi Yayınları.
- Aston, G., ve Burnard, L. (1998). *The BNC handbook: Exploring the British National corpus with SARA*. Edinburg: Edinburg University Press.
- Biber, D., Reppen, R., Schnur, E., ve Ghanem R. (2016). On the (non)utility of Juilland's D to measure lexical dispersion in large corpora. *International Journal of Corpus Linguistics*, 21(4), 439–464.
- Davies, M. ve Gardner, D. (2010). *A frequency dictionary of contemporary American English: Word sketches, collocates, and thematic lists*. London: Routledge.
- Davies, M. ve Preto-Bay, A. (2008). *A frequency dictionary of contemporary Portuguese*. London: Routledge.

- Fillmore, C.J. (2008). 'Corpus linguistics' or 'computer-aided armchair linguistics'. T. Fontenelle (yay.), *Practical lexicography: A reader* (ss.105–122) içinde. Oxford / New York: Oxford University Press.
- Gardner, D. ve Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, 34(5), 1–24.
- Göz, İ. (2003). *Yazılı Türkçenin kelime sıklığı sözlüğü*. Ankara: TDK Yayınları.
- Gries, S. Th. (2008). Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics*, 13(4), 403–437.
- İlgen B. ve Karaođlan B. (2007). Investigation of Zipf's 'law-of meaning' on Turkish corpora, *22nd International Symposium on Computer and Information Sciences, Middle East Technical University, Cultural and Convention Center, Ankara, Turkey*, IEEE Computer Society 7-9 Kasım.
- Juilland, A. ve Chang-Rodriguez, E. (1964). *Frequency dictionary of Spanish words*. Lahey: Mouton de Gruyter.
- Juilland, A.G., Brodin, D.R., ve Davidovitch, C. (1970). *Frequency dictionary of French words*. Lahey: Mouton de Gruyter
- Kilgarriff, A., Miloř Husák, M., McAdam, M., Rundell, M., ve Rychly, P. (2008) GDEX: Automatically finding good dictionary examples in a corpus. Bernal, Elisenda, ve Janet DeCesaris (yay.) *Proceedings of the 13th EURALEX international congress* (ss. 425-432) içinde. Barcelona: Institut Universitari de Linguística Aplicada, Universitat Pompeu Fabra.
- Kučera, H. ve Francis, W.N. (1967). *Computational analysis of presentday American English*. Providence, RI: Brown University Press.
- Leech, G., Rayson, P., ve Wilson, A. (2001). *Word frequencies in written and spoken English: Based on the British National Corpus*. Londra: Longman.
- Lyne, A. (1985). *The vocabulary of French business correspondence*. Cenevre: Slatkine.
- Memođlu-Süleymanođlu, H. (2006). *Türkçenin ters sıklık sözlüğü*. Ankara: Kurmay Yayınları.
- Ölker, G. (2011). *Yazılı Türkçenin kelime sıklığı sözlüğü (1945-1950 arası)*. Konya: Kömen Yayınları.
- Pierce, J.E. (1960). *A frequency count of Turkish words*. Ankara: MEB Yayım Müdürlüğü Basılı Eđitim Malzemeleri Hazırlama Merkezi.
- Pierce, J.E. (1961). A frequency count of Turkish affixes. *Anthropological Linguistics*, 3(9), 31–40.
- Pierce, J.E. (1962). Frequencies of occurrence of affixes in written Turkish. *Anthropological Linguistics*, 4(6), 30–41.
- Pierce, J.E. (1963). A statistcal study of grammar and lexicon in Turkish and Sahaptin (Klikitat). *International Journal of American Linguistics*, 29(2), 96–106.
- Tekcan, A. İ., Göz, İ., Yalçın, S., Akırmak, Ü., Serbest, S., Fırat, C., ve Rodoplu, S. (2002). Yazılı Türkçenin Kelime Sıklığı Sözlüğü'nün geçerliđi üzerine bir çalıřma. *Türk Psikoloji Dergisi*, 17(50), 27–42.
- Tezcan, N (1988). *Elementarwortschatz Türkisch – Deutsch*. Turkologie und Türkeikunde Dizisi, Bd.1 (yay K. Kreiser), Wiesbaden: Harrassowitz Verlag.
- Thorndike, E.L. (1921). *The teacher's work book*. New York: Teachers College, Columbia University Press.