



*Uluslararası Türkçe Edebiyat Kültür Eğitim Dergisi Sayı: 9/3 2020 s. 1018-1032, TÜRKİYE*

*Araştırma Makalesi*

## **SOKAKTAKİ ADAM ROMANINDAKİ ANLATICILARIN KULLANDIĞI SÖZCÜK KADROSUNUN SIKLIK VE DAĞILIMININ BİLGİSAYARLI İNCELEMESİ**

**Cengiz DEMİR\***

*Geliş Tarihi: Mart, 2020*

*Kabul Tarihi: Temmuz, 2020*

### **Öz**

Bilgisayarlı nicel metin incelemesi, okuma yoluyla elde edemeyeceğimiz verilere ve sonuçlara ulaşmamızı sağlar. Bilgisayar yardımıyla dilsel birimlerin sıklığı ve dağılımı ölçülür, dahası, çeşitli karşılaştırma ve istatistiksel incelemeler yapılır. Attilâ İlhan'ın *Sokaktaki Adam* adlı romanı 22 bölümden oluşur ve her bir bölüm dönüşümlü olarak 7 ayrı anlatıcı tarafından anlatılır. Farklı sosyal çevrelerden gelen bu anlatıcıların her biri için ayrı bir bireysel dilin gerekliliği ortaya çıkar. Bu çalışmada, anlatıcıların bireysel diline ait özellikleri ortaya çıkarmak için, anlatıcılara ait sözcük sıklık listesi, sözcük bulutları ve sözcük dağılım grafikleri çıkarılmış; sözcük uzunluğu, sözcük yoğunluğu ve tümce uzunluğu hesaplanmıştır. Bu çalışmanın amacı bilgisayar desteği ile metin incelemesinin üslup çalışmalarına ne derece katkısı olabileceğini göstermektir.

**Anahtar Sözcükler:** *Sokaktaki Adam*, anlatıcı, bireysel dil, sözcük sıklığı.

### **THE COMPUTATIONAL ANALYSIS OF THE FREQUENCY AND DISPERSION OF THE VOCABULARY USED BY THE NARRATORS IN THE NOVEL *SOKAKTAKİ ADAM***

#### **Abstract**

A computer-assisted quantitative text analysis enables us to reach the data and results that cannot be reached through reading. The frequency and dispersion of linguistic units are measured, moreover, various comparisons and statistical analyses can be made with computers. The novel *Sokaktaki Adam* by Attilâ İlhan has 22 chapters and each chapter is narrated sequentially by a narrator who comes from a different social background. Therefore, it is necessary to examine an idiolect for each of these narrators who represent different social environments. In this study, to reveal the idiolectic characteristics of the narrators, word frequency lists, word clouds and word dispersion plots were generated; word length, word density and sentence length were calculated. The objective of this study is to see, to what extent, the contribution to the stylistic studies.

**Keywords:** *Sokaktaki Adam*, narrator, idiolect, word frequency.

---

\* Dr.; Kocaeli Üniversitesi, Sosyal Bilimler Enstitüsü, cendemir@gmail.com

## Giriş

Nicel (sayısal) metin incelemesi düşüncesini ilk defa matematikçi Augustus de Morgan (1806-1871) ortaya atarak yazarın tercih ettiği sözcüklerin uzunluklarının ve sıklıklarının ortalamasının ölçümüyle yazarın kendine özgü üslup özelliğine ulaşılacağını savunmuştur. Bu düşüncenin ilk uygulamasını 1887 ve 1901 yıllarında iki ayrı çalışma olarak Mendenhall (1841-1924) yapmıştır. O günün şartlarında, ücretli iki sekreter tutup tek tek sayarak Shakespeare ve Bacon'ın tüm eserlerinin sözcük sıklık şemasını çıkararak karşılaştırma yapmış, Shakespeare ve Bacon'ın farklı üslup özelliklerine sahip olduğunu iddia etmiştir.

*The Statistical Study of Literary Vocabulary* adıyla bu sahadaki ilk kapsamlı çalışmalardan birini yapan Udney Yule (1871-1951) ise 1938'de yazarın ayırt edici özelliğinin tümce uzunlukları olduğu düşüncesiyle Bacon, Coleridge ve Macaulay üzerine çalışma yapmış, 1944'te ise çalışmasını sözcükler üzerine yoğunlaştırmıştır (Kenny, 1982, s. 1-5).

Bilgisayar destekli ilk çalışma 1948 yılında Roberto Busa (1913-2011) tarafından yapılmıştır. Aslen Cizvit bir rahip olan Busa, IBM şirketinin kurucusu Thomas J. Watson'ı sponsorluk için ikna ederek Thomas Aquinas'ın tüm çalışmalarını 65.000 sayfa olarak bilgisayar ortamına aktarmış ve *The Index Thomisticus* adlı çalışmayı yapmıştır.

Bugüne kadar felsefi, dinî ve edebî pek çok metin üzerine bilgisayar destekli çalışmalar yapılmıştır ancak 1967 tarihli Luis Milic'in *A Quantitative Approach to the Style of Jonathan Swift* ve J. F. Burrows'un Jane Austin'in romanlarını incelediği 1987 tarihli *Computation into Criticism* özellikle vurgulanması gereken öncü çalışmalardır.

Bilgisayarın akademik çevrelerde yoğun kullanımıyla *bilgisayarlı üslupbilim* (*computational stylistics*) adı öne çıksa da, *nicel üslupbilim* (*quantitative stylistics*) için *istatistiksel üslupbilim* (*stylostatistics*) ve *üslupölçüm* (*stylometry*) adları da kullanılmaktadır. Bilgisayar destekli nicel metin incelemesi özellikle uzun metinlerin, derlemin (*corpus*) ve belli bir tür ya da dönemin incelenmesinde oldukça kullanışlıdır. Sözcüksel ya da dil bilgisel birimin sıklığı, metin ve derlem içerisindeki dağılımı bilgisayar desteğiyle ölçülerek çeşitli istatistiksel yöntemlerle yorumlanıp karşılaştırmalar yapılabilmektedir.

Bir metin yazarının üslubu mevcut dil normları içerisinde yapmış olduğu çeşitlemeler (*variation*) ya da sapmaların (*deviation*) tespit edilmesiyle ortaya konur. Bu çeşitleme ve sapmalar, mevcut dil normları içerisinde yazarın bilinçli tercihinin ya da bilinçsiz alışkanlığının ürünleridir. Yazarın bu üslup özelliklerini belirleyebilmek için metin nitel ya da nicel inceleme yöntemleriyle incelenir.

Nitel metin incelemelerinde (*qualitative text analysis*) *ne, nasıl, niçin* sorularının cevabı aranırken, nicel metin incelemelerinde (*quantitative text analysis*) *ne kadar ve ne sıklıkta* sorularının cevabı aranmaktadır.

Nicel metin incelemesinde kişiye özgü üslup belirleyicileri (*style markers*) olarak sözcüklerin ya da dil bilgisel yapıların metindeki dağılımı ve sıklığı belirlenir. İntihal, telif davası gibi hukuksal süreçlerde, meçhul kalmış ve yazarının kim olduğu tartışmalı bir metnin yazarının bulunmasında bu üslup özelliklerinden yola çıkılmaktadır. Nicel metin incelemesi nesnel yöntemleriyle nitel metin incelemesinde ortaya konan görüşleri desteklemekte, okuyarak elde edilemeyecek veri ve sonuçlara ulaşılmasını sağlamaktadır.

### Sokaktaki Adam

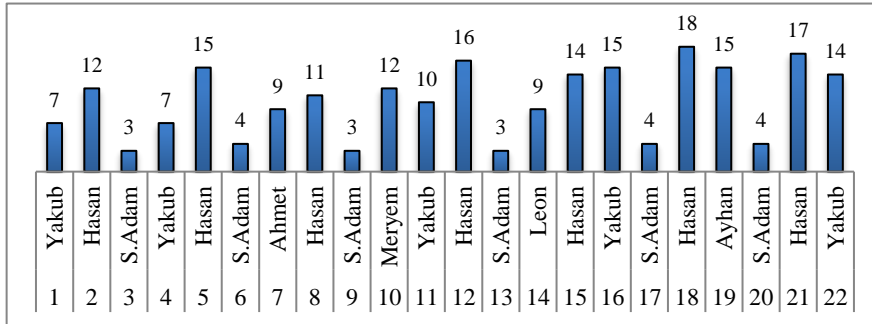
Nicel metin incelemesi yöntemiyle ne gibi sonuçlara ulaşılabileceğini görmek için Attilâ İlhan'ın *Sokaktaki Adam* adlı romanı üzerinde R istatistik programı, Excel ve çeşitli çevrim içi araçlarla bilgisayar destekli bir çalışma yapılmıştır. Bu çalışmada, 1953 yılında kitap olarak yayımlanan *Sokaktaki Adam* romanının 2019 yılında İş Bankası Kültür Yayınları arasında çıkan son baskısı kullanılmıştır. Yazarın 1968 yılında eklediği “Roman Dedik” adlı giriş yazısı ile birlikte toplam 240 sayfa olarak basılan eserin, 222 sayfa olan roman kısmı 22 bölümden oluşur.

Anlatıcı, romanda geçen olay ve durumları okuyucuya aktaran bir ara kişidir. Anlatıcı, romancının kendisi olabileceği gibi, roman kahramanlarından biri de olabilir. Anlatıcının kim olacağı romancının seçimidir ve bu seçim romanın gerçekçiliğini ve üslubunu etkiler (Şen, 2009, s. 558). *Sokaktaki Adam*'da her bir bölüm ayrı bir anlatıcı tarafından aktarılmak suretiyle çoğul anlatıcı tekniği kullanılmıştır. Çoğul anlatıcı (*multiple narrators*) roman kişilerinin kendi kendilerini ayrı ayrı anlatması, bir romanda birden çok anlatıcı tipine yer verilmesidir (Çetin, 2017, s. 115). Çoğul anlatıcı ile kurulan romanlarda farklı figürler aynı hikayeyi değişik bakış açılarından aktarır (Sazyek, 2015, s. 101; Tekin, 2001, s. 63).

Romanın olay örgüsü kısaca şöyledir: Romanın başkışisi Hasan, üniversite öğrenimini yarıda bırakıp kamarot olarak gemilerde çalışmaktadır. Arkadaşı Yakub ile birlikte yurtdışından kaçak kürk getirip satmaktadır. Çalıştıkları gemi İstanbul'a demirleyince getirdikleri kürkleri satmak isterler ama anlaştıkları alıcılardan Nubar tutuklanmıştır. Hasan diğer alıcı Leon'un ücreti karşılığında kürkleri geri götürme teklifini reddeder. Leon'u aradan çıkaran Triyandafilos Ahmet, büyük patron Şevket Nuri'yi ikna ederek kürkleri Leon'un metresi Meryem'in evinde Hasan'dan teslim alır. Hasan bu süreçte hayat kadını olan Meryem'le yakınlık kurar. Ahmet tutuklanır ve Leon İsrail'e kaçar. Tutuklanmadan önce Hasan üniversiteden âşık olduğu ve unutamadığı Ayhan'la da yüzleşir ve yolunu ayırır. Hasan bir sokak kavgasını ayırmaya çalışırken yaralanır ve ölür.

### Anlatıcılar ve Anlatıdaki Payları

Romanda dışöyküsel anlatıcı yoktur. Roman, Yakub'un “Ben Yakub’um: Kamarot Yakub” sözleriyle başlar ve her bir anlatıcı kendini tanıtarak anlatısını yapar. Anlatıcılar sırasıyla şunlardır: Yakub (kamarot / Hasan'ın arkadaşı), Hasan (başkışı / kamarot), Sokaktaki Adam (Olay örgüsünde yeri yoktur, sadece sokaktaki vatandaş olarak araya girer), Triyandafilos Ahmet (kaçakçı / Şevket Nuri'nin adamı), Meryem (hayat kadını), Leon (kürk tüccarı), Ayhan (Hasan'ın üniversiteden kız arkadaşı).

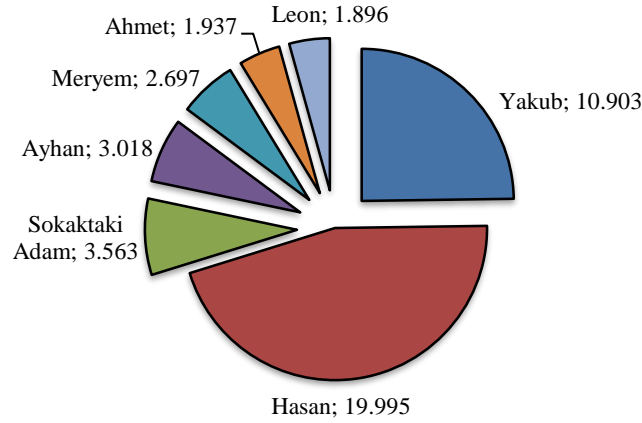


Şekil 1: Romanı oluşturan bölümler, anlatıcıları ve sayfa miktarları

Romanın hangi bölümünde hangi anlatıcının devreye girdiği ve her bir bölümün kaç sayfa yer tuttuğu Şekil 1'deki grafikte görülmektedir.

Roman; 1, 4, 11, 16 ve 22. bölümlerin anlatıcısı Yakub'un anlatımıyla başlar ve Yakub'la biter. Romanın başkışisi Hasan ise 2, 5, 8, 12, 15, 18 ve 21. bölümlerin anlatıcısıdır. Sokaktaki Adam adlı anlatıcı 3, 6, 9, 13, 17 ve 20. bölümlerde araya girerek olay örgüsünün dışında bazen sokaktaki biriymiş gibi anlatıda bulunur.

Yakub 53, Hasan 103, Sokaktaki Adam 21, Ahmet 9, Meryem 12, Leon 9 ve Ayhan 15 sayfa tutarında anlatıda bulunmuştur. Yaklaşık olarak anlatının 2/4'ü romanın başkışisi Hasan, 1/4'ü Yakub ve geriye kalan 1/4'ü ise diğer anlatıcılar tarafından yapılmıştır.



Şekil 2: Anlatıcılara ait tüm sözcük sayıları

Şekil 2'de görüldüğü üzere anlatıcı başına düşen toplam sözcük sayısı 1.895'ten 19.995'e kadar değişiklik göstermektedir. 1.895 olan Leon'un sözcük sayısı en az olduğundan, sağlıklı bir inceleme yapabilmek için diğer anlatıcıların sözcük sayıları da 1.895 olarak sınırlandırılmıştır.

Anlatıcılardan kesit metin alındıktan sonra aşağıdaki ölçüm ve uygulamalar yapılmıştır:

- Sözcük sıklık listesi (*word frequency list*)
- Sözcük yoğunluğu (*lexical density*)
- Sözcük uzunluğu (*word length*)
- Sözcük bulutu (*word cloud*)
- Dağılım grafikleri (*dispersion plots*)
- Sözcük dizisi (*n-gram*)
- Tümce uzunluğu (*sentence length*)

### Sözcük Sıklık Listesi

İlk önce, anlatıcılara ait olan her bir metin birim sözcüklere ayrılarak (*tokenization*) sözcük sıklık listesi (*word frequency list*) çıkarılmıştır. Sözcük sıklık listesinin ilk 20 sözcüğü ve tüm romandan alınan en sık 20 sözcük Tablo 1'de görülmektedir. Tablo incelendiğinde *bir*,

*bu, ben, o, ne, de* gibi işlevsel sözcüklerin listenin başında olduğu görülür. Dildeki işlevsel sözcüklerin en sık olması olağan bir durumdur. Bazı üslupbilim uzmanları yazarın tercihini yansıtan içerik sözcüklerin daha önemli olduğunu, işlevsel sözcüklerin listelenmeden inceleme yapılması gerektiğini iddia etseler de özellikle yazar belirleme çalışmalarında, yazarın bilinçsiz alışkanlık ürünü olan işlevsel sözcüklerin önemli rolü olmaktadır.

Tablo 1: Anlatıcıların ve tüm romanın sözcük sıklık listelerindeki ilk 20 sözcük

|       | Yakub | Hasan  | S.Adam | Ahmet | Meryem | Leon   | Ayhan | Roman   |    |        |    |       |    |       |      |
|-------|-------|--------|--------|-------|--------|--------|-------|---------|----|--------|----|-------|----|-------|------|
| bir   | 41    | bir    | 63     | bir   | 60     | bir    | 45    | bir     | 42 | bir    | 33 | bir   | 56 | bir   | 1458 |
| bu    | 27    | ve     | 38     | ben   | 41     | bu     | 36    | ne      | 37 | ben    | 30 | ne    | 40 | ve    | 612  |
| ne    | 27    | bu     | 32     | ve    | 30     | ne     | 23    | o       | 26 | ne     | 30 | Hasan | 29 | ben   | 499  |
| yok   | 27    | için   | 22     | bu    | 20     | heh    | 20    | da      | 24 | bu     | 18 | bu    | 21 | bu    | 478  |
| ben   | 24    | ben    | 19     | her   | 17     | o      | 20    | bu      | 20 | neden  | 18 | o     | 21 | ne    | 439  |
| ama   | 22    | de     | 18     | da    | 16     | de     | 19    | ben     | 19 | o      | 17 | gibi  | 19 | gibi  | 407  |
| mi    | 20    | gibi   | 16     | benim | 15     | da     | 17    | gibi    | 18 | be     | 16 | ben   | 17 | de    | 315  |
| de    | 18    | ne     | 15     | de    | 14     | ben    | 15    | ama     | 17 | mı     | 15 | beni  | 16 | o     | 312  |
| o     | 18    | Meryem | 14     | sonra | 13     | ama    | 12    | ya      | 17 | de     | 14 | diye  | 16 | da    | 288  |
| ulan  | 17    | ya     | 14     | gibi  | 10     | mi     | 12    | yok     | 16 | ki     | 14 | mı    | 15 | Hasan | 227  |
| mı    | 16    | o      | 12     | gün   | 9      | Meryem | 11    | mi      | 14 | var    | 14 | de    | 14 | diye  | 210  |
| gibi  | 15    | kadar  | 11     | iki   | 9      | para   | 11    | mı      | 11 | yok    | 14 | var   | 14 | mi    | 207  |
| Hasan | 12    | da     | 10     | kadar | 9      | yok    | 11    | daha    | 10 | mi     | 13 | ve    | 14 | ama   | 196  |
| var   | 12    | bana   | 9      | bazan | 8      | benim  | 10    | de      | 10 | başka  | 12 | bana  | 13 | yok   | 187  |
| be    | 11    | beni   | 9      | o     | 8      | biz    | 10    | değil   | 10 | şimdi  | 12 | onu   | 13 | mı    | 186  |
| değil | 11    | başka  | 8      | ya    | 8      | daha   | 10    | şekerim | 10 | ama    | 11 | ama   | 12 | sonra | 186  |
| ağbiy | 10    | böyle  | 8      | adam  | 7      | için   | 10    | var     | 10 | Allah  | 9  | dedi  | 12 | ya    | 173  |
| bak   | 10    | değil  | 8      | beni  | 7      | ki     | 10    | iki     | 9  | böyle  | 9  | diyor | 12 | daha  | 172  |
| da    | 10    | büyük  | 7      | çünkü | 7      | onun   | 10    | mu      | 9  | herkes | 9  | mi    | 12 | var   | 166  |
| mu    | 10    | hayır  | 7      | fakat | 7      | ya     | 10    | sonra   | 9  | hiç    | 9  | sonra | 12 | için  | 154  |

### Sözcük Yoğunluğu

Her bir anlatıcının sözcük yoğunluğuna (*type/token ratio*) bakılmıştır. “*Damlaya damlaya göl olur*” tümcesi 4 birim sözcükten (*token*) ve 3 tek sözcükten (*type*) oluşur. Sözcük yoğunluğu, metinde tekrarsız tek sözcüklerin (*type*) metinde bulunan tüm birim sözcüklere (*token*) bölünerek 100 ile çarpılmasıyla elde edilir.

İncelenen metin ya da derlem büyüdükçe sözcük yoğunluğu oranı düşer çünkü metin uzadıkça birim sözcükler (*tokens*) sürekli artarken tekrarlardan dolayı tek sözcükler (*types*) sürekli artmaz. Yazarın metin içerisindeki bütünlüğü yakalayabilmek için önceden kullandığı sözcükleri tekrarlamak zorunda olması ve dildeki söz varlığının sınırlı oluşundan dolayı, yazar yeni sözcükler bularak sözcük çeşitliliği sağlayamadığından sözcük yoğunluğu zorunlu olarak düşmektedir (Stubbs, 2014, s. 54).

Tablo 2: Anlatıcıların sözcük yoğunluğu

| Anlatıcı       | Sözcük Sayısı | Tek Sözcükler | Sözcük Yoğunluğu |
|----------------|---------------|---------------|------------------|
| Sokaktaki Adam | 1.895         | 1.233         | 65,06            |
| Hasan          | 1.895         | 1.143         | 60,31            |
| Ahmet          | 1.895         | 1.145         | 59,31            |
| Meryem         | 1.895         | 1.102         | 58,15            |
| Yakub          | 1.895         | 1.086         | 57,3             |
| Leon           | 1.895         | 1.044         | 55,09            |
| Ayhan          | 1.895         | 980           | 51,71            |

Tablo 2’de, anlatıcıların 1,895 sözcükten ne kadarını tek kullandığı görülmektedir. Tek sözcük (*type*) oranının yüksekliği metin içeriğinin sözcüksel (*lexical*) zenginliğini ifade eder. Tabloya göre Sokaktaki Adam 1,241 tek sözcük ve % 65,04 sözcük yoğunluğu oranıyla sözcüksel zenginliği açısından ilk sıradadır.

Tablo 3: Anlatıcıların toplam sözcük sayısı ve sözcük yoğunluğu oranı

| Anlatıcı       | Sözcük Sayısı | Tek Sözcükler | Sözcük Yoğunluğu |
|----------------|---------------|---------------|------------------|
| Roman (Toplam) | 44.009        | 13.112        | 29,77            |
| Hasan          | 19.995        | 7.129         | 35,65            |
| Yakub          | 10.903        | 4.329         | 39,71            |
| Sokaktaki Adam | 3.563         | 2.037         | 57,17            |
| Ayhan          | 3.018         | 1.348         | 44,66            |
| Meryem         | 2.697         | 1.451         | 53,81            |
| Ahmet          | 1.937         | 1.155         | 59,62            |
| Leon           | 1.896         | 1.044         | 55,06            |

Tablo 3’te romanın tamamı göz önüne alındığında, en fazla anlatıma sahip olan Hasan ve Yakub anlatıcılarının sözcük yoğunluğu oranındaki düşme ve romanın tümünün sözcük yoğunluğunun % 29,97 oluşu bunu ortaya koymaktadır.

### Sözcük Uzunluğu

Harf sayısı bakımından uzun sözcüklerin çokluğu metnin okuma zorluğunu ifade eder. Kullanılan uzun sözcükler bakımından en fazla sözcük 820 ile Sokaktaki Adam’a aittir. Yakub ve Meryem ise en fazla kısa sözcük ve en az uzun sözcük kullanarak konuşma diline en yakın, en sade anlatımı yapmışlardır.

Tablo 4: Anlatıcıları sözcük uzunlukları

| Anlatıcı       | Sözcük sayısı | Sözcük Uzunluk (Ort.) | Kısa Sözcükler (3 Harf ve altı) | Uzun Sözcükler (7 harf ve üzeri) |
|----------------|---------------|-----------------------|---------------------------------|----------------------------------|
| Sokaktaki Adam | 1.895         | 6,6                   | 369                             | 820                              |
| Hasan          | 1.895         | 6,3                   | 363                             | 760                              |
| Leon           | 1.895         | 6,2                   | 391                             | 696                              |
| Ayhan          | 1.895         | 6                     | 422                             | 660                              |
| Ahmet          | 1.895         | 5,8                   | 450                             | 635                              |
| Yakub          | 1.895         | 5,8                   | 465                             | 552                              |
| Meryem         | 1.895         | 5,6                   | 475                             | 582                              |

Tablo 4’te anlatıcıların kullandıkları sözcüklerin harf sayısı ortalaması, uzun ve kısa sözcüklerin metindeki sayıları görülmektedir. 7 harf ve üzeri sözcükler uzun sözcük olarak kabul edilmiştir. Anlatıcıların kullandıkları en uzun sözcükler: Yakub *okumayacakmış* (14 harf), Hasan *münasebetsizlikler* (18 harf), Ahmet *boşboğazlığının* (15 harf), Meryem *çekiştirilerken* (16 harf), Leon *anlattırıyorum* (16 harf) ve Ayhan *konuştuklarımızı* (16 harf) şeklindedir.

### Sözcük Bulutu

Sözcük bulutu (*word cloud*) metinden elde edilen sıklık listesinin görselleştirme şeklidir. Sözcük bulutları en sık sözcük en büyük olmak üzere kademeli olarak küçülerek belli sayıdaki sözcüğü görselleştirmek için kullanılır.



Şekil 3: Romanın tamamı için alınan sözcük bulutu

Şekil 3'te bütün romandaki en sık 45 sözcükten oluşan sözcük bulutu görülmektedir. Romanda *bir* (1458), *ve* (612), *ben* (499), *gibi* (407) ve *Hasan* (227) kez geçmiştir.



Şekil 4: Yakub için sözcük bulutu



Şekil 5: Hasan için sözcük bulutu



Şekil 6: Sokaktaki Adam için sözcük bulutu



Şekil 7: Ahmet için sözcük bulutu



Şekil 8: Meryem için sözcük bulutu



Şekil 9: Leon için sözcük bulutu



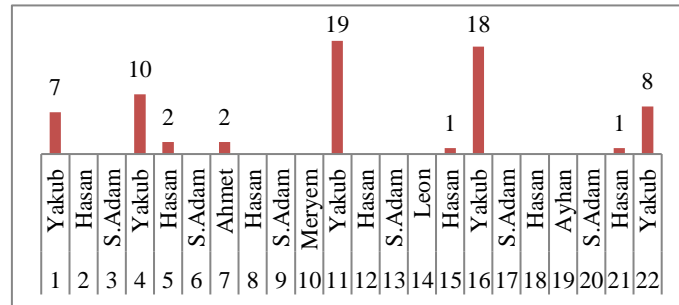
Şekil 10: Ayhan için sözcük bulutu

Romadaki her bir anlatıcı için oluşturulmuş olan yukarıdaki sözcük bulutları sık kullanılan işlevsel sözcükler çıkarıldıktan sonra yapılmıştır. Böylece sözcük bulutları incelendiğinde, her bir anlatıcının aynı olay örgüsü içerisinde kendine göre farklı şeyler anlattığı gözlemlenebilmektedir: Yakub, futbol ve Hasan'dan; Hasan, Meryem'den; Sokaktaki Adam kendinden ve zamandan; Ahmet, Meryem ve paradan; Leon, polisten ve yakalanma korkusundan, Ayhan da Hasan'dan söz etmektedir.

Sözcük bulutları yardımıyla her anlatıcıda birey dilinin yansımalarını görmek mümkündür: Yakub *ulan, ağbiy, be, bak, karı*; Hasan *daima, fakat*; Sokaktaki Adam *bazan, fakat*; Ahmet *heh heh, halbukisem*; Meryem *şekerim, herif, ayol*; Leon *be, efendim* çok kullanarak diğerlerinden ayrılmıştır. Hasan ve Sokaktaki Adam konu ve sözcük seçimi açısından benzeşmekte, konu açısından diğerlerinden daha soyut görünmektedir.

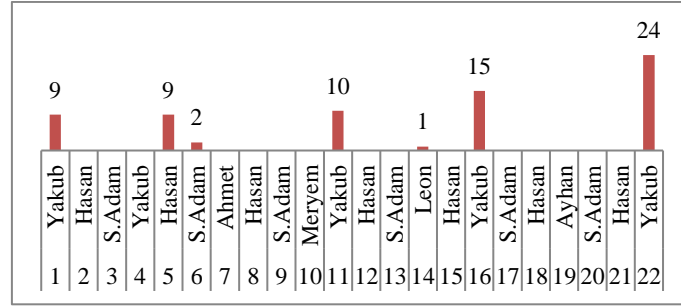
### Sözcük Dağılım Grafikleri

Metin içerisinde sözcüğün dağılımını görselleştirmek için sözcük dağılım grafikleri (*dispersion plots*) kullanılmaktadır. Çalışmamızın konusu olan romanda her anlatıcı ayrı bölüme anlatıda bulunduğundan hangi anlatıcının sözcüğü kullandığı ya da kullanmadığı dağılım grafikleri yardımıyla görülmektedir. Sözcük bulutlarıyla yapılan tespitleri, tüm romanın sıklık listesinden belirlenen bazı anahtar sözcüklerin roman içerisindeki dağılımını gösteren aşağıdaki dağılım grafikleri desteklemektedir.



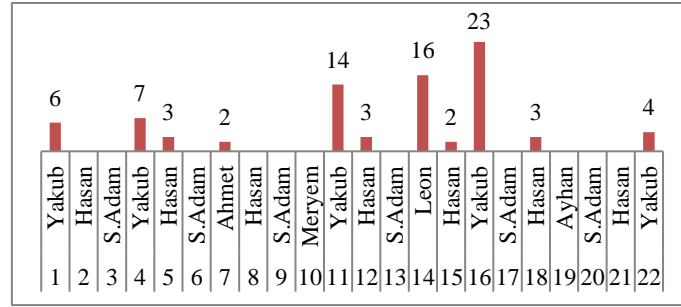
Şekil 11: ulan sözcüğünün romandaki dağılımı



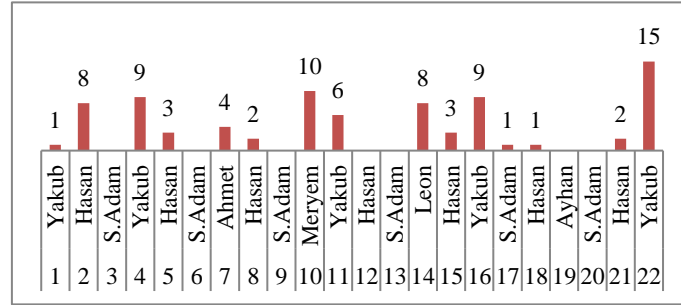


Şekil 12: ağbiy sözcüğünün romandaki dağılımı

Şekil 11’de görüldüğü üzere 68 kez geçen *ulan* sözcüğünün 62’sini ve Şekil 12’de 70 kez kullanıldığı görülen *ağbiy* sözcüğünün 67’sini Yakub kullanmıştır.

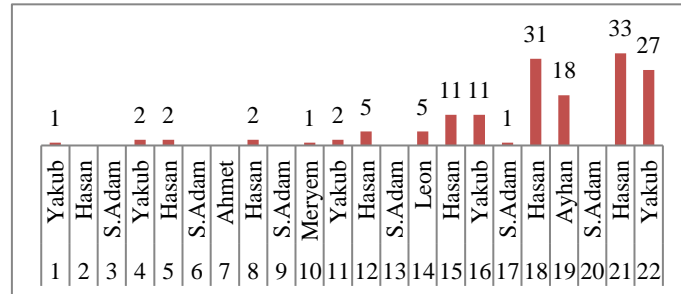


Şekil 13: be sözcüğünün romandaki dağılımı

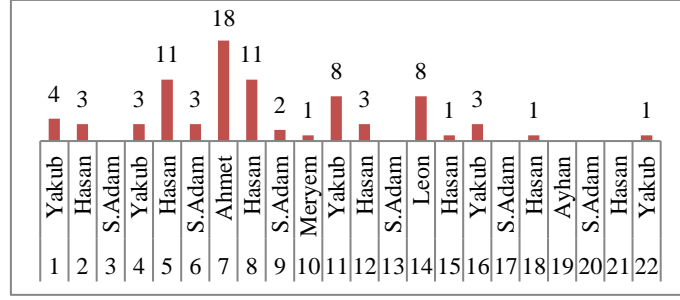


Şekil 14: herif sözcüğünün romandaki dağılımı

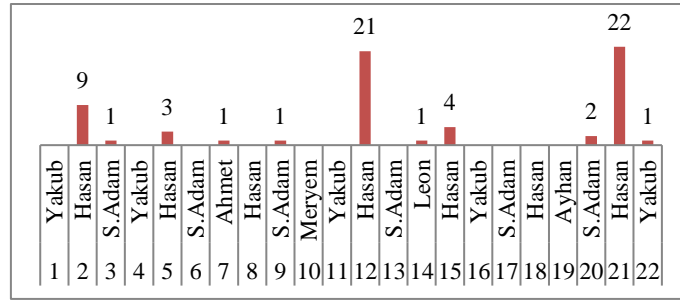
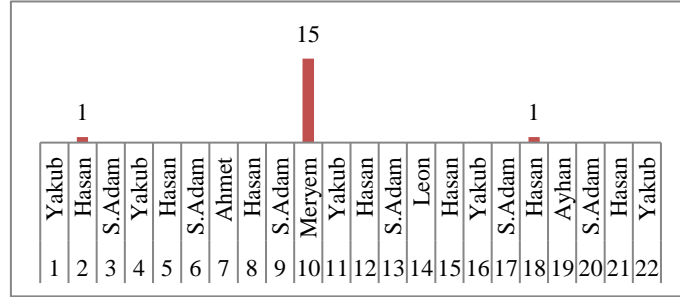
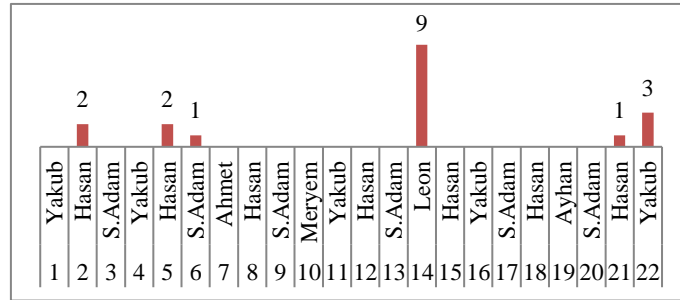
Toplam 87 kez geçen *be* (Şekil 13) sözcüğünü 54 kez Yakub ve 16 kez Leon kullanmıştır. *Herif* (Şekil 14) sözcüğünde ise ağırlık yine Yakub, Meryem ve Leon’dadır.



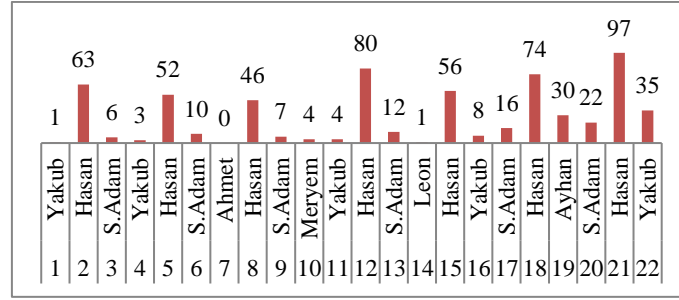
Şekil 15: diyorum sözcüğünün romandaki dağılımı

Şekil 16: *para* sözcüğünün romandaki dağılımı

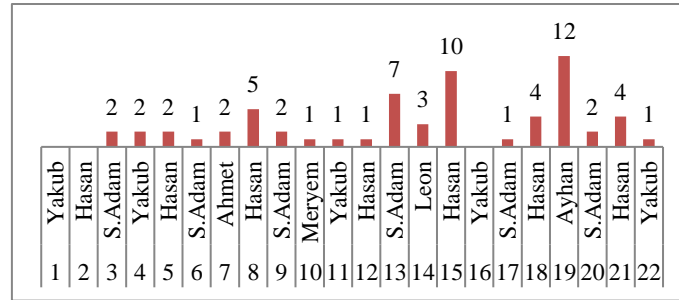
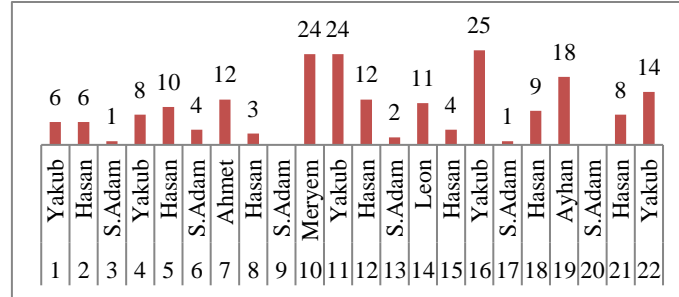
*Diyorum* (Şekil 15) romanın sonlarına doğru Hasan ve Yakub tarafından oldukça fazla kullanılırken *para* (Şekil 16) sonlara doğru geçmez olmuştur.

Şekil 17: *yağmur* sözcüğünün romandaki dağılımıŞekil 18: *şekerim* sözcüğünün romandaki dağılımıŞekil 19: *polis* sözcüğünün romandaki dağılımı

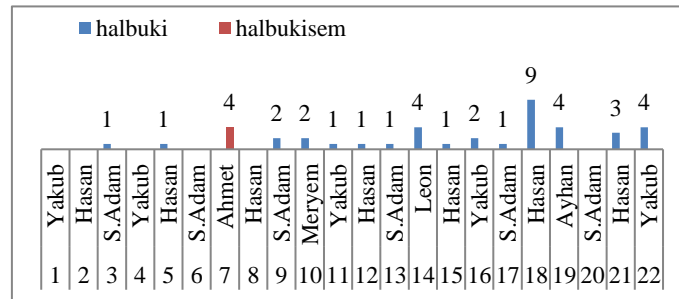
*Yağmur* (Şekil 17) sözcüğünün neredeyse tamamı Hasan, *şekerim* (Şekil 18) Meryem, *polis* (Şekil 19) ise Leon tarafından kullanılmıştır.

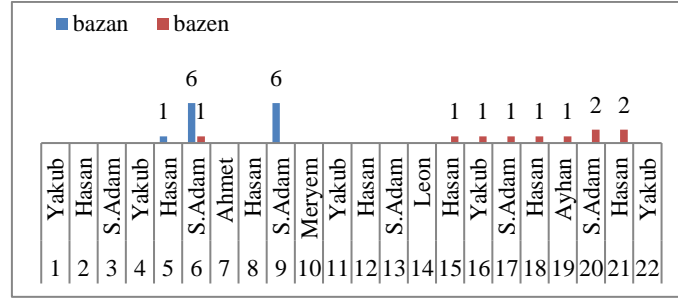
Şekil 20: *ve* bağlacının romandaki dağılımı

628 kez kullanılan *ve* bağlacı 470 kez Hasan ve 73 kez Sokaktaki Adam tarafından kullanılmıştır. Ahmet hiç *ve* bağlacını kullanmazken Ayhan da tek bölümde 30 kez kullanmıştır.

Şekil 21: *fakat* bağlacının romandaki dağılımıŞekil 22: *ama* bağlacının romandaki dağılımı

*Fakat* (Şekil 21) bağlacı tüm anlatıcılar tarafından kullanılmış olsa da ağırlık Hasan ve Sokaktaki Adam'dadır. *Ama* (Şekil 22) bağlacında Sokaktaki Adam ve Hasan çok az kullanmışken ağırlık Yakub ve Meryem'dedir.

Şekil 23: *halbuki* bağlacının romandaki dağılımı

Şekil 24: *bazan* /*bazen* sözcüklerinin romandaki dağılımı

*Halbuki* (Şekil 23) bağlacı romanın sonlarına doğru diğer anlatıcılar tarafından kullanılmışken, *halbukisem* sadece Ahmet tarafından kullanılmıştır. Diğer anlatıcılar az da olsa *bazen* (Şekil 24) kullanırken Sokaktaki Adam hep *bazan* kullanmıştır.

### Sözcük dizisi ( *n-gram* )

Metin içerisindeki sözcük dizilerinin sıklıklarını çıkarmak için sözcük dizisi (*n-gram*) taraması yapılır. *1-gram* (*monogram*), *2-gram* (*bigram*), *3-gram* ve daha fazla sözcükten oluşan diziler metin ya da derlem içerisinde çıkarılabilir. *Damlaya damlaya göl olur* tümcesinin ikili sözcük dizisi (*2-gram*) sonucu şu şekildedir: *damlaya damlaya*, *damlaya göl*, *göl olur*. Sözcük bazında yapılan *n-gram* incelemelerinde metin içerisinde sık kullanılan ikilemeleri, kalıplaşmış ifadeleri (*collocation*), özellikle yazar tarafından sık olarak tekrar edilmiş sözcük dizilerini tespit etmek mümkündür.

Tablo 5'te tüm anlatıcılar için alınmış sözcük dizisi listesinde anlatıcıların farklı konulardan, farklı ifadelerle söz ettikleri görülmektedir. *Ya da*, *ben de*, *o da* gibi birkaçı dışında birbirlerinden tamamen farklıdır ve ne ortak bir ikileme, ne deyim ne de kalıp bir ifade vardır.

Tablo 5: Anlatıcıların ilk 10 sözcük dizileri

| Yakub             | Hasan                  | Sokaktaki Adam | Ahmet                 | Meryem            | Leon              | Ayhan           |
|-------------------|------------------------|----------------|-----------------------|-------------------|-------------------|-----------------|
| yok ağbiy         | 4 Cumartesi için       | 6 her gün      | 5 heh heh             | 10 on beş         | 6 ben de          | 3 ben de        |
| gibi bir          | 4 bu gece              | 5 ben hep      | 3 o da                | 5 o da            | 6 bir de          | 3 ya da         |
| bir kere          | 3 ya da                | 5 sonra bütün  | 3 Şevket Nuri         | 4 ama ne          | 4 kara bıyıklı    | 3 biliyor musun |
| bu kadar          | 3 çarşambadır çarşamba | 3 aynı zamanda | 2 az kalsın           | 3 var mı          | 4 korkmaz mısımız | 3 bana ne       |
| dedin mi          | 3 bu yüzden            | 2 başka bir    | 2 bu işte             | 3 iki iki         | 3 ne var          | 3 onunla yalnız |
| değil mi          | 3 ellerimle yüzümü     | 2 bir iki      | 2 değil mi            | 3 daha dört       | 3 ondan sonra     | 3 hiçbir şey    |
| Seyfi Kaptan      | 3 şu bardakları        | 2 bir insan    | 2 diye sordu          | 3 küfürler ediyor | 2 bereket versin  | 2 bu kadar      |
| var mı            | 3 çukur aynasında      | 2 bir kahvede  | 2 Triyandafilos Ahmet | 3 hoşuma gidiyor  | 2 başka türlü     | 2 bak sana      |
| ben Beşiktaşlıyım | 2 ne de olsa           | 2 bir kahvenin | 2 ben de              | 2 ama bu          | 2 kan çanağına    | 2 ne diyeceğim  |
| sapına kadar      | 2 arasında ne          | 2 bir yere     | 2 bir de              | 2 başka çaresi    | 2 olacaksa olsun  | 2 az kalsın     |

### Tümce Uzunluğu

Sözdizimi açısından baktığımızda romanın tümü değerlendirildiğinde tümce başına düşen sözcük sayısı ortalama 5.43'tür. Anlatıcıların ise tümce başına düşen sözcük sayısı Tablo 6'da görüldüğü şekilde 5-7 civarındadır. Bütün anlatıcıların tümce uzunlukları listelendiğinde 10 sözcükten yukarı tümce adedi toplam adedin %15-20'si civarında oluşu romanın büyük oranda kısa, birleşik olmayan ve basit tümcelerden oluştuğunu göstermektedir.

Tümce oranlarının birbirine çok yakın oluşu, birleşik tümce oranlarının da her anlatıcıda aynı oranlarda oluşu yazarın anlatıcılar arasında roman boyunca aynı sözdizimsel yapıyı koruduğunu göstermektedir. Bu durum, anlatımın kurmaca kişilere bırakılıp, konuşma diliyle yapılmasından kaynaklanıyor görünmektedir.

Tablo 6: Toplam tümce sayısı ve tümce başına düşen sözcük sayısı

| Anlatıcı       | Sözcük Sayısı | Toplam Tümce Sayısı | Tümce Başına Düşen Sözcük Sayısı |
|----------------|---------------|---------------------|----------------------------------|
| Sokaktaki Adam | 1.895         | 271                 | 6,99                             |
| Ahmet          | 1.895         | 286                 | 6,63                             |
| Yakub          | 1.895         | 309                 | 6,13                             |
| Ayhan          | 1.895         | 319                 | 5,94                             |
| Leon           | 1.895         | 320                 | 5,92                             |
| Meryem         | 1.895         | 322                 | 5,89                             |
| Hasan          | 1.895         | 330                 | 5,74                             |

*Sokaktaki Adam* romanı üzerine dil incelemesi yapan Bahar Eriş'in tespiti de elde edilen bu sonucu desteklemektedir: *Sokaktaki Adam*, 1.507 basit isim tümcesinden, 4.822 basit fiil tümcesinden, buna göre de toplamda 6.329 basit tümceden ve 1.189 birleşik tümceden oluşmaktadır. *Sokaktaki Adam*'da 573 eksilteli olmak üzere toplam tümce sayısı 8.091'dir. %84 oranında basit, %16 oranında birleşik tümceden oluşur (Eriş, 2010, s. 17).

### Sonuç

*Sokaktaki Adam* romanının anlatıcı dünyası üzerine yapılan bilgisayar destekli sıklık ve sözcük dağılımı incelemesi sonucunda, yazarın anlatıcıların bireysel dillerini gözettiği gözlemlenmiştir. Elde edilen sonuçlara göre diğer anlatıcılardan sözcük kadrosu ve üslup olarak ayrılan Sokaktaki Adam adlı anlatıcı kurmaca içerisinde belli bir figür olmadığından yazarın kendisi olarak değerlendirilmiştir.

İncelenen sözcük sayısı 1895 ile sınırlandırıldığı için anlatıcıların sözcük yoğunlukları birbirine yakın olsa da Sokaktaki Adam ve Hasan en yüksek değere sahiptir. Üniversiteden terk bir aydını temsil eden başkişi Hasan ikinci sıradadır. Eğitim durumu en yüksek Ayhan olmasına rağmen en düşük değer Ayhan'a aittir. Şekil 10'daki sözcük bulutunda görüldüğü üzere bu durum Ayhan'ın kendi monologunda Hasan ile arasındaki ilişki odaklı konuşmasına bağlanabilir. Yakub ve Hasan aynı gemide çalışan iki kamarottur ve arkadaşır ama sözcük tercihlerinde büyük farklar vardır. Hasan düzeyli bir dil kullanırken Yakub argolu sokak dili kullanmaktadır.

Sözcük sıklık tablosunda işlevsel sözcükler açısından tüm anlatıcıların tercihleri çoğunlukla örtüştüğü görülmektedir ancak Hasan'ın *ve*, Yakub'un *ulan* ve *ağbiy*, Meryem'in *şekerim* tercihleri sıklık tablosunun öne çıkan sonuçlarıdır. Sözcük uzunluğu açısından elde edilen sonuçlar birbirine yakındır ama sözcük uzunluğunda Sokaktaki Adam ve Hasan öndedir.

Sözcük bulutları yardımıyla her anlatıcının kendi monologunda ne hakkında konuştuğunu, en sık hangi sözcükleri tercih ettiğini ve varsa kendine özgü hitap tarzını görmekteyiz. Yakub, Meryem ve Ahmet'in sokak dili tercih etmelerine karşın Sokaktaki Adam ve Hasan daha düzeyli ve daha soyut bir dil kullanmaktadır. Dağılım grafikleri de romandan seçilen anahtar sözcüklerin hangi anlatıcı kişilerce tercih edildiğini ortaya koymaktadır. Sözcük dizileri tablosu anlatıcılar arasında ortak sözcük dizisi olmadığını ortaya koymaktadır.

Sözdizimsel olarak romanın dengeli olduğu söylenebilir. Uzun ve bileşik tümceler varsa da çoğunluk kısa ve basit tümcedir. Anlatıcıların monologlarında konuşma dilini kullanmaları sözdizimsel yapıya sadelik kazandırmıştır.

### Kaynaklar

- Çetin, N. (2017). *Roman çözümleme yöntemi*. Ankara: Akçağ Yayınları.
- Desagulier, G. (2017). *Corpus linguistics and statistics with r*. London: Springer.
- Eriş, B. (2005). *Attilâ İlhan'ın dili üzerine bir inceleme: Sokaktaki Adam ve Gazi Paşa*. Yüksek Lisans Tezi. İzmir: Ege Üniversitesi Sosyal Bilimler Enstitüsü.
- İlhan, A. (2019). *Sokaktaki adam*. İstanbul: İş Bankası Kültür Yayınları.
- Jockers, M. L. (2014). *Text analysis with r for students of literature*. London: Springer.
- Kenny, A. (1982). *The computation of style*. Oxford: Pergamon Press.
- McMenamin, G. R. (2002). *Forensic linguistics: advances in forensic stylistics*. Florida: CRC Press.
- Sazyek, H. (2015). *Roman terimleri sözlüğü*. Ankara: Hece Yayınları.
- Stubbs, M. (2014). Quantitative methods in literary linguistics. *The Cambridge Handbook of Stylistics / ed.* Stockwell, P. & Whiteley, S. Cambridge: Cambridge University Press.
- Şen, C. (2009). Anlatım tekniği açısından Attilâ İlhan'ın Sokaktaki Adam romanı. *Türk Dili*, 690.
- Tekin, M. (2001). *Roman sanatı*. İstanbul: Ötüken Neşriyat.

### Extended Abstract

Computer-assisted quantitative text analyses enable us to reach the data and results that cannot be obtained through reading. With computers, the frequency and dispersion of lexical and linguistic units in a text are measured, various comparisons and statistical analyses can be made. Computational text analyses are especially useful in the study of long texts, corpora, a particular genre or a period.

To see what results can be achieved with the quantitative text analysis method, a computer-aided study was carried out on Attilâ İlhan's novel, *Sokaktaki Adam*. It was first published in 1953 and it comprises 222 pages and 22 chapters. The novel has seven narrators and each chapter is narrated alternately by a different narrator coming from a different social background. The aim of this study is to see the idiolectic characteristics of the narrators in terms of lexicon and sentence structure. The results are presented with visual tools such as charts, tables and word clouds.

The plot of the novel is: Hasan, the protagonist, works on ships as a steward and brings contraband furs from abroad and sells them in İstanbul with his friend Yakub. When the ship anchors in İstanbul, Hasan wants to sell the furs but Nubar, the buyer, has been arrested. Hasan delivers the furs Ahmet at the house of Meryem, the prostitute. In the meantime, Hasan has a close relationship with Meryem and meets Ayhan, his ex-girlfriend. Ahmet is arrested and Leon, the other buyer, escapes to Israel. While he was trying to separate a street fight, Hasan gets wounded and dies.

Nearly the half of the narration is done by Hasan, the protagonist, a quarter of it by Yakub and the rest of the narration by The Man on the Street, Ahmet, Meryem, Leon and Ayhan. The narration technique is called multiple narration in which more than one figure is in the narrator position in the novel. Therefore, it is necessary to examine an idiolect for each of these narrators who have different social and educational backgrounds.

First, the lengths of narration varies from Leon's 1,895 words to Hasan's 19,995. To give each narrator the same weight, each monologue has been reduced to the length of Leon's. Then, each part has been tokenized and created a sorted frequency list of words. The top 20 words of the each narrator's frequency list are shown in table 1 and it is clearly seen that the function words are at the top of the list.

Later, the lexical density of each narrator has been calculated. The lexical density or the type/token ratio is calculated by dividing the total number of unique word types by the total number of word tokens. High ratio of types indicates the lexical richness of the text. The Man on the Street is in the first place in terms of his lexical richness with the number of 1,241 single words (types) and 65,04% lexical density. When the text or corpus gets bigger, the type/token ratio decreases. The lexical density of the whole novel is 29,77 %.

Table 4 shows the average word length of each narrator. Words longer than 7 letters are regarded as long words. Long words deeply affect the readability of a text.

A word cloud is used to visualize a certain number of words from the text by gradual decrease in size and rank. The word clouds show that each narrator describes different things in his/her own perspective: Yakub talks a lot about football and Hasan; Hasan talks about Meryem very much; The Man on the Street, himself and time; Ahmet, Meryem and money; Leon, police, and Ayhan talks about Hasan.

It is also possible to see the difference among the lexical choices of each narrator. Yakub uses *ağbiy*, *bak*, *ulan*, *karı*; Hasan *daima*, *fakat*; The Man on the Street *bazan*, *fakat*; Ahmet *heh heh*, *halbukisem*; Meryem *şekerim*, *herif*, *ayol*; Leon *be*, *efendim*. Hasan and The Man on the Street are similar in terms of subject and word choice, and they appear more abstract than the others.

The results from word clouds are supported by the dispersion plots of the keywords selected in the frequency list of the whole novel. For instance, Yakub uses almost all *ulan* and the word *ağbiy*. The conjunction *ve* is used 628 times in total and 470 of them by Hasan and 73 by The Man on the Street. Meryem uses *şekerim* and *ayol*. *Halbukisem* is used only by Ahmet whereas *halbuki* is used by the other narrators towards the end of the novel.

In addition to frequency list and word clouds, the co-occurrences of words in the text or corpora can be shown in an n-gram list. N-gram lists are used to find out collocations in the text or corpus.

Regarding syntax, the average number of words per sentence is 5.43. The number of words per sentence of the narrators is around 5-7. This low ratio indicates that the novel is composed of short and simple sentences. The sentence ratios of narrators are very close to each, that is, the author keeps the same syntactic structure throughout the novel without making any syntactic distinction among the narrators.

In conclusion, the narrators show their idiolectic characteristics regarding lexicon. The Man on the Street, not considered as a figure in the plot, has the highest lexical density and Hasan, the protagonist, follows him. The Man on the Street and Hasan show identical characteristics and they use more courteous and abstract vocabulary than the others. Though Hasan and Yakub are friends and both stewards, they differ in lexical choices.