



Öğrenen Derlemlerinin Dil Öğretiminde Kullanımı Üzerine Kuramsal Tartışmalar*

*Anna GOLYNSKAIA***
*Fatma BÖLÜKBAŞ KAYA****

Öz

Derlem, belli bir dili veya dil değişkesini temsil edebilmek amacıyla, yazılı ve/veya sözlü dil kullanımlarını dengeli ve katmanlı örnekleme yoluyla derleyip temel dilbilimsel çözümlemelerle birlikte elektronik ortamlarda sunan kaynaklardır. Bu çalışmanın amacı, öğrenen derlemlerinin dil öğretiminde kullanımı üzerine farklı yaklaşımları incelemek ve mevcut öğrenen derlemlerini çeşitli ölçütler bakımından değerlendirerek araştırmacılara kuramsal bir çerçeve sunmaktır. Ayrıca dil öğretiminde dünyada uzun yıllardır yararlanılan öğrenen derlemlerinden Türkçe öğretiminde de yararlanılması için bir farkındalık oluşturmaktır. Bu amaç doğrultusunda derlem, derlem dil bilimi, derlem türleri ve öğrenen derlemleriyle ilgili alan yazın taraması yapılmış, öğrenen derlemlerinin dil öğretiminin hangi aşamasında nasıl kullanıldığıyla ilgili saptamalar ortaya konulmuştur. Öğrenen derlemlerinin genellikle aradil çözümlemelerinde, derslerde kullanılacak ders kitabı, dil bilgisi kitabı, sözlük gibi kaynakların hazırlanmasında, dil öğrenenlerin hedef dilin kullanım örneklerini inceleme çalışmalarında, öğretimin müfredat oluşturma ve ölçme-değerlendirme aşamalarında, özel amaçlı dil öğretimi ve akademik amaçlı dil öğretimi çalışmalarında kullanıldığı sonucuna ulaşılmıştır.

Anahtar Kelimeler: Derlem, derlem dilbilimi, öğrenen derlemleri, dil öğretimi

Theoretical Discussions on the Use of Learner Corpora in Language Teaching

Abstract

Corpora are electronically stored and processed collections of written and/or spoken language data compiled by means of balanced and layered sampling and representing a specific language or language variety. The purpose of this study is to examine different approaches to the use of learner corpora in language teaching and to present a theoretical framework to researchers by evaluating existing learner corpora in terms of various criteria. Learner corpora have long been used in the field of language instruction; therefore, the present study also aims at raising awareness of the use of learner corpora in teaching Turkish language as well. To attain this objective, literature review of corpus, corpus linguistics, corpus types, and learner corpora was made and the ways that learner corpora were used in each stage of language teaching were described. It was determined that learner corpora were usually used in interlanguage analysis, during the preparation of reference materials such as textbooks, grammar books and dictionaries, in studies examining the learners' use of the target language, in the process of curriculum design and language testing and assessment, as well as LSP and LAP.

Keywords: Corpus, corpus linguistics, learner corpora, language teaching

* Bu makale, İstanbul Üniversitesi-Cerrahpaşa Lisansüstü Eğitim Enstitüsünde Doç. Dr. Fatma Bölükbaş Kaya danışmanlığında yürütülen "İleri Düzey Türkçe Öğrencilerinin Dilbilimsel Yeterlilikleri: Derlem Tabanlı Bir Araştırma" adlı doktora tezinden üretilmiştir.

** Doktora Öğrencisi, İstanbul Üniversitesi-Cerrahpaşa, Lisansüstü Eğitim Enstitüsü, Türkçe Eğitimi Doktora Programı, İstanbul, anna201188@hotmail.com, ORCID: 0000-0001-6921-8110

*** Doç. Dr., İstanbul Üniversitesi-Cerrahpaşa, Hasan Ali Yücel Eğitim Fakültesi, Türkçe ve Sosyal Bilimler Eğitimi Bölümü, İstanbul, fbolukbas@istanbul.edu.tr, ORCID: 0000-0002-1592-4885

Giriş

Derlem dilbilim, dil değışkeleri ile dil kullanımının deneysel incelemelerini kolaylařtıran ve diđer yöntemlere göre genellenebilirlik ve geçerliliđi daha yüksek olan bulguları ortaya koyan bir araştırma yöntemidir. Derlem dilbilim kapsamında yapılan arařtırmaların birtakım ortak hedefleri ile özellikleri bulunmaktadır. Biber, Conrad ve Reppen (1998:4), bu özellikleri řöyle sıralamaktadır: (I) Deneyseldirler, dođal metinlerdeki gerçek kullanım örüntülerini incelerler. (II) Dođal dil kullanımlarını örnekleyen metinlerin büyük bir koleksiyonu olarak tanımlanabilen derlemin çözümlmesine dayanırlar. (III) Çözümlleme sürecinde geniş ölçüde bilgisayar teknolojilerinden yararlanırlar. (IV) Hem nicel hem de nitel çözümlleme tekniklerini kullanırlar.

Derlem dilbilim teriminin ilk kez Aarts ve Meijs (1984) tarafından yayınlanan “Corpus Linguistics: Recent Developments in the Use of Computer Corpora in English Language Research” adlı çalışmada geçtiđi kabul edilmektedir. Ancak Aarts bu terimi 1980 yılında da kullandığını, ayrıca bu adlandırmanın çok başarılı olmadığını dile getirmektedir. Aarts, derlem dilbilimden “başlıca araştırma aracı ile veri kaynağının adını taşıyan garip bir disiplin” olarak bahsetmektedir (Taylor, 2008: 179).

Alanyazında derlem dilbilim bir araç, yöntem, metodoloji, metodolojik yaklaşım, disiplin, teori, teorik yaklaşım, (teorik veya metodolojik) paradigma veya bunların karışımı olarak değerdendirilmektedir. Leech (1992: 106)’e göre “bilgisayar temelli derlem dilbilim, sadece dili incelemek için ortaya çıkan yeni bir metodoloji deđil, aynı zamanda yeni bir araştırma girişimi ve konuya ışık tutan yeni bir felsefi yaklaşımdır”.

Stubbs (1993: 23-24) derlemi sadece dilbilimsel bir çözümlleme aracı deđil, dilbilim kuramında önemli bir kavram olarak tanımlar. Stubbs (2001: 243), derlem dilbilimle fen bilimleri arasında bir paralel çizerek řunları dile getirir: “Jeologlar, doğrudan gözlemlenemeyen süreçlerle ilgilenirler, çünkü inceledikleri malzemeler uzun zaman dilimleri içinde meydana gelir. Derlem dilbilimciler de tıpkı jeologlar gibi doğrudan gözlemlenemeyen süreçlerle ilgilenirler, çünkü derlemler de birçok farklı konuşmacı ve yazarın uzun zaman dilimleri içinde gerçekleşen dil kullanımı örneklerinden oluşur.

Teubert (2005:2) derlem dilbilimi, “dil çalışmaları için teorik bir yaklaşım” olarak değerdendirir. Teubert (2005: 4)’e göre, derlem dilbilim kendi başına bir yöntem deđildir; derlem verilerinin işlenmesi ve çözümlenmesinde birçok farklı yöntem kullanılmaktadır. Derlem dilbilim daha ziyade, belli ilkelere göre söylemden alınarak derlenen gerçek dilsel veriler ile çalışılması gerektiğini iddia eden bir yaklaşımdır.

Thompson ve Hunston (2006: 8) derlem dilbilimin “dile yönelik herhangi bir teorik yaklaşıma göre ayarlanabilen bir metodoloji” olduđu fikrini savunur. McEnery ve Wilson (2001: 2) derlem dilbilimin, “dilbilimin belli bir alanı olmaktan ziyade neredeyse bütün dilbilimsel alanlarda kullanılabilecek bir metodoloji” olduđu üzerinde durur. McEnery, Xiao ve Tono (2006: 7-8) ise derlem dilbilimi “Derlemin dil çalışmalarına ve dil öğretimine nasıl uygulanabileceğini gösteren yöntemler ve ilkeler sistemi.” olarak tanımlayarak derlem dilbilimin kesinlikle teorik bir statüye sahip olduđunun altını çizer.

Derlem dilbilimi eleştiren ve “standart bilimsel yöntemlerinden yararlanarak dil hakkında daha fazla şey öğrenebileceğimizi” iddia eden Chomsky (2004: 97) de derlem dilbilimi bir yöntem olarak değerdendirir. Teubert ve Krishnamurthy (2007) ise derlem dilbilimi, Saussure ve Chomsky’nin “dil (langue) dilbilimi” ne karşı koyarak “söz (parole) dilbilimi” olarak betimler.

Aksan (2011: 339)’a göre, derlem dilbilim, bir dili konuşanların ve yazarların kullandıkları dilin sözcük varlığını ve dilbilgisi olanaklarını nasıl kullandıkları sorusuna cevap vermeyi amaçlamaktadır. Bu sorunun cevabı, dil kullanımını örnekleyen yazılı ve sözlü metinlerin toplamı olan derlemin oluşturulmasını gerektirmektedir.

Bu çalışmanın amacı, öğrenen derlemlerinin dil öğretiminde kullanımı üzerine farklı yaklaşımları incelemek ve mevcut öğrenen derlemlerini çeşitli ölçütler bakımından değerdendirerek arařtırmacılara kuramsal bir çerçeve sunmaktır. Ayrıca dil öğretiminde dünyada uzun yıllardır yararlanılan öğrenen derlemlerinden Türkçe öğretiminde de yararlanılması için bir farkındalık oluşturmaktır. Kuramsal olarak planlanan bu çalışmada doğrudan raporlařtırmaya gidilmiştir.

Derlem ve Derlem Türleri

Derlem, belli bir dili veya dil değişkesini temsil edebilmek amacıyla, belli bir zaman aralığında yazılı ve/veya sözlü dil kullanımlarını coğrafi bölge, yazar/konuşan özellikleri (cinsiyet, yaş, eğitim vb.) ve iletişim ortamlarının alan ve türlerine göre dengeli ve katmanlı örnekleme yoluyla derleyip ayrıntılı veri bilgisi ve temel dilbilimsel çözümlerle birlikte elektronik ortamlarda sunan kaynaklardır (Biber, Conrad ve Reppen (1998), Gries (2009), McEnery ve Hardie (2012); Akt. Ruhi, Aksan ve Aksan, 2014: 264).

Oluşturulma amaçlarına ve oluşturulma yöntemlerine göre farklı derlem türleri bulunmaktadır. Günümüzde mevcut olan derlem türleri aşağıdaki şekilde özetlenebilir:

1. Örnekleme Derlemi: Örnekleme derlemi, belirli bir dilin belli bir zaman dilimindeki “anlık görünümü”dür (Örn. Brown Derlemi). Örnekleme derleminin başlıca amacı, bir dilin ve dil değişkesinin standart özelliklerini genel kullanımındaki benzer oranda yansıtmaktır. Bir dilin tüm özelliklerini sergileyen örnekleme derlemi, referans derlemi olarak adlandırılır. Referans derlemi en az 100 milyon kelime içerir ve bir toplulukta kullanılan tüm metin türlerini (yazılı, sözlü, kamuya ait, özel, bilgilendirici, kurgu vs.) kapsar (Tognini Bonelli, 2010: 20).

2. Kıyaslanabilir Derlemler: İki veya daha fazla derlem, aynı tasarım ölçütleri üzerine kurulduğunda veya benzer boyutta olduğunda, kıyaslanabilir nitelikte olur. Bunlara kıyaslanabilir derlemler denir. 1961’de yayımlanan belgelerden oluşan Brown Derlemi ile 1978’de tamamlanan Lancaster-Oslo-Bergen (LOB) Derlemi kıyaslanabilir derlemlere örnek olarak verilebilir.

3. Genel ve Özel Derlemler: Genel derlemler tek bir metin türü, konu alanı veya kesite ait olmayan metinlerden oluşur (Örn. British National Corpus). Özel derlemler ise bir dili bir bütün olarak temsil etmeyi amaçlamayan, dilin sadece belli bir kesimini temsil eden derlemlerdir (Örn. MICASE – Michigan Sözlü Akademik İngilizce Derlemi) (Lee, 2010: 114).

4. Yazılı ve Sözlü Derlemler: Yazılı derlemler belirli bir dilin yazılı metinlerinin bir araya getirilmesiyle oluşur (Örn. BEC – İş İngilizcesi Derlemi). Yazılı derlem oluşturmanın, sözlü derlem oluşturmaya göre daha kolay olduğuna dair yaygın bir inanç vardır. Bu kısmen doğru olsa da yazılı derlem oluşturmanın metinleri seçme, ihtiyaç duyulan metinlere erişim sağlama ve metinleri bilgisayarca okunur hâle getirme gibi birçok zorluğu bulunmaktadır (Nelson, 2010: 53).

Sözlü derlemler, bir dilin türlü kullanım alanlarında ses kayıtlarıyla elde edilmiş konuşmaların yazıya aktarılması ile oluşturulur (Örn. CANCODE – Cambridge ve Nottingham İngilizce Söylem Derlemi) (Akın, 2015: 6). Sözlü derlemler, yazılı derlemlere kıyasla daha küçük olur ve bu nedenle bazı sözcük birim ve kelime öbeklerinin yazılı derlemlerdekiyle aynı oranda tekrarlanmasını sağlayamaz. Bunun dışında sözlü söylemin çözümlenmesi tonlama, jest ve söylem yapısı gibi metin ötesi özelliklere dikkat etmeyi gerektirir, dolayısıyla da sırf yazılı derlemin analizinde başvurulan sıklığa dayalı tekniklerin aracılığıyla sözlü derlemlerin analizi yapılamaz (Adolph ve Knight, 2010: 53).

5. Artzamanlı / Tarihî ve Monitör Derlemler: Artzamanlı derlem, farklı dönemlere ait metinler içeren bir derlem olup dildeki gelişimi veya değişimi incelemek için kullanılır (Örn. Helsinki Derlemi) (Lee, 2010: 113). Monitör derlemde ise sürekli olarak yeni içerik eklenerek ve bu içerik durağan bir nitelik taşıyan referans derlemi ile karşılaştırılarak dil değişiminin takip edilmesi amaçlanır (Walter, 2010: 430).

6. İki Dilli ve Çok Dilli Derlemler: İki dilli ve çok dilli derlemler başlangıçta bilgisayar dilbiliminin asıl hedefi olan makine çevirisinin ihtiyaçlarını karşılamak üzere geliştirilmiştir. Biri diğerinin çevirisi olan iki tek dilli derlemden oluşan derleme paralel derlem denir. En çok bilinen paralel derlem, Kanada’nın resmi dilleri olan İngilizce ve Fransızca olarak yayınlanan Kanada Parlamentosu tartışmalarından oluşan Kanada Hansard Derlemi’dir (Tognini Bonelli, 2010: 23).

Aston (1999) iki dilli ve çok dilli derlemleri aşağıdaki başlıklar altında incelemektedir (Akt. Lee, 2010: 119):

Tablo 1.
İki Dilli ve Çok Dilli Derlemler

Derlem Türü	A Dili	B Dili
Kıyaslanabilir (çok dilli) derlemler	Özel derlem	Aynı tasarıma sahip özel derlem
Tek yönlü paralel derlemler	Özel derlem A1	A dilindeki metinlerin çevirisi B1
İki yönlü paralel derlemler	Özel derlem A2 B1 çevirileri	A1 ile aynı tasarımı olan özel derlem B2 A1 çevirileri

7. Ayrıştırılmış / Ağaç Yapılı Derlemler: Ayrıştırılmış, biçimbirimsel ve sözdizimsel etiketlerle işaretlenen derlemlerdir. Böyle derlemler için bütün bir derlemin tamamen ve tutarlı bir şekilde etiketlendirildiğinin kontrol edilmesi zor olduğundan genellikle küçük boyutlu olmaları tercih edilir (Agan ve Diri, 2016: 1). Ağaç yapılı ilk derlemlerden biri, LPC (Lancaster Ayrıştırılmış Derlemi)'dir.

8. Çoklu Ortam Derlemleri: Çoklu ortam derlemleri, orijinal ses veya video kayıtları ile senkronize edilmiş transkripsiyonlara sahip olan derlemlerdir (Örn. SBCSAE – Santa Barbara Amerikan İngilizcesi Derlemi). Bu, derlem araştırmacılarının ve sıradan kullanıcıların prozodi, jestler, bakışlar, aralık vb. gibi söylemin işitsel ve görsel öğelerine erişmelerine izin verir (Lee, 2010: 114-115).

9. Pedagojik Derlemler: Pedagojik derlemler, öğrencilerin çeşitli kanallar aracılığıyla maruz kaldığı veya maruz kalması muhtemel olduğu sözlü ile yazılı dilin büyük ve temsil gücü yüksek olan örneğidir (Örn. ELISA, SACODEYL ve BACKBONE derlemleri). Bu derlemler, öğrencilerin sınıf içinde veya kendi kendilerine çalışmalarını esnasında karşılaşabilecekleri veya teknoloji destekli iletişim sırasında rastlayabilecekleri yazılı ve sözlü metinleri içerir (Chambers, 2015: 447-448).

10. Gelişimsel Dil Derlemleri: Gelişimsel dil derlemleri, çeşitli gelişim aşamalarında bulunan ve bir dilin anadili konuşucusu olan çocukların ürettiği metinlerden oluşan derlemlere verilen isimdir. CHILDES veritabanı ve Wales Politeknik Derlemi (POW), gelişimsel dil derlemlerine örnek olarak verilebilir (Lee, 2010: 116).

Derlem türlerinden bir diğeri de bu makalenin konusu olan öğrenen derlemleridir. Bu makalede dil öğreniminde öğrenen derlemlerinden nasıl yararlanılabileceği irdelendiği için, “öğrenen derlemleri” konusu ayrı bir başlık altında incelenecektir.

Öğrenen Derlemleri

Teknolojideki ilerlemelere paralel olarak gelişen bilgisayar destekli öğrenen derlemleri, “ikinci dil edinimi ve yabancı dil öğretimine yönelik belirli bir amaç için belirli tasarım ölçütlerine göre bir araya getirilen hedef dile ait özgün metin verilerinin elektronik koleksiyonları” olarak tanımlanmaktadır (Granger, 2002: 7).

Günümüze kadar dil öğretiminin en çok ilgi gören boyutları; hedef dili tanımlama, eğitim-öğretim (görev, izlenec, müfredat, öğretim yöntemleri vb.) ortamları ve öğrenen özellikleri (güdü, öğrenme stilleri, tutum vb.) ile ilgili konular olmuştur. Öğrenen dili ise gerek araştırmacılar gerekse eğitimciler tarafından hep göz ardı edilmiştir (Mark, 1998: 84). Öğrenen dilinin, derlem dilbiliminkine benzer yöntemlerle incelenebileceği fikri ilk öğrenen derlemi oluşturulmadan çok daha önce ortaya çıkmıştır. 1980’li yıllarda ilk olarak Yunoven (1989), Cornu&Delahaye (1987) ve Huebner (1983, 1985) gibi araştırmalarda dil öğrenenlerin yaptığı yanlışların çözümlemesi ve sınıflandırması üzerinde durulmaya başlanmıştır. Sözü geçen bu araştırmalar küçük, kâğıt tabanlı öğrenen derlemleri temel alınarak gerçekleştirilmiştir (Akt. McEnery, Brezina, Gablasova ve Banerjee, 2019: 77).

Bilgisayar öncesi dönemde öğrenenlerden gelen veriye az başvurulmasının başlıca sebebi, bu tür verilerin toplanması ve çözümlenmesinin çok zaman ve emek istemesiydi. Günümüz teknolojileri ise çok miktarda veriyi toplayıp bilgisayarda depolamaya ve bu verileri otomatik veya yarı otomatik şekilde çözümlemeye imkân tanımaktadır. Bilgisayar destekli ilk öğrenen derlemi olan International Corpus of Learner English (Uluslararası İngilizce Öğrenen Derlemi) 2002 yılında CD-ROM olarak

Öğrenen Derlemlerinin Dil Öğretiminde Kullanımı Üzerine Kuramsal Tartışmalar

yayınlanmıştır. Farklı dilleri (Bulgarca, Çince, Çekçe, Felemenkçe, Fince, Fransızca, Almanca, İtalyanca, Japonca, Norveççe, Lehçe, Rusça, İspanyolca, İsveççe, Tsvanaca, Türkçe) konuşan, İngilizceye orta ve ileri düzeyde hâkim olan öğrencilerin ortaya koyduğu tartışma türündeki yazılardan oluşan derlemin birinci sürümü, yaklaşık 2 milyon sözcük içermektedir. 2009'da çıkan derlemin ikinci sürümüne dâhili bağımlı dizin oluşturma programı eklenmiştir.

Tono (2003: 800)'ya göre öğrenen derlemi hazırlamada kullanılacak tasarım ölçütleri dille ilgili olanlar, görevle ilgili olanlar ve öğrenenle ilgili olanlar olmak üzere üç başlık altında toplanabilmektedir:

Tablo 2.

Öğrenen Derlemi Hazırlamada Kullanılacak Tasarım Ölçütleri

Dille İlgili Olanlar	Görevle İlgili Olanlar	Öğrenenle İlgili Olanlar
B biçim (yazılı/sözlü)	Veri toplama (kesitsel/boylamsal)	İçsel-Bilişsel (yaş/bilişsel tarz)
Tür (mektup/günlük/kurgu/deneme)	Çıktı (hazırlıksız/hazırlıklı)	İçsel-Duygusal (güdü/tutum)
Anlatım biçimi (öyküleme/tartışma)	Referans kullanımı (sözlük/kaynak metin)	Birinci dil artalanı
Konu (genel/serbest zaman vb.)	Süre kısıtlılığı (sabit/serbest/ev ödevi)	İkinci dil ortamı (ikinci dil/yabancı dil)/ İkinci dil yeterliği (standart test sonucu)

Öğrenen derlemleri, sıklık taraması ve dökümü gibi temel çözümlerlerin yanı sıra, gelişmiş veri çözümlerini de olanaklı kılan çeşitli yazılımlar ile çözümlenerek farklı anadili artalanlarından gelen öğrencilerin aradıl özelliklerini ortaya koymaktadır (Can, 2009: 19). Öğrenenlerden elde edilen veriler, metin dışı bilgiler (dil/görev/öğrenen değişkenleri), transkripsiyon düzeyi (yazım) ve etiketleme düzeyi (cümle sınırlarını belirleme, sözcükbirimleştirme, sözcük türü etiketleme, başsözcükleştirme, ögelere ayırma, semantik etiketleme, söylemsel etiketleme, hata etiketleme, prozodi etiketlemesi, gönderimsel etiketleme) göz önünde bulundurularak işlenmektedir (Tono, 2003: 801).

Öğrenen derlemleri, tek dilli/iki dilli, genel/mesleki, eşzamanlı/artzamanlı, yazılı/sözlü olmak üzere dört başlık altında incelenebilmektedir (Granger, 2002: 9). Bunun dışında öğrenen derlemleri küresel /yerel, ticari amaçlı/akademik amaçlı, büyük ölçekli /küçük ölçekli, İngilizce olan / İngilizce olmayan, ileri zamanda pedagojik amaçlı kullanılacak olanlar/hemen öğrencilerin kullanımına sunulanlar, aynı ana dili konuşan öğrencilerle ana dilleri farklı olan öğrencilerden elde edilen verilerden oluşan derlemler gibi gruplara ayrılmaktadır (Granger, 2008: 260-262; Gilquin, 2015: 12-15).

Öğrenen Derlemi Araştırmaları

Öğrenen derlemi araştırmaları, ikinci dil ya da yabancı dil kullanımını ortaya koyan derlemlerin üzerine yapılan araştırmalardır (McEnery vd., 2019: 76). 1980'li yılların sonunda ortaya çıkan öğrenen derlemi araştırmaları, derlem dilbilim ve ikinci/yabancı dil araştırmaları olmak üzere birbirinden bağımsız olarak gelişen iki alanın arasında önemli bir köprü vazifesini üstlenmiştir. Öğrenen derlemi araştırmaları ne yeni bir dilbilim alanı ne de yeni bir dil kuramıdır; kullandığı veri türü onu dile olan yaklaşımları değiştirecek kadar etkili bir yöntem hâline getirmiştir (Granger, 2002: 4). Meunier (2016: 476) öğrenen derlemi araştırmalarının temel amaçlarını şu şekilde özetlemiştir: a. Kesitsel ve boylamsal yaklaşımlar aracılığıyla öğrenenlerin dil kullanımını ve dilsel gelişimlerini tanımlamak. b. İkinci dil edinim süreçlerine ışık tutmak. c. Pedagojik materyalleri ve veriye dayalı öğrenme etkinliklerini geliştirmek.

Genel derlem dilbilimde olduğu gibi öğrenen derlemi araştırmalarında da derleme dayalı (corpus-informed), derlem temelli (corpus-based) ve derlem çıkışlı (corpus-driven) olmak üzere üç

yaklaşımdan bahsedilebilir. Derleme dayalı araştırmalarda derlem, araştırmacının sezgilerini doğrulamak veya belli bir dilbilimsel özelliğin var olduğunu kanıtlamak için kullanılan başlıca başvuru kaynağı olarak karşımıza çıkar. Bu yaklaşımı benimseyen araştırmacılar derlem verilerine doğrudan erişim sağlamadan derlemde elden edilen verilerle meta düzeyde çalışır. Derlem temelli araştırmalarda derlemlerin ortaya çıkmasından önceki varsayım ve kuramların doğrulanması veya çürütülmesi hedeflenir. Bunun yanı sıra “derlem temelli” ifadesinin derlem üzerine kurulan tüm çalışmaları betimlemek için geniş anlamda da kullanıldığı görülür. Tümevarımsal bir niteliğe sahip olan derlem çıkışlı araştırmalar ise mevcut varsayımları sınamaktan ziyade araştırmacının yeni sorular sorması ve elindeki verileri değerlendirip bir sonuca varması için imkân tanır (Callies, 2015: 36).

Callies (2015: 36)’e göre, günümüzde öğrenen derlemi araştırmalarının çoğu derlem temelli, nicel, kesitsel ve karşılaştırmalıdır. Öğrenen derlemleri, karşılaştırmalı aradil analizi ve bilgisayar destekli hata çözümlemesi gibi metodolojik yaklaşımlarda kullanılmaktadır.

Karşılaştırmalı aradil analizi iki tür karşılaştırma içerir. Ana dili konuşucuları ile ana dili konuşucusu olmayanların dil kullanımlarının karşılaştırması, öğrenenlerin yazılı ve sözlü anlatım örneklerindeki hataları tespit etmeye ve bazı kelime, ifade ve yapıların ana dili konuşucularındakiyle kıyasla daha az veya daha çok temsil edilmelerinin nedenlerini araştırmaya olanak sağlar. Bazı dilbilimciler aradilin başlı başına incelenmesi gereken bir konu olup ana dili ile karşılaştırılmaması gerektiği fikrini savunarak bu tür bir karşılaştırmaya karşı çıkar (Granger, 2002: 12). Bunun dışında, karşılaştırmalı aradil analizi farklı ana dili artalanlarından gelen öğrenenlerin aradillerinin karşılaştırılarak incelenmesini mümkün kılar. Böyle bir karşılaştırma hangi aradil özelliklerinin farklı öğrenen gruplarında bulunarak gelişimsel bir nitelik taşıdığını, hangilerinin ise aynı uyruğa sahip belli bir öğrenen grubuna ait olup o grubun ana diline bağlı geliştiğini ortaya koyar (Granger, 2002: 13).

Bilgisayar destekli hata çözümlemesi, belli bir hata etiketleme sistemine göre etiketlenmiş öğrenen derlemini gerektirir. Hata etiketleme sistemleri dilbilimsel kategorilerinin üzerine kurulu olup biçim, dilbilgisi, sözcük düzlemi, kesit, tarz, sözcük, sözcüksel dilbilgisi ve noktalama kategorilerine göre sınıflandırılmış hataların taksonomisini içerir. Ancak geleneksel hata çözümlemesi yönteminde olduğu gibi bilgisayar destekli hata çözümlemesinde de bazı sorunlar ve sınırlamalar mevcuttur. Birinci olarak, bu yöntem öğrenenlerin dil kullanımını bir bütün olarak değerlendirmeyi, başka bir deyişle hataların üzerine odaklanarak hatalı olmayan kullanımları göz ardı eder. Oysaki öğrenen dilinin özelliklerine ışık tutması açısından hatalı olmayan kullanımlar hatalar kadar bilgilendirici olabilir. İkinci olarak, hataların ölçünlü dil kullanımından bir sapma olduğunu varsayarsak ölçünlü kullanımın nasıl tanımlanabildiği (Örneğin hangi dil değişkesinin temel alınacağı, sadece dilbilgisel doğruluğun mu yoksa edimsel uygunluğunun mu göz önünde bulundurulacağı, yazılı dilin normlarının mı yoksa sözlü dilin normlarının mı standart olarak değerlendirileceği, vb.) sorunu ortaya çıkar. Üçüncü sınırlılık ise örtülü hataların saptanmasının zorluğu konusundadır. Açık hataların ölçünlü dilden bir sapma olduğu çıplak gözle görülürken örtülü hatalar ilk bakışta doğru gelen ama öğrenenin yüklediği anlamdan farklı bir anlam ifade eden bir nitelikte olur. Sonuç olarak sistematik yapılan ve öğrenenin belli bir alandaki yetersizliğinden kaynaklanan hatalarla dikkatsizlikten kaynaklanan yazım yanlışların ayırt edilmesi güçleşir. Zira öğrenen hakkında yeterince bilgi sahibi olmadan bilgi eksikliğinden kaynaklanan hatalar ile dikkatsizlikten kaynaklanan yanlışlar arasında ayırım yapmak son derece zordur (Callies, 2015: 40-41).

Öğrenen Derlemlerinin Dil Öğretiminde Kullanımı

Öğrenen derlemlerinin dil eğitimine katkısı çoğunlukla dolaylı olarak gerçekleşir (Stubbs, 2004: 11). Bu tür derlemler daha ileriki zamanlarda kullanılmak üzere akademisyenler ya da yayıncılar tarafından oluşturulur. Ayrıca derlem verilerini sağlayan kitleyle bu verilerden yararlanan kitle farklıdır (Granger, 2015: 486).

Granger (2002) öğrenen derlemlerinin dil eğitimine dönük potansiyel kullanım alanlarını ders izlencesi oluşturma, ders materyali geliştirme ve ders içi etkinlikleri hazırlama olarak sıralar. Meunier (2016) öğrenen derlemlerini gramer kitapları, öğrenen sözlükleri ve diğer yazılı öğretim materyalleri ve ölçme - değerlendirme için birer girdi olarak değerlendirir. Nesselhauf (2006) ise öğrenen derlemlerinin yabancı öğrencilerce üretilen anlatım örneklerinde hedef dile ait olmayan unsurların

tespit edilmesi ve ikinci dil edinim sürecinin incelenmesi amacıyla kullanılabilceğini, ayrıca veri güdümlü öğrenme çerçevesinde hazırlanan etkinliklerin temelini oluşturabileceğini ifade eder.

Öğrenen derlemleri üzerine yapılan çalışmalardan yola çıkarak öğrenen derlemlerinin Türkçe öğretiminde kullanım amaçları şu şekilde sıralanabilir:

1. Öğrenen derlemleri aradil çözümlemelerinde kullanılır. Öğrenen derlemleri farklı yaş, cinsiyet, eğitim düzeyi, meslek ve anadillere sahip olan öğrencilerin aradil özelliklerine ışık tutar. Selinker (1972) tarafından öne sürülen aradil (interlanguage) terimi, amaç dil ile ana dili arasındaki etkileşimin zihinde gerçekleşme sürecinin bir ürünü olarak tanımlanır. Aradilin en önemli özelliklerinden biri fosilleşmedir. Fosilleşme, dil öğrenenlerin amaç dili erken yaşta ya da uzun süredir öğreniyor olmalarına rağmen aradillerinde belli başlı dilbilimsel birimleri, kural ve alt sistemleri muhafaza etmeleri durumudur. İkinci dil edinimi ve yabancı dil eğitimi çalışmaları açısından bakıldığında, öğrencilerin aradil verilerinin toplanması ve çözümlenmesi iki açıdan önemlidir. Birincisi aradil araştırmaları ikinci dil edinim sürecinin, onu etkileyen etmenlerin ve aradilin zihinsel simgelenmesinin daha iyi anlaşılabilmesine katkıda bulunur; ikinci olarak da yabancı dil sürecinde yaşanan zorlukların giderilmesi için öğretim ve öğrenim araçları tasarlayan araştırmacılara kaynak sağlar (Can, 2009: 18).

Cobb (2003: 394), aradil incelemelerinin, erken yaştaki ikinci dil edinim sürecini değerlendirmede yararlı olduğunu, ancak orta ve ileri düzey öğrenenlerin devam eden dil gelişimlerini anlamaya daha az katkıda bulunduğunu ifade eder. Cobb (2003)'a göre orta ve ileri düzey aradilin yeterince araştırılmamış olmasının iki nedeni vardır: Bunlardan biri veri eksikliğidir. Orta ve ileri düzey dil kullanıcılarının verileri, bu kişilerin ne zaman dil öğrenmeye başladıkları ve öğrenme tarzları açısından çeşitlilik gösterdiği için üzerinde incelemeye yapmaya çok olanak tanımazlar. Ayrıca, ileri düzey öğrenenlerin dil gelişiminin büyük bir kısmı ikinci dilde okuma ve yazmayı öğrenmeyi kapsar, yazılı etkileşim ise veri toplanması açısından öğrenciler arasındaki sözlü etkileşim kadar uygun olmaz. İkinci neden, geç dil edinimi hakkında yeterince bilgi sahibi olmamaktan kaynaklanır. İleri düzeyde dil öğrenmenin birtakım ortak noktaları mevcut olsa bile bunlar dilbilgisi, söylem ve pragmatik gibi biçimbilimle sözdiziminin ötesinde kalan alanların içinde yer alır ki bu alanlar ana dil konuşucuları açısından bile yeterince irdelenmemiştir.

Derlem çalışmaları, aradil, özellikle de ileri düzey öğrenenlerin aradili ile ilgili varsayımların sınanmasında büyük bir rol oynar. Cobb (2003), bu tür varsayımların derlem dilbilimi yöntemleriyle ele alındığı üç araştırmadan bahseder. Ringbom (1998) öğrenen dilinin belirsiz ve basmakalıp olduğu savını kanıtlayarak bu durumun öğrenenlerin kullanım sıklığı yüksek olan öğeleri gerektiğinden fazla kullanmalarından ve genel kavramlar arasındaki farkları ayırt etmekte zorlanmalarından kaynaklandığını ileri sürmüştür. Cock, Granger, Leech ve McEney (1998), öğrenen dilinin hata içermemesine rağmen kulağa yabancı geldiği görüşünü ele alarak bunu öğrenen dilinde sınırlı sayıda kalıp ifadenin bulunmasına bağlamıştır. Petch-Tyson (1998), ileri düzey öğrenenlerin yazmalarının bağlam, durum ve yazı türüne bakılmaksızın yazıya dökülen konuşma dili olduğu varsayımının doğruluğunu araştırmıştır. Cobb (2003) yukarıda adı geçen araştırmaların ve kendi araştırmalarının bulgularına dayanarak, derlem dilbilimi aracılığıyla ispatlanmış sorunların dil eğitimi sürecinde nasıl çözülebileceği konusunda önerilerde bulunmuştur.

Öğrenen derlemlerinin eğitimcilerle sağladığı en büyük katkı, dil öğrenenlerin kullandığı belli başlı yapı ve kelimelerin sıklığı hakkında bilgi vermesidir. Kullanım sıklığı, farklı dil kullanım düzeylerinde (sözcüksel, dilbilgisel ve pragmatik) ve öğelerin tek başına ya da birlikte kullanılması (eşdizimlilik) açısından analiz edilebilir. Konuşmacıların dilsel öğeleri kullanma sıklığı aradillerinin durumuna dair fikir verir ve bu öğelerin kendi dillerinde kullanılmalarına sebep olan etmenlerin araştırılmasına olanak tanır (Gablasova, Brezina ve McEney, 2017).

Sıklık listelerinin oluşturulması, bazı dilbilgisel yapıların ya da sözcüklerin öğrencilerce yeterince kullanılıp kullanılmadığının saptanmasına hizmet eder. Öğrencilerin hangi dil bilgisel yapıları ve sözcükleri kullanmaktan kaçındıklarının saptanması, öğretmenlerin derslerde hangi konular üzerinde daha fazla durmaları gerektiği konusunda onlara yol gösterici olabilir. Bazı dilsel öğelerin gereğinden fazla kullanıldığının ortaya konması ise öğretmenlere öğrencilerinin kendilerini hedef

dilde ifade etmekte güçlük çektiklerini gösterir, dolayısıyla da ifade biçimlerinin nasıl zenginleştirilebileceği üzerinde öğretmenlere düşünme fırsatı yaratır.

2. Öğrenen derlemleri derslerde kullanılacak kaynaklarının hazırlanmasında yol gösterici olur. Dil öğretiminde kullanılan kaynaklardan biri “öğrenen sözlükleri”dir. Öğrenen sözlükleri, genellikle bir yabancı dili orta veya ileri seviyede öğrenenleri hedef alan sözlüklerdir. Ana dil konuşucuları için hazırlanan sözlüklerden farklı olarak öğrenen sözlükleri sözcüklerin dilbilgisel, eşdizimsel, sözdizimsel, pragmatik vb. özellikleriyle ilgili ayrıntılı bilgiler verir. Öğrenen sözlüklerinde, ana dil konuşucuları sözlüklerine göre daha az sözcük bulunur ve sözcük tanımları olabildiğince basit tutulmaya özen gösterilir. Bunların dışında öğrenen sözlüklerinde etimolojik ve tarihî bilgilere yer verilmez (Heuberger, 2016: 25).

Günümüzde sözlükbilimciler sözlük maddelerini oluştururken neolojizmlere, kelime sıklıklarına, tipik kelime kullanımlarına, bağlamsallaştırılmış tanımlara ve özgün örneklerle kolayca erişmek için ana dili derlemleri ile derlem dilbilim araç ve yöntemlerini etkin olarak kullanırlar. Öğrenen derlemlerinden elde edilen verilerin sözlükbilime olan katkısından ilk bahsedenler ise Maingay ve Rundell (1987) olmuştur. “Öğrenenlerin hatalarını öngörme: Sözlük hazırlayanlar için çıkarımlar” başlıklı makalelerinde öğrenen derlemlerinin öğrencilerin hata yapmaya meyilli oldukları sözcüklerin maddesine eklenecek uyarı notlarının oluşturulmasında kullanılabileceği dile getirilmiştir (Akt. Granger, 2015: 489). Öğrenen derlemleri bir yandan dil öğrenenlerin yaptığı dilbilgisel ve sözcüksel yanlışlarla yazım ve üslup hatalarına dair bilgi verirken öbür yandan hedef dildeki sözcük ve ifadeleri doğru oranda kullanıp kullanmadıklarını gösterir. Bu bilgiler, doğrudan ya da dolaylı olarak öğrenen sözlüklerine dâhil edilir:

tasteful



Şekil 1. Longman Dictionary of Contemporary English Sözlüğündeki Uyarı Notu Örneği

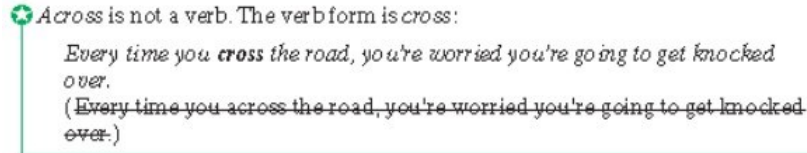
Yukarıdaki resimde, öğrencilerin “tasteful” ve “tasty” sözcüklerini karıştırmamaları için sözlük yazarlarınca uyarı notu düşüldüğü görülmektedir. Öğrenen derlemleri, öğrencilerin en çok hata yaptıkları sözcükleri sözlük yazarlarına göstermesi bakımından önemli bir işleve sahiptir.

Öğrenen sözlüklerindeki tanımlar genellikle tanımlayıcı sözlükçe kullanılarak yazılır. Tanımlayıcı sözlükçe; öğrencilerin bilmeleri gereken, kullanım sıklığı yüksek olan sözcüklerin listesidir. Örneğin, LDOCE'nin tanımlayıcı sözlükçesi yaklaşık 2000 sözcükten oluşur. Tanımlayıcı sözlükçe, bilinmeyi bilinmeyenle tanımlama riskini azaltarak kullanıcıların tanımları anlama şansını artırır (Adamska-Sałaciak, 2012: 327; Ostermann, 2015: 36). Öğrenen derlemlerinin öğrenen sözlükbilimine sağladığı bir başka katkı, tanımlarda kullanılan sözcüklerin öğrenenlerce anlaşılmasını kolaylaştırmasıdır. Örneğin, Longman tanımlayıcı sözlükçesindeki kelimelerin Longman Öğrenen Derlemi'ndeki kullanım sıklığı ve doğru kullanım oranlarıyla karşılaştırılması gerek bütün öğrenenler gerekse belli ana dili artalanlarından gelen öğrenenler için zorluk oluşturabilecek kelimelerin tespit edilmesine imkân tanımıştır (Granger, 2015: 489).

Öğrenen derlemleri, sözlükbilimcilerin belirli öğrenci kategorilerini ele almalarını da sağlar. Örneğin öğrenen derlemi öğrencilerin yeterlilik düzeyine göre filtrelenebilir, bu da bilgilerin sözlüğün hedef kitlesi için uygun olacak şekilde tasarlanmasına katkıda bulunur. Öğrenen derlemleri, öğrencilerin kelimeleri nasıl kullandıklarının yanı sıra hangi kelimeleri kullandıklarına da ışık tutar. Bu da öğrencilerin dil üretimi esnasında kullanmaya eğilim gösterdikleri kelimelerin dil bilgisi ve eşdizimlilik gibi kullanım özelliklerinin sözlük maddelerine dâhil edilmesine imkân verir (Walter, 2010: 440).

Anadili ile öğrenen dilinin karşılaştırılması da öğrenen sözlüklerinin tasarımını etkileyebilir. Örneğin, anadili konuşucuları tarafından yazılan metinlerde bazı sözcükler daha az tekrarlanırken, yabancı dil öğrenenlerce yazılan metinlerde sözcük tekrarı ile daha fazla karşılaşılabilir; çünkü dil öğrenenler o dili anadili olarak konuşanlar kadar eş anlamlı sözcük bilmemektedirler. Bu durum, sözlüklere kavram dizini türündeki bilgilerin eklenmesi gerekip gerekmediği konusunda sözlük yazarlarına bilgi verebilir.

Ders kitapları ile dil bilgisi kitapları da dil öğretiminde kullanılan öğretim materyallerinin başında gelmektedir. Granger (2015: 491), günümüzde dilbilgisi kitaplarının çoğunun hâlâ sezgiye dayalı olduğunu, oradaki kuralların yapay ve genellikle kullanım olasılığı düşük olan örneklerle açıklandığını ifade eder. Bu açıdan ana dili derlemleri gerçek dil kullanım durum ve örnekleri için kaynaklık edebilir. Hata etiketlemesi yapılan öğrenen derlemi ise öğrencilerin yaptığı tipik hatalar hakkında bilgi içeren uyarı notlarının oluşturulması için kullanılabilir. Genel öğrenci kitlesini hedefleyen dil bilgisi kitaplarında sadece en yaygın olan hatalara yer verilmesi gerekirken belirli bir öğrenci grubu (Örneğin İngilizce veya Rusça konuşucuları vb.) için tasarlanmış kitaplarda o gruba özgü hatalar gösterilebilir (Granger, 2015: 492). Örneğin Ronald Carter ve Michael McCarthy (2006) tarafından hazırlanan Cambridge Grammar of English adlı İngilizce dil bilgisi kitabı, Cambridge Uluslararası Derlemi'nden elde edilen kullanım örneklerinin yanı sıra İngilizce öğrenenlerin yaptığı başlıca hatalarla da ilgili bilgi içerir:



Şekil 2. Cambridge Grammar of English Kitabındaki Uyarı Notu Örneği

Ana dili derlemi ile öğrenen derlemi verileri üzerine kurulan bir başka dilbilgisi serisi Blass, Iannuzzi, Savage ve Reppen'in (2012) Grammar and Beyond adlı kitabıdır. Öğrenci derlemi, tüm ünitelerde yer alan "Yaygın Hatalardan Kaçınma" ve "Metin Düzeltme" bölümlerini oluşturmak için kullanılmıştır. "Yaygın Hatalardan Kaçınma" bölümleri, öğrencilerin yapmaya eğilimli olduğu hata örneklerini ve uyarıları içerir. "Metin Düzeltme" bölümleri ise tespit edilmesi ve düzeltilmesi gereken hataları içeren bir metinden oluşur:

5 Avoid Common Mistakes

- 1 Use the past perfect or past perfect progressive to give background information for a past tense event.
had
I ~~have~~ never seen my sister in real life, so I was nervous the first time we met.
had been dreaming
I ~~have dreamed~~ about meeting her, and I finally did.
- 2 Use the past perfect or past perfect progressive to give a reason for a past event.
had been crying
Her eyes were red and puffy because she ~~cried~~.
- 3 Use the past perfect (not the past perfect progressive) for a completed earlier event.
arranged
They ~~had been arranging~~ a time to meet, but both of them forgot about it.
- 4 Use the past perfect (not present perfect) to describe a completed event that happened before a past event.
had
I ~~have~~ visited her in Maine twice before she came to visit me.

Editing Task

Find and correct seven more mistakes in the paragraphs about sibling differences.

had
I ~~have~~ never really thought about sibling differences until my own children were born. When we had our first child, my husband and I have lived in Chicago for just a few months. We have not made many friends yet, so we spent all our time with our child. Baby Gilbert was happy to be the center of attention. He depended on us
5 for everything.

Şekil 3. Grammar and Beyond Kitabındaki Uyarı Notu Örneği

Geleneksel başvuru kaynakları olan sözlük ve dil bilgisi kitaplarının yanı sıra, ikinci dil edinilenlerle yabancı dil öğrenenlerin ihtiyaçlarını karşılamak üzere derleme dayalı yeni nesil kaynakların da tasarlandığı görülmektedir. Bilgisayar destekli dil öğrenimi (CALL) geniş anlamda teknolojinin dil öğrenimi ve öğretimi için kullanılması olarak tanımlanabilir. Öğrenen derlemlerinin bilgisayar destekli dil öğreniminde doğrudan kullanımına nadir olarak rastlanır. Bunun sebebi, öğretmenlerin öğrencilerini hatalı girdiye maruz bırakmayı istememeleridir. Bununla birlikte, öğrenen derlemlerinin bilgisayar destekli dil öğreniminde kullanımı üzerine yapılan bazı çalışmalar bulunmaktadır. Bunlardan biri, Hegelheimer ve Fisher'in (2006) Internet Writing Resource for the Innovative Teaching of English (iWRITE) sistemidir. Sistemin geliştirilmesinin birinci aşamasında Iowa Eyalet Üniversitesi'nde öğrenim gören yabancı öğrencilerden dört farklı konuda açıklayıcı anlatım türünde birer kompozisyon yazmaları istenmiş ve bunun sonucunda 45 metinden oluşan 12.839 kelimelik bir derlem oluşturulmuştur. Hata etiketlemesi yapıldıktan sonra öğrencilerin metinleri HTML olarak görüntüleyebilmeleri ve onlarla etkileşime girebilmeleri için birtakım uygulamalar yapılmıştır. Sonunda öğrencilere çeşitli hataların nasıl düzeltilebileceğini gösteren, onların ülkeye ve TOEFL puanına göre sıralanmış yazıları görüntülemelerini sağlayan, edisyon ve hata düzeltme becerilerini geliştirecek boşluk doldurma alıştırmalarını Word formatında indirmelerine ve eğitimcilerin hata düzeltme sürecini izlemelerine olanak tanıyan, ayrıca öğrencilere kelimelerin nasıl kullanıldığını görmeleri için eşdizimlilik programından yararlanma olanağı veren iWRITE sistemi geliştirilmiştir.

Öğrenen derlemlerinin bilgisayar destekli dil öğrenimindeki dolaylı kullanımını Grammar Talk (Chuang ve Nesi, 2006), ESL Writing Tutor (Cowan, Choo ve Lee, 2014) ve Check My Words ve Mark My Words (Milton, 2006) gibi kaynaklar örneklendirmektedir. Grammar Talk İngilizce öğrenen Çinli öğrencilerin akademik yazmalarında en sık yaptığı hatalar üzerine odaklanmaktadır. Program farkındalık artırma etkinlikleri ile üretim odaklı etkinlikleri içerir. ESL Writing Tutor da benzer şekilde Koreli öğrencilerin İngilizcede yaptıkları dil bilgisi hatalarını tespit etme ve düzeltme becerilerini geliştirebilen bir eğitim yazılımıdır. Check My Words programı, başlangıç düzeyinde İngilizce öğrenen öğrencilerin yazı yazarken çevrimiçi dilbilgisi kılavuzuna erişmelerini sağlar. Mark my Words ise öğretmenlerin öğrenci metinlerine dönüt verirken Check My Words araç çubuğundaki kaynaklara bağlantılar içeren yorumlar eklemek için kullanabildiği bir başka araç çubuğudur. Bahsedilen bütün bu elektronik kaynaklar öğrenen derlemlerinden elde edilen verilere dayanmaktadır.

Akıllı bilgisayar destekli dil öğrenim (ICALL) uygulamaları, gerçek zamanlı otomatik hata çözümü ve öğrencilere anında geri bildirim sağlama olanağı sunar. Öğrenen derleminin kullanıldığı ilk ICALL uygulaması, Akıllı Web Tabanlı Etkileşimli Dil Öğrenimi (Intelligent Web-based Interactive Language Learning - IWiLL) adlı uygulamadır. IWiLL başlangıçta öğrencilerin yazılarını toplamak için geliştirilen bir öğrenme platformu olarak kullanılmıştır. Yazılar bir araya getirilip hata etiketlemesi yapıldıktan sonra İngilizce fiil ve isimlerin eşdizimsel açıdan yanlış kullanım örneklerini tespit eden birtakım CALL araçları geliştirilmiştir (Meunier, 2016: 479-480). Bu alanda oluşturulmuş başka bir çalışma da on yılı aşkın bir zamandır kullanılan E-Tutor Almanca öğrenme platformudur. E-Tutor öğrenci girdisinin dilbilimsel analizini gerçekleştirerek ve geri bildirim mesajlarını öğrencinin uzmanlığına uygun şekilde ayarlayarak hatalara dair kişiselleştirilmiş geri bildirim sağlamaktadır (Heift, 2000).

3. Öğrenen derlemleri dil öğrenenlere hedef dilin kullanım örneklerini sunar. Öğrenen derlemlerinin bir kısmı dil öğrenenlerin doğrudan kullanmaları amacıyla hazırlanır ki bu derlem türü "pedagojik derlem" olarak adlandırılır. Öğrenen derlemlerinin dolaysız bir şekilde kullanılması durumunda, derlemler öğretmenler tarafından hazırlanır ve hemen akabinde öğrencilerin kullanımına sunulur. Öğrenciler ise hem derlem verilerini sağlayan hem de bu verilerden faydalanan taraf olur (Granger, 2015: 486).

Braun (2005) tarafından hazırlanan English Language Interview Corpus as a Second-language Application (ELISA) adlı derlem, pedagojik derlem türüne örnek olarak verilebilir. Bu derlem İngilizcenin farklı varyantlarını konuşan ve farklı yaşam alanlarından gelen anadili İngilizce olan kişilerle yapılan röportajları içerir. ELISA yaşamın mesleki, sosyal ve kültürel boyutlarına ışık tutan çeşitli konuları kapsar. Başka bir deyişle ELISA'nın mesleki ve kişisel alanlarla ilgili farklı dil

kullanımlarını yansıttığı söylenebilir. Ayrıca söz konusu derlem sözlü anlatım örneklerine dayandığı için öğrencilerin konuşma dili konusundaki farkındalığını artırmak için de kullanılabilir.

Ana dili konuşucularının yanı sıra yabancı dil öğrenenlerden elde edilen veriler de pedagojik amaçlarla kullanılabilir. Eğiticilerin kendi öğrencilerinin dil kullanım örneklerinden oluşturduğu yerel öğrenen derlemleri derlem dilbilimi araştırmalarıyla sınıf içindeki dil öğretiminin arasında köprü görevi görebilir (Mukherjee ve Rohrbach, 2006). Eğiticilerin kendi öğrencilerinin çıktıkları üzerine odaklanmaları, daha fazla öğretmenin derlem temelli etkinliklere karşı ilgisini arttıracak; öğrencilerin öğrenen verileri ile aşına olmaları ise birçok öğrenciyi dil kullanımları üzerine düşünmeye ve böylece dil farkındalıklarını artırmaya yönlendirecektir (Mukherjee ve Rohrbach, 2006: 228). Yerel öğrenen derlemlerinin kullanımı, tipik hataların nesnel olarak değerlendirilmesine olanak sağlar. (Mukherjee ve Rohrbach, 2006: 225). Bunun dışında bu tür derlemler gerek sınıf düzeyinde gerekse bireysel düzeyde öğrenenlerin dilsel gelişimlerini kayıt altına alır ve öğrenenlerin yaptıkları hataları görerek doğru dil kullanımlarını öğrenmelerini sağlar (Mukherjee ve Rohrbach, 2006: 225, 227).

Öğrencileri hata içeren verilere maruz bırakmanın sorun teşkil edeceğini öne süren Granger (2002: 26), öğrenen derlemlerinin şu iki bağlamda kullanılması gerektiğini belirtir: a. Yapı odaklı öğretim (form-focused instruction) bağlamında, ana dili konuşucularıyla ana dili konuşucusu olmayanların ürettiği bağımlı dizinlerin karşılaştırılması yoluyla öğrencilere, bu iki kitlenin dil kullanımları arasındaki farkın gösterilmesi yararlı olabilir. b. Öğrenen derlemleri veri güdümlü öğrenimden (data-driven learning) ziyade öğrenim güdümlü veriler (learning-driven data) bağlamında değerlendirilebilir. Bu durumda öğrencilerden kendi ürettikleri metinleri çözümlenmeleri istenir, böylelikle öğrenciler kendi öğrenim çıktıkları üzerinde çalışmış olur.

4. Öğrenen derlemleri müfredat oluşturma sürecinde yol gösterici olur. Öğrenen derlemlerinin müfredattaki konuların seçimi ve sıralanışında önemli bir yeri vardır. Öğrenen derlemlerinden elde edilen veriler hangi sözcüklerin ve dilbilgisel yapıların öğrenciler için sorun oluşturduğuna, dolayısıyla da öğretmenlerin hangi yapılara daha fazla zaman ayırması gerektiğine ışık tutar. Derlemlerin sağladığı sıklık bilgileri ise öğretilecek öğelerin önem derecesine göre, yani çok kullanılanlardan az kullanılanlara doğru sıralanmasına imkân tanır (Granger, 2002: 21-24).

Öğrenen derlemi çalışmalarının müfredat oluşturmaya katkıda bulunabilecek başka bir çıktısı, hatalı kullanımlardan ziyade doğru kullanımlara odaklanarak dilbilgisel yapıların başlangıç düzeyinden ileri düzeye kadar nasıl edinildiğinin takip edilmesi ve bu yapıların edinim aşamasına göre sıraya konulması ile ilgilidir. Örneğin, Cambridge Öğrenen Derlemi ile Cambridge İngilizce Profili Derleminden yararlanılan English Profile programı, İngilizce öğrenenlerin her düzeyde bildikleri ve kullandıkları tipik dilsel öğeleri tanımlayarak İngilizce öğretmen ve öğrencilerinin Avrupa Dilleri Ortak Başvuru Metninde tanımlanan dil düzeylerinin İngilizce ile nasıl ilişkili olduğunu anlamalarına yardımcı olmayı amaçlar (Harrison, 2015: 1). Programın bir alt dalı olan İngilizce Dilbilgisi Profili, öğrencilerin dil seviyeleri yükseldikçe, dilbilgisel biçim ve anlamın yanı sıra pragmatik uygunluk konusunda nasıl yeterlilik kazandıklarını görmemizi sağlar (bkz. <https://www.englishprofile.org/english-grammar-profile>).

Öğrenen derlemleri vasıtasıyla kelime dağarcığı ile sözdizimsel özelliklerin kullanımları arasında yapılan karşılaştırma, öğrencilerin hangi sözcük veya yapıları yetersiz kullandıklarını, dolayısıyla da hangi konuların üzerine odaklanılması gerektiğini ortaya çıkarmaktadır. Hangi sözcüklerle hangi dilbilgisel yapıların ne zaman öğretilmesi gerektiği sorusu bu alanda çalışanların en çok zorlandıkları konulardan biridir. O'Donnel (2013: 577) yeterlilik seviyelerinin gerçekte mevcut olmadığını, her dil öğrenenin yabancı dili farklı şekilde edindiğini, ancak dilbilimsel öğelerin belli bir sıraya göre edinildiğini, dolayısıyla da bu öğelerin edinim zorluklarına göre sıralanabileceğini ileri sürmektedir. Bir başka deyişle O'Donnell (2013: 575), öğrenen derlemlerinin dilbilgisel öğelerin hangi yeterlilik düzeyinde daha çok kullanıldığını ortaya koymak için değil, dilbilgisel öğelerin hangi düzeyden itibaren kullanılmaya başlandığını saptamak için kullanılması gerektiğini ifade etmektedir.

5. Öğrenen derlemleri ölçme ve değerlendirme sürecine rehberlik eder. Dil yeterliliğinin ölçülmesi ve değerlendirilmesi, disiplinlerarası bir alan olup dilbilim, uygulamalı dilbilim, dil edinimi, dil öğretimi ile ölçme ve değerlendirmenin sağladığı kuram ve tanımlar üzerine kurulmuştur. Dil yeterliliğini ölçme ve değerlendirme, dili kullanma kalitesinin ölçülüp değerlendirilmesi için uygun

olan dilsel araçların araştırılması, kuramsallaştırılması ve geliştirilmesini amaçlar. Dil yeterliliğini ölçme ve değerlendirme birbiriyle bağlantılı olan iki alt alandan oluşur: Bunlardan biri “neyin” değerlendirilmesi üzerine odaklanırken diğeri değerlendirmenin “nasıl” yapılacağını konu alır (Shohamy, 2008: xiii).

1960’lardan bu yana dil yeterliliğini ölçme ve değerlendirmedeki gelişmelerin incelenmesi, bu alana ait kuram ve uygulamaların her zaman dil yeterliliği tanımlarıyla yakından ilişkili olduğunu ortaya koymaktadır. Ayrık nitelikli sınavların (discrete-point testing) kabul gördüğü dönemde dilin sözel ve yapısal öğelerden oluştuğu düşünülmuş, dolayısıyla o dönemin dil testleri birbirinden izole olan dilsel öğelerin ölçülüp değerlendirilmesini hedeflemiştir. İçerik temelli öğretim döneminde dil sınavları dilin bütünlük ve söylemsel boyutlarını öne çıkarmıştır. İletişimsel yaklaşım çağında dil yeterliliğini ölçme ve değerlendirmede özgün sözlü ve yazılı metinler kullanılarak gerçek etkileşim biçimlerinden yola çıkılmıştır. Edimbilimin yaygınlık kazandığı dönemde dil sınavlarına katılanlardan “gerçek hayat” bağlamındaki görevleri yerine getirmeleri istenmiştir. Dilin tek bir uygulama kullanılarak değerlendirilemeyen karmaşık bir olgu olduğu anlaşıldıktan sonra ise alternatif ölçme ve değerlendirme yöntemleri oluşturulmuştur (Shohamy, 2008, xiii-xiv).

Öğrenen derlemleri dil yeterliliğini ölçme ve değerlendirme alanına da birtakım yeni olanaklar sunmuştur. Öğrenen derlemleri demografik değişkenlerin (cinsiyet, ana dili, dil öğrenme deneyimi vb.), ölçme usulü (kâğıt üzerinde/bilgisayar tabanlı) ve öğrenme ortamının öğrencilerin ürünlerini ne derecede etkilediğini ortaya koyarak dil sınavlarının tasarımına katkıda bulunmaktadır. Ayrıca belli yeterlilik düzeyindeki öğrencilerin neler yapabildiklerini gösteren öğrenen derlemleri, değerlendirme ölçeklerinin şekillendirilmesi ve otomatik değerlendirme sistemlerinin geliştirilmesi aşamalarında yol gösterici olmaktadır (Barker, Salamoura ve Saville, 2015: 521-524).

Callies, Diez-Bedmar ve Zaytseva (2014), öğrenen derlemlerinin dil ölçme ve değerlendirmedeki kullanımını (I) derlem verilerinin nasıl kullanıldığına, (II) dil ölçme ve değerlendirmeye olan katkısına ve (III) araştırmacının verilerin elde edilmesi, çözümlenmesi ve yorumlanmasında aktif rol alıp almadığına göre üç başlık altında incelemektedir. Derleme dayalı yaklaşımda öğrenen derlemleri testlerin içeriğini şekillendirmek ve değerlendirme yapanların kararlarını doğrulamak amacıyla kullanılmaktadır. Bir başka deyişle, derlemler dil öğrenenlerin neler yapabildiklerini ortaya koyar, bu da belli bir seviyede nelerin ölçüleceğini ve bunların nasıl değerlendirileceğini etkiler (Barker, 2010: 639).

Derlem temelli yaklaşımda öğrenen derlemleri sık sık ana dili derlemleriyle karşılaştırılarak araştırmacının varsayımını doğrulamak veya çürütmek için aradığı ampirik kanıtların kaynağı olarak değerlendirilir. Derlem güdümlü yaklaşım ise araştırmacının herhangi bir varsayımda bulunmasını gerektirmez. Araştırmacı, derlem verilerini istatistiksel analize tabi tuttukten sonra ortaya çıkan verilere dayanarak fikir yürütür.

6. Öğrenen derlemleri akademik amaçlı ve özel amaçlı dil öğretiminde yol gösterici olur. Öğrencilerin özel amaçlı dil öğretimi veya alan dersleri çerçevesinde ortaya koydukları metinleri içeren özel amaçlı dil öğrenen derlemleri, gelecek vaat eden ancak henüz yaygınlık kazanmamış bir öğrenen derlemi türüdür. Özel amaçlı dil öğrenen derlemleri, belirli bir alana ait makale, vaka takdimi, proje raporları gibi metin örneklerini içerir. Özel amaçlı dil öğrenen derlem araştırmalarının amacı, yabancı dil öğrenenlerin bu alanlarda yaşadıkları zorluklarını ortaya çıkarmak suretiyle kişiye özel öğretim yöntem ve araçlarını tasarlamaktır (Granger ve Paquot, 2013).

Özel amaçlı dil öğretimiyle ilgili başka bir derlem türü, tercüman adaylarının yaptığı çevirileri içeren çeviri öğrenen derlemidir. Student Translation Archive (Bowker, 2002) ve MeLLANGE Learner Translator Corpus (Castagnoli vd., 2011) gibi çeviri öğrenen derlemleri ana diline yapılan çevirilerden oluşurken, PELCRA çeviri öğrenen derlemi (Uzar ve Waliński, 2001) hedef dile çevrilen metinleri içermektedir (Akt. Granger ve Paquot, 2013). Öğrenen derleminin özel amaçlı dil öğretimi kapsamında kullanıldığı çalışmalardan biri Lee ve Swales’e (2006) aittir. Bu çalışmada Akademik Amaçlı İngilizce kursunda öğrenim gören doktora öğrencilerinden iki derlem oluşturmaları istenmiştir. Bunlardan biri öğrencilerin kendi yazdıklarından (dönem sonu projeleri, tez taslakları, düzenlenmemiş makale taslaklarından), diğeri ise uzmanlık alanlarıyla ilgili olan profesyonel makalelerden oluşmaktadır. Lee ve Swales’e (2006) göre bu çalışmanın öğrenim çıktısı, ileri düzey öğrencilerin kendi

yazdıkları metinlerden oluşan derlemi, ana dili konuşucularının yazma örnekleriyle karşılaştırmaları yoluyla alternatif ifade biçimlerini öğrenmelerini sağlamaktır.

Özel amaçlı dil öğrenen derlemlerinin dilbilimsel analizi söylem ve metasöylem çözümlemesine dayanır (Granger ve Paquot, 2013; Flowerdew, 2015). Örneğin, Luzón (2009) İspanyol öğrencilerin yazdığı 55 rapordan oluşan bir derlemde “biz” zamirinin söylem işlevlerini incelemiştir. Luzón (2009) öğrencilerin, “biz” zamirini içeren deyimsel kalıpların akademik yazı türlerinde belirli işlevleri yerine getirmek için kullanıldığı konusunda bilgi sahibi olmadıklarını tespit etmiştir. Yabancı öğrencilerin problem çözme sürecinde sözcüksel ve dilbilgisel örüntü kullanım oranlarını inceleyen Flowerdew (2008), öğrencilerin sınırlı sayıda fiili kullandıklarını, neden/etki ile sonuç/etki fiillerini karıştırdıklarını saptamıştır. Hewings ve Hewings (2002) ana dili İngilizce olmayan MBA öğrencilerinin yazdığı tezlerde meta söylemin dilbilgisel özelliklerinden biri olan ve okuyucuları iddiaların geçerliliği konusunda ikna etmek için başvurulan öngörülü “it” ile ertelenen özne kullanımlarını inceleyerek öğrencilerin bu yapıyı gereğinden fazla kullandıklarını, ayrıca önermeleri daha güçlü bir şekilde ifade ettiklerini ortaya çıkarmıştır.

7. Öğrenen derlemleri doğal dil işleme sürecinde kullanılabilir. Öğrenen derlemlerinin kullanıldığı bir başka alan, doğal dillerin kurallı yapısının çözümlenerek anlaşılmasını veya yeniden üretilmesini sağlayan Doğal Dil İşleme (Natural Language Processing – NLP) sürecidir. Meurers (2015: 537-538), öğrenen derlemlerinin NLP’de üç farklı şekilde kullanılabileceğini ileri sürer: (I) NLP araçları öğrenen derlemlerini etiketlemek ve dil edinim süreciyle öğrencilerin ihtiyaçları hakkında bilgi edinmek için kullanılabilir. (II) NLP araçları öğrenenlerin ana dilini veya yeterlilik düzeylerini belirleme gibi öğrenen dilinin belli başlı çözümlene tekniklerini sunar. (III) Öğrenen derlemlerine NLP araçlarının, özellikle bu araçların istatistik ve makine öğrenimi (machine learning) bileşenlerinin geliştirilmesi için başvurulur.

Öğrenen derlemlerinin NLP araçlarının geliştirilmesinde kullanıldığı örnek çalışmalardan biri Zechner, Higgins, Xi ve Williamson’a (2009) aittir. Araştırmacılar TOEFL®Practice Online (TPO) testinin konuşma bölümünde verilen sözlü cevaplardan oluşturulan öğrenen derleminden yararlanarak, sınav ortamında ana dili İngilizce olmayan bireylerin doğaçlama olarak yaptığı konuşmaları otomatik olarak puanlandıran Speech Rater sistemini geliştirmiştir. Mitton ve Okada (2007) İngilizceyi yabancı dil olarak öğrenen Japon öğrencilerinin yazılı anlatım örneklerinin hata çözümlenmesini yaparak Japonlara has yazım hatalarını tespit etmişler ve Britanya İngilizcesi konuşucuları için tasarlanan yazım denetleyicisini Japon öğrencilere uyarlamışlardır.

Öğrenen derlemlerinden elde edilen veriler ana dilini tanımlama alanında da yaygın olarak kullanılmaktadır. Ana dilini tanımlama, yazar veya konuşurun ana dilinin yalnızca hedef dilde ürettiği yazılı veya sözlü metinlere dayanarak otomatik olarak belirlenmesidir (Jarvis ve Paquot, 2015: 605). Bu tür çalışmalar konuşma tanıma yazılımlarının geliştirilmesine de katkı sağlamaktadır.

Sonuç

Öğrenen derlemi araştırmaları 1980’li yılların sonunda bağımsız bir disiplin olarak şekil almaya başlayan, oldukça yeni olmasına rağmen dinamik olarak gelişen bir derlem dilbilim dalıdır. Öğrenen derlemleri, dil öğrenenlerin ürettiği yazılı ve sözlü metinlerin elektronik koleksiyonları olarak tanımlanabilir. Öğrenen derlemleri belli bir amaç doğrultusunda belirli tasarım ölçütlerine göre oluşturulur. Bu ölçütler dille ilgili olanlar (biçim, tür, taraz, konu), görevle ilgili olanlar (veri toplama, çıktı, referans kullanımı, süre kısıtlılığı) ve öğrenenle ilgili olanlar (yaş, tutum, ana dili, hedef dil) olmak üzere üç başlık altında incelenebilir. Öğrenen derlemleri tek dilli/iki dilli, genel/mesleki, eşzamanlı/artzamanlı, yazılı/sözlü, küresel /yerel, ticari amaçlı/akademik amaçlı, büyük ölçekli /küçük ölçekli gibi türlere ayrılmaktadır.

Öğrenen derlemi çalışmalarında araştırmacı ve eğitimcilerin sezgilerine dayalı incelemeler yerine, özgün metin kullanımını temel alan dil verilerinin tercih edildiği görülür. Sezgiye dayalı araştırmalar ile öğrenen kitlesini temsil gücü düşük olan küçük örneklemelerin kullanıldığı çalışmalarda, araştırmacının kendi bireysel dil kullanımı, kısıtlı deneyimi, dış etkilenmeleri, önyargıları gibi incelemeyi etkileyen etmenler söz konusuysen, derlem tabanlı araştırmalarda bu tür sorunlar yaşanmamaktadır.

Öğrenen derlemlerinde en çok başvurulan iki teknik, bilgisayar destekli hata çözümlemesi ve karşılaştırmalı aradil analizidir. Bu teknikler öğrencilerin sıklıkla yaptıkları hataların tespit edilmesini sağlar, ayrıca öğrenen diline özgü özellikleri ortaya çıkararak dilbilimci ve eğitimcilerin çeşitli öğrenci grupları hakkında genellemeler yapmalarına imkân tanır. Derlem kullanımı, büyük miktarda özgün dilsel verinin kısa sürede işlenebilmesini sağladığından elde edilen bulgulara nesnel bir nitelik kazandırır. Türkçe öğrenen öğrencilerin yazılı ve sözlü anlatım örneklerinin derlem dilbilimi yöntemleri ile incelenmesi de bu alanda yapılan çalışmaların genellenebilirliğini ve nesnellliğini sağlayacaktır.

Öğrenen derlemleri, ikinci dil edinimi ile yabancı dil öğrenimi araştırmalarında çok önemli bir yere sahiptir. Çeşitli yazılımlar aracılığıyla çözümlenen öğrenen derlemleri, dil öğrenenlerin sözcüksel ve sözdizimsel düzeyde hedef dili nasıl edindikleri konusunda araştırmacılara bilgi verir. Öğrenen derlemleri farklı yaş, cinsiyet, eğitim düzeyi, meslek ve anadillere sahip olan öğrencilerin aradil özelliklerine ışık tutar. Aradil araştırmaları bir yandan dil öğrenim sürecinin, bu süreci etkileyen faktörlerin ve aradilin zihinsel temsilinin daha iyi anlaşılmasına katkıda bulunurken, diğer yandan dil öğrenme sürecinde yaşanan zorlukların giderilmesi için öğretim ve öğrenim araçlarını tasarlayan araştırmacılara kaynak sağlar.

Öğrenen derlemlerinden elde edilen veriler, dilbilgisi kitapları ve öğrenen sözlüklerine dil öğrenenlerin en sık yaptığı dilbilgisel ve sözcüksel yanlışlarla yazım ve üslup hatalarını içeren uyarı notlarının eklenmesine olanak sağlar. Ayrıca öğrenen derlemleri, bilgisayar destekli dil öğrenme araçlarının geliştirilme sürecinde de dolaylı olarak kullanılır. Öğrenen derlemlerinden ders içi etkinliklerde de yararlanılabilir. Öğrencilerden kendileri veya başka öğrenciler tarafından üretilen metinleri çözümlemeleri ya da bu metinleri ana dili konuşucularına ait metinlerle karşılaştırmaları istenebilir.

Öğrenen derlemleri müfredattaki konuların seçiminde ve sıralanışında da belirleyici bir role sahiptir. Öğrenen derlemleri, farklı dil düzeylerinde bulunan öğrencilerin ne gibi yeterliliklere sahip olduklarını, ayrıca o düzeylerin öğrencilerin gerçek hayattaki dil edinme özelliklerine tekabül edip etmediğini gösterir. Dil öğrenenlerin neler yapıp yapamadıklarını ortaya koyan öğrenen derlemleri, nelerin ölçülmesi ve değerlendirilmesi gerektiğine de ışık tutar. Dolayısıyla ölçme ve değerlendirme esnasında kullanılan sınav ve ölçeklerin şekillendirilmesine katkıda bulunur. Öğrenen derlemleri özel amaçlı dil öğretimi bağlamında da özellikle öğrencilerin pragmatik yeterliliklerini geliştirmek için kullanılabilir veya özel amaçlı dil öğretim materyallerin içeriğini şekillendirebilir.

Araştırma ve Yayın Etiği

Bu çalışmada “Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi” kapsamında uyulması belirtilen tüm kurallara uyulmuştur. Yönergenin ikinci bölümü olan “Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler” başlığı altında belirtilen eylemlerden hiçbiri gerçekleştirilmemiştir.

Yazarların Katkı Oranı

1. yazar %50, 2. yazar %50

Çıkar Çatışması

Bu çalışmanın hazırlanmasında ve yayınlanmasında çıkar çatışması teşkil edebilecek bir durum ve ilişki yoktur.

Kaynaklar

- Aarts, J., & Meijs, W. (1984). *Corpus Linguistics: recent developments in the use of computer corpora in English language research*. Amsterdam: Rodopi.
- Adamska-Sałaciak, A. (2012). Dictionary definitions: Problems and solutions. *Studia Linguistica Universitatis Iagellonicae Cracoviensis*, 129(4), 323-339.
- Adolphs, S., & Knight, D. (2010). Building a spoken corpus: what are the basics? A. O’Keeffe ve M. McCarthy (Ed.), *The Routledge handbook of corpus linguistics* (pp. 38-52). London: Routledge.

- Agan, C., & Diri, B. (2016). Türkçe derlemler için söz dizimsel görselleştirme ve sorgulama aracı. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 9(1), 1-10.
- Akın, L. (2015). Konuşma dili grameri ve sözlü derlemlerden yararlanarak konuşma dili özelliklerinin tespiti. *Türkiyat Mecmuası*, 25, 1-27.
- Aksan D. (2011). Derlem dilbilim ve dil eğitimi. L. Uzun ve Ü. Bozkurt (Ed.), *Türkçenin eğitimi-öğretiminde kuramsal ve uygulamalı araştırmalar* (s. 349-343). Essen: Verlag Die Blaue Eule.
- Barker, F. (2010). How can corpora be used in language testing? A. O'Keeffe ve M. McCarthy (Ed.), *The Routledge handbook of corpus linguistics* (pp. 633-645). London: Routledge.
- Barker, F., Salamoura, A., & Saville, N. (2015). Learner corpora and language testing. S. Granger, G. Gilquin ve F. Meunier (Ed.), *The Cambridge handbook of learner corpus research* (pp. 511-534). Cambridge: Cambridge University Press.
- Biber, D., Conrad, S. ve Reppen, R. (1998). *Corpus linguistics. Investigating language structure and use*. Cambridge: Cambridge University Press.
- Blass, L., Iannuzzi, S., Reppen, R., & Savage, A. (2012). *Grammar and beyond level 3 student's book*. Cambridge: Cambridge University Press.
- Braun, S. (2006). ELISA - a pedagogically enriched corpus for language learning purposes. S. Braun, K. Kohn ve J. Mukherjee (Ed.), *Corpus technology and language pedagogy: New resources, new tools, new methods* (pp. 25-47). Frankfurt/M: Peter Lang.
- Callies, M. (2015). Learner corpus methodology. S. Granger, G. Gilquin ve F. Meunier (Ed.), *The Cambridge handbook of learner corpus research* (pp. 35-56). Cambridge: Cambridge University Press.
- Callies, M., Díez-Bedmar, M. B., & Zaytseva, E. (2014). Using learner corpora for testing and assessing L2 proficiency. P. Leclercq, H. Hilton, A. Edmonds (Ed.), *Proficiency assessment measures in SLA research: measures and practices* (pp.71-90). Clevedon: Multilingual Matters.
- Can, C. (2009). İkinci dil edinimi çalışmalarında bilgisayar destekli bir Türk öğrenici İngilizcesi derlemi: ICLE'nin bir alt derlemi olarak TICLE. *Dil Dergisi*, 144, 16-34.
- Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English: A comprehensive guide*. Cambridge: Cambridge University Press.
- Chambers, A. (2015). The learner corpus as a pedagogic corpus. S. Granger, G. Gilquin ve F. Meunier (Ed.), *The Cambridge handbook of learner corpus research* (pp. 445-464). Cambridge: Cambridge University Press.
- Chomsky, N. (2004). The master and his performance: An interview with Noam Chomsky. *Intercultural Pragmatics*, 1(1), 93-111.
- Chuang, F., & Nesi, H. (2007). Grammar talk: Developing computer-based materials for Chinese EAP students. O. Alexander (Ed.), *New Approaches to Materials Development for Language Learning* (pp. 315-330). Oxford: Peter Lang.
- Cobb, T. (2003). Analyzing late Interlanguage with learner corpora: Québec replications of three European studies. *Canadian Modern Language Review/ La Revue canadienne des langues vivantes*, 59(3), 393-424.
- Cock, S., Granger, S, Leech, G., & McEnery, T. (1998). An automated approach to the phrasicon of EFL learners. S. Granger (Ed.), *Learner English on computer* (pp. 67-79). London: Longman.
- Cowan, R., Choo, J., & Lee, G. S. (2014). ICALL for improving Korean L2 writers' ability to edit grammatical errors. *Language Learning & Technology*, 18(3), 193-207.
- Flowerdew, L. (2008). *Corpus-based analyses of the problem solution pattern*. Amsterdam/Philadelphia: John Benjamins.
- Flowerdew, L. (2015). Learner corpora and language for academic and specific purposes. S. Granger, G. Gilquin ve F. Meunier (Ed.), *The Cambridge handbook of learner corpus research* (pp. 465-484). Cambridge: Cambridge University Press.
- Gablasova, D., Brezina, V., & McEnery, T. (2017). Exploring learner language through corpora: Comparing and interpreting corpus frequency information. *Language Learning*, 67(S1), 130-154.

- Granger, S. (2002). A Bird's-eye view of computer learner corpus research. S. Granger, J. Hung ve S. Petch-Tyson (Ed.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 3-33). Amsterdam and Philadelphia: Benjamins.
- Granger, S. (2008). Learner corpora. A. Lüdeling ve M. Kytö (Ed.), *Corpus linguistics: An international handbook* (pp. 259-275). Berlin: Mouton de Gruyter.
- Granger, S. (2015). The contribution of learner corpora to reference and instructional materials design. S. Granger, G. Gilquin ve F. Meunier (Ed.), *The Cambridge handbook of learner corpus research* (pp.486-510). Cambridge: Cambridge University Press.
- Granger, S., & Paquot, M. (2013). Language for specific purposes learner corpora. C.A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 3142-3146).Oxford: Wiley.
- Guilquin, G. (2015). From design to collection of learner corpora. S. Granger, G. Gilquin ve F. Meunier (Ed.), *The Cambridge handbook of learner corpus research* (pp. 9-34). Cambridge: Cambridge University Press.
- Harrison, J. (2015). The English grammar profile. J. Harrison ve F. Barker (Ed.), *English profile in practice* (pp. 28-48). Cambridge: Cambridge University Press.
- Hegelheimer, V., & Fisher, D. (2006). Grammar, writing, and technology: A sample technology-supported approach to teaching grammar and improving writing for ESL learners. *CALICO Journal*, 23(2), 257-279.
- Heift, T. (2010). Developing an intelligent language tutor. *CALICO Journal*, 27(3),443-459.
- Heuberger, R. (2016). Learner's dictionaries: History and development; current issues. P. Durkin (Ed.), *The Oxford handbook of lexicography* (pp. 25-43). Oxford: Oxford University Press.
- Hewings, M.,& Hewings, A. (2002). "It is interesting to note that ...": a comparative study of anticipatory 'it' in student and published writing. *English for Specific Purposes*, 21(4), 367-383.
- Jarvis, S., & Paquot, M. (2015). Native language identification. S. Granger, G. Gilquin ve F. Meunier (Ed.), *The Cambridge handbook of learner corpus research* (pp. 605-627). Cambridge: Cambridge University Press.
- Lee, D. Y.W. (2010). What corpora are available? A. O'Keeffe ve M. McCarthy (Ed.), *The Routledge handbook of corpus linguistics* (pp. 107-121). London: Routledge.
- Lee, D.,& Swales, J. (2006). A Corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*, 25(1), 56-75.
- Leech, G. (1992). Corpora and theories of linguistic performance. J. Svartvik (Ed.), *Directions in corpus linguistics. Proceedings of Nobel symposium 82* (pp. 105–122). Berlin: Mouton de Gruyter.
- Luzon, M. J. (2009). The use of "we" in a learner corpus of reports written by EFL engineering students. *Journal of English for Academic Purposes*, 8(3), 192-206.
- Mark, K.L. (1998). The Significance of learner corpus data in relation to the problems of language teaching. *Bulletin of General Education*, 312, 77-90.
- McEnery, T., Brezina, V., Gablasova, D., & Banerjee, J. (2019). Corpus linguistics, learner corpora, and SLA: Employing technology to analyze language use. *Annual Review of Applied Linguistics*, 39, 74–92.
- McEnery, T., & Wilson, A. (2001). *Corpus linguistics*. Edinburgh: Edinburgh University Press.
- McEnery, T., Xiao,R.Z., & Tono, Y.(2006). *Corpus-based language studies: An advanced resource book*. London: Routledge.
- Meunier, F. (2016). Learner corpora and pedagogic applications. A. O'Keeffe ve M. McCarthy (Ed.), *The Routledge handbook of language learning and technology* (pp. 376-387). London: Routledge.
- Meurers, D. (2015). Learner corpora and natural language processing. S. Granger, G. Gilquin ve F. Meunier (Ed.), *The Cambridge handbook of learner corpus research* (pp. 537-565). Cambridge: Cambridge University Press.

- Milton, J.(2006). Resource-rich Web-based feedback: Helping learners become independent writers. K. Hyland ve G. Hyland (Ed.), *Feedback in second language writing* (pp. 123-139). Cambridge: Cambridge University Press.
- Mitton, R.,& Okada, T. (2007). *The adaptation of an English spellchecker for Japanese writers*. London: Birkbeck ePrints. Erişim adresi: <https://eprints.bbk.ac.uk/id/eprint/592/>
- Mukherjee, J.,& Rohrbach, J.-M. (2006). Rethinking applied corpus linguistics from a language-pedagogical perspective: New departures in learner corpus research. B. Kettemann ve G. Marko (Ed.), *Planing, gluing and painting corpora: Inside the applied corpus linguist's workshop* (pp. 205-232). Frankfurt: Peter Lang.
- Nelson, M. (2010). Building a written corpus: what are the basics? A. O'Keeffe ve M. McCarthy (Ed.), *The Routledge handbook of corpus linguistics* (pp. 53-65). London: Routledge.
- Nesselhauf, N. (2006). Researching L2 production with ICLE. S. Braun, K. Kohn ve J. Mukherjee (Ed.). *Corpus technology and language pedagogy: New resources, new tools, new methods* (pp. 141-156). Frankfurt: Lang.
- O'Donnell, M. (2013). From learner corpora to curriculum design: An empirical approach to staging the teaching of grammatical concepts. *Procedia - Social and Behavioral Sciences*, 95,571-580.
- Ostermann, C. (2015). *Cognitive lexicography: A new approach to lexicography making use of cognitive semantics*. Berlin/Boston: De Gruyter Mouton.
- Petch-Tyson, S. (1998). Writer/reader visibility in EFL written discourse. S. Granger (Ed.), *Learner English on computer* (pp. 107-118). London: Longman.
- Ringbom, H. (1998). Vocabulary frequencies in advanced learner English: Across-linguistic approach. S. Granger (Ed.), *Learner English on computer* (pp. 41-52). London: Longman.
- Ruhi, Ş., Aksan, M., & Aksan, Y. (2014). Derlem dilbilim yöntemlerinin etkin olarak araştırmalarda kullanımı: Uygulamalar. *Proceedings of the 27th national linguistics conference* (ss. 264-270). Ankara: Hacettepe Üniversitesi.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10(3),209-231.
- Shohamy, E. (2008). Introduction to volume 7: Language testing and assessment. E. Shohamy ve N. Hornberger (Ed.), *Language testing and assessment: Encyclopedia of language and education*, 7 (pp. xiii-xxii). New York: Springer.
- Stubbs, M. (1993). British traditions in text analysis: From Firth to Sinclair. M. Baker, G. Francis ve E. Tognini Bonelli (Ed.), *Text and technology: In honour of John Sinclair* (pp.1-36). Amsterdam: John Benjamins.
- Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Oxford: Blackwell.
- Stubbs, M. (2004). Language corpora. A. Davies ve C. Elder (Ed.). *The handbook of applied linguistics* (pp. 106-132). London: Blackwell Publishing.
- Taylor, C. (2008). What is corpus linguistics? What the data says. *ICAME Journal*, 32, 179-200.
- Thompson, G., & Hunston, S. (2006). System and corpus: two traditions with a common ground. G. Thompson ve S. Hunston (Ed.), *System and corpus: Exploring connections* (pp. 1-14). Bristol: Equinox.
- Teubert, W. (2005). My version of corpus linguistics. *International Journal of Corpus Linguistics*, 10(1), 1-13.
- Teubert, W., & Krishnamurthy, R. (2007). General introduction to corpus linguistics. W. Teubert ve R. Krishnamurthy (Ed.), *Critical concepts in linguistics* (pp. 1-37). London and New York: Routledge.
- Tognini Bonelli E. (2010). Theoretical overview of the evolution of corpus linguistics. A. O'Keeffe ve M. McCarthy (Ed.), *The Routledge handbook of corpus linguistics* (pp. 14-27). London: Routledge.
- Tono, Y. (2003). Learner corpora: Design, development and applications. D. Archer, P. Rayson, A. Wilson ve T. McEnery (Ed.), *Proceedings of the 2003 Corpus Conference* (pp. 800-809). Erişim adresi: <http://ucrel.lancs.ac.uk/publications/cl2003/papers/tono.pdf>

- Walter, E. (2010). Using corpora to write dictionaries. A. O’Keeffe ve M. McCarthy (Ed.), *The Routledge handbook of corpus linguistics* (pp. 428-443). London: Routledge.
- Zechner, K., Higgins, D., Xi, X., & Williamson, D.M. (2009). Automatic scoring of non-native spontaneous speech in tests of spoken English. *Speech Communication*, 51(10), 883-895.

Extended Abstract

Introduction

Corpora are electronically stored and processed collections of written and/or spoken language data compiled by means of balanced and layered sampling and representing a specific language or language variety. There are different types of corpora according to the design criteria and the purpose for which a corpus is built:

1. Sample Corpus: A sample corpus is a “snapshot” of a particular language in a particular period of time. The main purpose of a sample corpus is to reflect the standard features of a language or language variety similar to those that can be found in general use. It is also called a reference corpus as it displays all the features of the given language.

2. Comparable Corpora: Two or more corpora are comparable when they are based on the same design criteria or are of similar size. These are called comparable corpora.

3. General Language and Specialized Corpora: General language corpora consist of texts that do not belong to a single text genre, subject area, or register. Specialized corpora, on the other hand, instead of representing the language as a whole, only represent a certain segment of the language.

4. Written and Spoken Corpora: Written corpora are compiled of written texts of a particular language. Spoken corpora are created by transcribing the spoken data obtained by means of sound recordings in various areas of language use. Spoken corpora are smaller than written ones, therefore they cannot ensure that some phrases are repeated at the same rate as in written corpora.

5. Diachronic / Historical and Monitor Corpora: Diachronic corpus is a collection of texts from different periods of time used to examine the development of language (Lee, 2010, p. 113). Monitor corpus aims to keep track of the language development by continuously adding new content and comparing this content with the reference corpus that has a static character (Walter, 2010, p. 430).

6. Bilingual and Multilingual Corpora: Bilingual and multilingual corpora were originally developed to meet the needs of machine translation, which was the main goal of computer linguistics. A parallel corpus is a corpus consisting of two monolingual corpora, one of which is a translation of the other. The most well-known parallel corpus is the Canadian Hansard Corpus, which consists of the Canadian Parliament debates published in English and French, which are the official languages of Canada (Tognini Bonelli, 2010, p. 23).

7. Parsed Corpora / Treebanks: A treebank is a parsed text corpus that annotates syntactic or semantic sentence structure. Since it is difficult to ensure a consistent approach in tagging of an entire corpus, such corpora are usually small in size (Agan & Diri, 2016, p. 1).

8. Multimedia Corpora: Multimedia corpora are collections of transcripts synchronized with the original audio or video recordings. This type of corpora allows researchers and users to access auditory and visual elements of discourse such as prosody, gestures, gaze, interval, etc. (Lee, 2010, pp. 114-115).

9. Pedagogic Corpora: Pedagogic corpora are large and representative collections of spoken and written language which learners are exposed to or are likely to be exposed to through various channels. These corpora include written and spoken texts that learners may encounter inside and outside the classroom as well as in the course of technology-supported communication (Chambers, 2015, pp. 447-448).

10. Developmental Language Corpora: Developmental language corpora are collections of texts produced by native-speaking children at various stages of development (Lee, 2010, p. 116).

Learner Corpora

Defining the target language, education environments (task, curriculum, teaching methods, etc.) and learner characteristics (motives, learning styles, attitudes, etc.) have been the most studied

aspects of language teaching. As for the learner language, it has always been ignored by researchers and educators (Mark, 1998, p. 84).

Learner corpus research refers to research conducted on corpora representing foreign or second language use (McEnery et al., 2019, p. 76). Learner corpus research, which emerged in the late 1980s, bridged the gap between the two fields that developed independently of each other, namely, corpus linguistics and second/foreign language research.

According to Callies (2015, p. 36), most of the learner corpus studies today are corpus-based, quantitative, cross-sectional and comparative. Learner corpora are used within such methodological approaches as contrastive interlanguage analysis and computer-aided error analysis. Contrastive interlanguage analysis involves carrying out quantitative and qualitative comparisons between native and non-native data or between different varieties of non-native data, while computer-aided error analysis focuses on errors in interlanguage and uses computer tools to tag, retrieve and analyze them.

Meunier (2016, p. 476) summarized the main objectives of learner corpus research as follows: a. Describing the language use and linguistic development of learners through cross-sectional and longitudinal approaches. b. Shedding light on second language acquisition processes. c. Developing pedagogical materials and data-driven learning activities.

Result and Discussion

The results of research on learner corpora demonstrate that learner corpora can be used in teaching Turkish in the following ways:

1. Learner corpora are used in interlanguage analysis. They shed light on the interlanguage of learners of different ages and genders, coming from various educational, professional and L1 (first language) backgrounds.

2. Learner corpora are used during the design process of reference materials to be used in lessons. They give information about the grammatical, lexical, spelling, and stylistic errors made by language learners. This information about language learners contributes to the process of development of learner dictionaries, textbooks, and grammar resources.

3. Learner corpora provide language learners with examples of the use of the target language. Some of the learner corpora are compiled for the direct use of language learners. These are called "pedagogic corpora". In this case corpora are prepared by teachers and subsequently made available to students. Therefore, learners both provide and benefit from the corpus data (Granger, 2015, p. 486).

4. Learner corpora inform the curriculum design. They play an important role in the selection and ordering of teaching content in the curriculum. The data obtained from the learner corpora shed light on the words and grammatical structures that cause problems for students and therefore require more time and attention from teachers.

5. Learner corpora have the potential to contribute to language testing and assessment. They demonstrate what students at a certain level of proficiency can do, which helps to shape assessment scales. In addition, learner corpora contribute to the design of language exams by revealing to what extent demographic variables (gender, mother tongue, language learning experience, etc.), the testing mode (paper/computer-based) and learning environment affect learners' output (Barker et al., 2015, pp. 521-524).