

**KOMPOZİSYON TİPİ SINAVLARDA KOMPOZİSYONUN
BİÇİMSEL ÖZELLİKLERİNDEN KESTİRİLEN
PUANLARIN ANAHTARLA VE GENEL İZLENİMLE
PUANLANMASINDAN ELDE EDİLEN PUANLARLA
İLİŞKİSİ**

113375

Bayram ÇETİN

Hacettepe Üniversitesi
Sosyal Bilimler Enstitüsü

113375

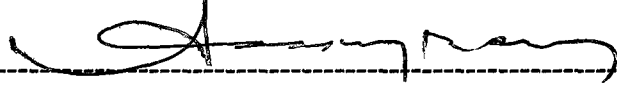
Lisansüstü Eğitim, Öğretim ve Sınav Yönetmeliğinin Eğitim Bilimleri Anabilim Dalı
Eğitimde Ölçme ve Değerlendirme Bilim Dalı için öngördüğü
YÜKSEK LİSANS TEZİ
olarak hazırlanmıştır.


DOKÜMANTASYON MERKEZİ

Ankara
Haziran, 2002

Sosyal Bilimler Enstitüsü Müdürlüğü'ne


Bu Çalışma, jürimiz Tarafından Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı'nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Başkan 
Doç. Dr. A. Ata TEZBAŞARAN

Üye 
Doç. Dr. Aysun UMay

Üye 
Doç. Dr. Selahattin GELBAL

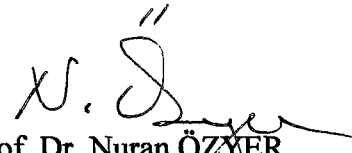
Üye 
Yrd. Doç. Dr. Hülya KELECİOĞLU (Danışman)

Üye 
Yrd. Doç. Dr. Ömer KUTLU

Onay

Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylarım.

22.1.2002


Prof. Dr. Nuran ÖZYER
Enstitü Müdürü

Bu günlere gelmemde büyük katkıları olan anneme ve babama...



ÖZET

Bu arařtırmada özel bir yazılı yoklama řekli olan kompozisyon sınavlarında kompozisyonun biçimsel özelliklerinin, anahtarla puanlamayı ve genel izlenimle puanlamayı ne derece yordadığının, en iyi yordayıcıların hangi biçimsel özellikler olduğunun ve anahtarla puanlama ile genel izlenimle puanlamanın puanlama güvenilirliklerinin belirlenmesi amaçlanmıştır.

Bu amaçlar doğrultusunda, 2001-2002 öğretim yılında Hacettepe Üniversitesi Eğitim Fakültesi İlköğretim Bölümü Sınıf Öğretmenliği Anabilim Dalı 1. sınıfta öğrenim gören 143 lisans öğrencisine kompozisyon sınavı uygulanmıştır. Öğrencilerin yazmış oldukları kompozisyonlar iki tane Türk Dili ve Edebiyatı öğretmenine hem anahtarla hem de genel izlenimle puanlatılmıştır.

Öğrencilerin yazmış oldukları kompozisyonlarda içerik analizi yapılarak daha önceden tanımlanan biçimsel özellikler ve anahtar kelimeler belirlenmiştir. Bu belirleme işlemi, anahtar kelimeler sayılarak; biçimsel özelliklerin de bir kısmı sayılarak bir kısmı da var olup olmadığına bakılarak yapılmıştır. Bu analiz sonucunda kompozisyonlar için 29 farklı değişkene ait veri elde edilmiştir. I. puanlayıcının önce anahtarla daha sonra genel izlenimle verdiği puanlar bağımlı değişken olarak alınmış, biçimsel özellikler için belirlenen 29 bağımsız değişken ile çoklu regresyon analizi yapılmıştır. Analizler sonucunda biçimsel özelliklerin anahtarla ve genel izlenimle verilen puanları ne derece yordadığı ve en iyi yordayıcı olan biçimsel özelliklerin neler olduğu belirlenmiştir.

Anahtarla ve genel izlenimle elde edilen puanların puanlayıcı güvenilirliğini hesaplamak için her iki puanlayıcının önce anahtarla verdikleri puanlar arasında daha sonra da genel izlenimle verdikleri puanlar arasında Pearson Momentler Çarpımı Korelasyon Katsayısı hesaplanmıştır. Aynı puanlayıcının anahtarla ve genel izlenimle verdiği puanlar arasındaki tutarlılık hesaplanırken her iki puanlayıcının da anahtarla ve genel izlenimle verdikleri puanlar arasında Pearson Momentler Çarpımı Korelasyon Katsayısı hesaplanmıştır.

Sonuç olarak; kompozisyonun anahtarla puanlanmasından elde edilen puanlarla kompozisyonun belirlenen biçimsel özellikleri arasında yapılan regresyon analizi sonucunda elde edilen denklem istatistiksel olarak $\alpha=0,01$ düzeyinde önemli bulunmuştur. Bağımlı değişken olan anahtarla puanlama ile biçimsel özellikler arasında 0,72 çoklu korelasyon katsayısı ve biçimsel özelliklerin anahtarla puanlamayı açıklayabilme derecesi olarak 0,52 belirtme katsayısı hesaplanmıştır. Kompozisyonun genel izlenimle puanlanmasından elde edilen puanlarla biçimsel özellikler arasında yapılan regresyon analizi sonucunda elde edilen denklem de istatistiksel olarak $\alpha=0,01$ düzeyinde önemli bulunmuştur. Bağımlı değişken olan genel izlenimle puanlama ile biçimsel özellikler arasında 0,77 çoklu korelasyon katsayısı ve biçimsel özelliklerin genel izlenimle puanlamayı açıklayabilme derecesi olarak 0,60 belirtme katsayısı hesaplanmıştır. Anahtarla puanlamanın puanlayıcı güvenilirliği 0,45, genel izlenimle puanlamanın puanlayıcı güvenilirliği 0,53 bulunmuştur. I. puanlayıcının iki puanlama yöntemi arasındaki tutarlılığı 0,78, II. Puanlayıcının iki puanlama yöntemi arasındaki tutarlılığı 0,55 bulunmuştur.

Kompozisyon sınavlarında kompozisyonun biçimsel özelliklerinden anahtarla puanlamanın ve genel izlenimle puanlamanın kestirebileceği sonucuna varılmıştır. Anahtarla ve genel izlenimle puanlamaların puanlayıcı güvenilirlikleri arasında istatistiksel bir fark bulunmamıştır. Öğretmenlerin anahtarı kullanabilme becerisindeki yetersizliğinin anahtarla puanlamanın puanlayıcı güvenilirliğini düşürücü bir neden olabileceği sonucuna varılmıştır.

SUMMARY

The purpose of this research is to determine the extent to which structural features of essays and keywords predict scoring by key and scoring by overall impression in writing test; and which structural features are the best predictors; and the degree of inter-rater reliability of scoring by key and scoring by overall impression.

For that purposes, a writing test has been administered to 143 undergraduate freshman students enrolled at Hacettepe University Faculty of Education Department of Elementary Education program in Primary Education in 2001-2002 school year. The writings of the students has been scored by two Turkish Language and Literature teachers using both scoring key and overall impression scoring methods.

Predetermined structural features and keywords were identified by doing content analysis for the compositions written by the students. This identification process was conducted by counting keywords, and some of the structural features and checking the existance of some of these feature. As a result of this analysis, 29 different variables were identified for the compositions. Key scores and overall impression scores of the I. rater are taken as dependent variables, multiple regression analysis was made on 29 independent variables for structural features. As a result of the analysis, to what extent structural features can predict scores obtained through scoring key and overall impression and the best predictive structural features were determined.

In order to compute inter-rater reliability of the two methods, Pearson correlation coefficients was computed between two rater's key scores and overall impression scores. While computing the consistency between the key scores and overall impression scores of the same rater, Pearson correlation were computed between key scores and overall impression scores of the both raters.

As a result, the equation obtained through regression analysis made between key scores and determined structural features of the compositions was found significant statistically at $\alpha=0,01$ level. The multiple correlation coefficient between dependent variable key scoring and structural features was found 0,72 and the determination coefficient for the key scoring explanation level of structural features was calculated 0,52. The equation obtained through regression analysis made between overall impression scores and structural features was found significant statistically at 0,01 level too. The multiple correlation coefficient between overall impression scores and structural features was found to be 0,77 and the determination coefficient was calculated 0,60. Inter-rater reliability of key scoring was found to be 0,45, and overall impression scoring was 0,53. the consistency between two scoring methods for I. rater was 0,78 and for II. rater was 0,55.

It was concluded that in the examination of essays written by students, the key scoring and overall impression scoring would be predict from structural features. Statistically significant difference was not found between two scoring method's inter-rater reliability. The inadequacy of the raters skill among using the key would decrease the inter-rater reliability of key scoring was concluded.

İÇİNDEKİLER

	Sayfa
ÖZET.....	i
SUMMARY.....	iii
İÇİNDEKİLER.....	v
TABLolar.....	vii
ŞEKİLLER VE GRAFİKLER.....	viii
ÖNSÖZ.....	ix
BÖLÜM I	
GİRİŞ.....	1
Problem Durumu.....	1
Problem Cümlesi.....	14
Alt Problemler.....	14
Sayıtlar.....	15
Sınırlılıklar.....	15
Tanımlar.....	15
Araştırmanın Amacı ve Önemi	16
İlgili Araştırmalar.....	16
BÖLÜM II	
YÖNTEM.....	19
Araştırmanın Türü.....	19
Araştırmanın Yapıldığı Grup.....	19
Verilerin Toplanması ve Veri Toplama Araçları.....	20

Verilerin Analizi.....	24
BÖLÜM III	
BULGULAR VE YORUM.....	27
1. Alt Probleme İlişkin Bulgular ve Yorum.....	27
2. Alt Probleme İlişkin Bulgular ve Yorum.....	31
3. Alt Probleme İlişkin Bulgular ve Yorum.....	35
BÖLÜM IV	
SONUÇ VE ÖNERİLER.....	39
Sonuçlar.....	39
Öneriler.....	40
KAYNAKÇA.....	42

TABLOLAR

Sayfa No

TABLO 1. Cümle Yapısı Ve Anlatım Yapısı Yordayıcıları.....	21
TABLO 2. Anahtar Kelimeler.....	22
TABLO 3. Puanlamalara Ait Betimleyici İstatistikler.....	27
TABLO 4. Anahtarla Puanlama ile Kompozisyonun Biçimsel Özelliklerinin Regresyon Analizine İlişkin ANOVA Tablosu.....	28
TABLO 5. Kısmi Regresyon Katsayıları, Standart Hataları, t değerleri ve Manidarlık Düzeyleri.....	30
TABLO 6. Genel İzlenimle Puanlama ile Kompozisyonun Biçimsel Özelliklerinin Regresyon Analizine İlişkin ANOVA Tablosu.....	31
TABLO 7. Kısmi Regresyon Katsayıları, Standart Hataları, t değerleri ve Manidarlık Düzeyleri.....	34

ŞEKİLLER

Sayfa No

ŞEKİL 1. Anahtarla puanlama Gözlenen Değerler ve Standartlaştırılmış Kestirilen Değerler Saçılma Grafiği.....	29
ŞEKİL 2. Genel İzlenimle Puanlamanın Gözlenen Değerleri ile Standartlaştırılmış Kestirilen Değerlerine Ait Saçılma Grafiği.....	32
ŞEKİL 3. Anahtarla Puanlama'da İki Puanlayıcının Verdiği Puanlar Arasındaki İlişkiye Ait Grafik.....	35
ŞEKİL 4. Genel İzlenimle Puanlama'da İki Puanlayıcının Verdiği Puanlar Arasındaki İlişkiye Ait Grafik.....	36
ŞEKİL 5. Birinci Puanlayıcının Anahtarla ve Genel İzlenimle Verdiği Puanlara Ait İlişki Grafiği.....	38
ŞEKİL 6. İkinci Puanlayıcının Anahtarla ve Genel İzlenimle Verdiği Puanlara Ait İlişki Grafiği.....	38

ÖNSÖZ

Bu araştırma, kompozisyon sınavlarında anahtarla ve genel izlenimle puanlama yöntemlerinin kompozisyonun biçimsel özelliklerinden ne derece kestirilebileceğinin belirlenmesi ve bu yöntemlerin puanlayıcı güvenilirliklerini belirlemek amacıyla yapılmıştır.

Bu araştırmanın yürütülmesi sırasında bir çok kişinin yardım ve desteğini gördüm. Araştırmanın her aşamasında yaptığı yardımlar ve rehberliği konusunda danışmanım olan Yrd. Doç. Dr. Hülya Kelecioğlu'na, bu aşamaya gelmemde emeği olan hocalarım Doç. Dr. A. Ata Tezbaşaran'a ve Doç. Dr. Selahattin Gelbal'a, kompozisyon sınavı uygulamasındaki yardımlarından ve alan uzmanı olarak görüşlerinden faydalandığım Dr. Fatma Türkyılmaz'a, kompozisyonları puanlayan Türk Dili ve Edebiyatı öğretmenlerine ve evdeki çalışmalarımda bana destek olan eşim Menekşe'ye teşekkür ederim.

BÖLÜM I

GİRİŞ

Bu bölümde problem durumu, problem cümlesi, alt problemler, sayılılar, sınırlılıklar, ilgili arařtırmalar ile tanım ve kısaltmalara yer verilmiřtir.

Problem Durumu

Eđitimde bařarının ölçülmesi ve deđerlendirilmesi eđitim-öđretim sürecinin ayrılmaz bir parçasıdır. Bireyde istendik davranıř deđişikliklerinin meydana gelip gelmediđi ölçme ve deđerlendirme süreciyle belirlenebilir. “İstenilen niteliklerde bir davranıř deđişikliđi, uygulanan eđitimin bařarisına, beklenen davranıř deđişikliđinin gerçekleřmemesi de bařarısızlıđa kanıttır” (Turgut 1997).

Eđitim programlarında yenileřmelere gitmede ölçme ve deđerlendirme yöntemlerindeki yenilikler önemli derecede katkıları sađlayacaktır. Wiggins (Akt: Miller ve Linn 2000) eđitimsel reformların en iyi řekilde deđerlendirme programlarında deđişikliklerle bařarılabilirliđini, çünkü deđerlendirmenin öđretmenlerin ne öđreteceđini ve öđrencilerin ne öđreneceđini belirlemede önemli olduđunu belirtmektedir.

Deđerlendirme eđitim sisteminin kontrol görevini yapar. Baykul (2000) deđerlendirmenin, kendisi de dahil olmak üzere, eđitim sistemindeki öđelerin iyi iřleyip iřlemediđini, varsa iřlemeyen yönlerini ortaya koyacađını; sistemin onarılması için bunun gerektiđini belirtmektedir. Turgut (1997) deđerlendirmeyi ölçme sonuçlarının bir ölçütü olarak kıyaslanarak karara varılması iři olarak tanımlamıřtır. Burada ölçme, deđerlendirme için gerekli verilerin toplanması iřidir. Deđerlendirme kararının isabetlilik derecesi, ölçütün uygunluđu yanında ölçme sonuçlarının geçerlik ve güvenilirliđine bađlıdır (Baykul 2000).

Eđitimde kullanılan farklı ölçme araçları vardır. Bunlar yazılı yoklamalar, kısa cevaplı testler, dođru-yanlıř tipi testler, çoktan seçmeli testler, sözlü yoklamalar olarak

sınıflandırılabilir. Bu araçlardan hangisinin kullanılacağı yoklanmak istenen özelliğin doğasına göre veya pratik nedenlerden dolayı uygulayıcı tarafından belirlenir.

Yazılı yoklamalar, az sayıda soru sorularak, öğrencilerin cevaplarını düşünüp istedikleri gibi düzenleyerek yazmalarına olanak tanıyan bir sınav türüdür. Öğrencilerin cevaplarını kendilerinin yazması ve büyük ölçüde cevaplayıcı bağımsızlığının olması, puanlamanın da puanlayıcının kişisel kanılarından etkilenmesine yol açar.

Cevaplayıcının; bir kelime, bir rakam, bir ibare veya en çok bir cümle ile cevaplayabileceği maddelerden oluşan test tipine kısa cevaplı test denir (Turgut 1997). Kısa cevaplı testler, cevapları öğrencilerin düşünüp yazması açısından yazılı yoklamalara benzemektedir. Bunun yanında öğrencinin cevabı istediği gibi düzenlemesi yazılı yoklamalara göre daha sınırlı olduğundan puanlanması yazılı yoklamalardan daha objektiftir. Tekin (1993) kısa cevap maddesinin cevabı öğrenci tarafından hatırlanıp yazılan tek objektif madde olduğunu söylemektedir.

Bazıları doğru, bazıları yanlış önermeler halinde verilen maddelerden oluşan testlere doğru – yanlış testleri denir. Öğrencilerden bu önermelerin doğru veya yanlış olduğuna ilişkin bir karar vermesi beklenir. Şans başarısı ile doğru cevabın bulunma olasılığı yüzde ellidir. Puanlanması objektif olarak yapılır.

Çoktan seçmeli test maddeleri ise bir madde kökü ve seçeneklerden oluşan maddelerdir. Madde köklerine veya seçeneklerine göre farklı yapılandırılmış çeşitleri vardır. Şans başarısı, doğru – yanlış tipi maddelere oranla daha azdır. Objektif olarak puanlanabilir.

Sözlü yoklamalar, soruların genellikle sözlü olarak sorulduğu ve öğrencilerden sözlü olarak cevaplama beklenen sınav türüdür. Puanlanması tamamen kişisel kanı gerektirdiğinden oldukça öznelidir.

Yukarıda belirtilen sınav türlerinden yazılı yoklamalar öğretmenler tarafından sıkça kullanılan bir sınav türüdür. Çünkü bazı bilişsel özellikler objektif testlerle

yoklanmaya uygun değildir. Kompozisyon sınavları, belirlenmiş bir doğru cevap bulunmaması yönüyle yazılı yoklamaların diğer türlerinden ayrılır. Kompozisyon sınavları öğrencilerin dili kullanması, bilgileri ve düşünceleri organize etmesi; etkili bir anlatımla ifade etmesi gibi özellikleri ölçmeye yöneliktir. Bu nedenle kompozisyon sınavında ölçülen özelliklerin çoktan seçmeli testlerle yoklanması geçerlik açısından sakıncalar içerebilir. Çünkü kompozisyon öğrencinin herhangi bir düşünceyi planlamasını, yapılandırmasını ve kendi kelimeleriyle yazmasını gerektirir. Yazılı yoklamalar objektif sınav türlerinden bazı önemli farklılıklar gösterirler. Bu farklara göre amacımıza ulaştıracak sınav tipini belirleyebiliriz. Ebel (1965) bu farkları aşağıdaki gibi sınıflamıştır:

1. Bir yazılı yoklama sorusu; öğrencilerin cevaplarını düşünmelerini, planlamalarını ve kendi kelimeleriyle ifade etmelerini gerektirir. Objektif bir test sorusu öğrencinin cevabını birden çok düzenlenmiş seçenek arasından seçmesini gerektirir.
2. Yazılı yoklamalar objektif testlere göre daha az sayıda, gereksiz ve istenilenin dışında cevap vermeye elverişli ve daha genel sorular içerir. Objektif test ise çok daha fazla ve kısa cevaplar gerektiren sınırlanmış sorular içerir.
3. Öğrenciler yazılı yoklamalarda zamanlarının çoğunu düşünme ve yazmaya ayırırlar. Objektif sınavlarda zamanın çoğu okuma ve düşünmeye ayrılır.
4. Objektif bir testin niteliğini daha çok test geliştiricinin becerisi belirler. Yazılı sınavın niteliğini ise daha çok öğrencilerin cevaplarını puanlayan puanlayıcı belirler.
5. Yazılı yoklama tipi bir sınav, hazırlama bakımından daha kolaydır ama hem doğru olarak puanlanması güçtür hem de puanlama zaman alıcıdır.

İyi bir objektif test hazırlamak daha zor ve yorucudur ama doğru olarak puanlama yapmak da oldukça kolaydır.

6. Bir yazılı yoklama öğrenciye verdiği cevaba bireysel özelliklerini katması açısından daha fazla serbestlik verir, puanlayıcıya da puanlamada bireysel görüşünün rehberliğinde puanlama yapması açısından daha az sınırlar. Objektif bir test ise test geliştirmeciye kendi bilgi ve değerlerini katması yönünde serbestlik tanırken, öğrenciye ise sadece verdiği doğru cevapların oranı yoluyla ne kadar çok veya ne kadar az bildiğini gösterme olanağı verir.
7. Objektif test maddelerinde yazılı tipi sınavlara göre çok daha kesin bir şekilde kaç puan tayin edileceğine karar verilebilir.
8. Objektif testler öğrencileri doğru cevabı tahmin etmeye izin verir, bazen de cesaretlendirirken, yazılı yoklamalarda sorulan sorular öğrencilerin konudan saparak bildiği konulara yönelmesine izin verir ve bazen de cesaretlendirir.

Gerek objektif testler gerekse yazılı yoklamalar kullanılış amaçlarına ve olanaklarına göre bazı avantaj ve sınırlılıklara sahiptir. Yukarıda sıralanan farklar bu avantaj ve sınırlılıkları belirler. Sınav uygulamasında hangisinin kullanılacağına karar vermede ölçülecek davranışın niteliği, düzeyi, uygulanacak grubun özellikleri, konu alanı özellikleri göz önünde bulundurulmalıdır.

Yazılı yoklama tipi sınav sorularının en belirgin üstünlüğü objektif test maddeleriyle yoklanamayan daha karmaşık öğrenme ürünlerinin yoklanabilmesidir. Bu öğrenme ürünleri; düşünceler arasında tercihte bulunma, organize etme, bütünleştirme ve düşünceleri değerlendirme becerileri gibi üst düzey bilişsel becerilerdir. Ayrıca bu becerilerin ifadelerinde de cevaplayıcıya kendi kelimeleriyle ifade etme serbestliği sunmaktadır. Düşünceleri bütünleştirme, düşünme becerileri uygulamaları ve problem

çözme becerileri uygulamalarını ölçmede daha elverişli olması yazılı yoklamaların ikinci bir üstünlüğüdür (Gronlund 1967). Üçüncü üstünlük olarak, yazılı yoklamaların sınıfta kullanımlarında tercih sebebi olarak görülebilecek en önemli özelliği, kolay hazırlanabilir olmasıdır. Yazılı yoklamaların önemli özelliklerinden biri soruları cevaplandırmada objektif test maddelerine göre şansla cevaplama olanağı çok daha azdır. Bu da geçerliği ve güvenilirliği artırıcı yönde olumlu bir etkiye sahiptir.

Yazılı yoklamalarda en önemli sınırlılık, puanlamadaki güvenilirliğin düşük olmasıdır. Bir çok ciddi çalışmada bu konu incelenmiş ve görülmüştür ki bu tip sorular farklı öğretmenler tarafından farklı şekilde puanlanmıştır. Hatta aynı öğretmenler tarafından farklı zamanlarda farklı puanlandığı da görülmüştür. Puanlamadaki bu değişkenlik aynı kağıdı en iyiye yakın olmaktan en kötüye yakın olmaya kadar sınıflayabilmektedir. Bu şekilde yapılan puanlamaları Gronlund (1967) “Üç kör adamın fili tanımlamasına” benzetmektedir. Bir öğretmen içeriğe daha fazla puan verirken, bir diğeri düşünceyi organize etme, başkası ise yazma becerisine daha fazla puan vermektedir.

Diğer bir sınırlılık ise puanlama için gereken yeterli süredir. Eğer puanlayıcı dikkatli bir şekilde puanlama yapar ve gerekli notları yeri ve zamanı geldiğinde almaya özen gösterirse az sayıdaki sınav kağıtlarını puanlamak için bile saatler harcamak gerekecektir (Gronlund 1967).

Yazılı yoklamalarda diğer bir sınırlılık ise sorulabilecek soru sayısının azlığıdır. Bu bazı alanların yoğun olarak ölçülmesini sağlarken bazı alanların ölçülmesinin ihmaline sebep olabilmektedir. Bu sebeple sınavın kapsam geçerliği düşebilmektedir.

Her sınavda olduğu gibi yazılı yoklamalar ve özel bir hali olan kompozisyon sınavlarının sorularını hazırlarken de bazı esaslara uyulması gerekmektedir. Bu sınavların hazırlanmasında uyulması gereken esaslar aşağıdaki gibi sayılabilir:

1. Yazılı yoklama tipi sorular, sadece objektif test maddeleriyle geçerli ve güvenilir olarak yoklanamayacağı düşünülen nitelikleri yoklamak için kullanılmalıdır (Wrightstone, Justman ve Robbins 1956). Örneğin yazılı anlatım becerisi, objektif testlerle yoklanmaya uygun değildir. Bu durumda kompozisyon sınavlarını kullanmak kaçınılmazdır.
2. Yazılı yoklama soruları hazırlanırken ölçülecek nitelikler iyi belirlenmelidir (Micheels ve Karnes 1968). Hedefler açıkça tanımlanmalıdır. Bunun için iyi bir sınav planı hazırlanmalıdır. Bu planda sınavın amacı, ölçülecek davranışlar, kapsamdaki konular, kritik davranışlar ve konuların soru sayısına göre dağılımı açıkça belirlenmelidir (Tekin 1993). Kompozisyon sınavlarında ileri sürülen fikrin doğruluğu ya da yanlışlığı puanlamaya katılmadığından bir kapsam analizi söz konusu değildir. Ancak bu sınavlarda yazılı anlatım becerilerinin tanımlanması gerekmektedir.
3. Sorulara verilebilecek cevaplar sınırlandırılmalıdır (Ebel 1965; Micheels ve Karnes 1968; Turgut 1997). Sorular her öğrencinin doğru olarak anlayabileceği şekilde açık yazılmalıdır (Ebel 1965 ; Tekin 1993; Turgut 1997). Açıkça anlaşılmayan sorular başka yönler çekilebilir. Sınırlanmayan sorular hem uzunluk hem de ayrıntılara girme bakımından farklılaşır. Ayrıca iyi sınırlanmayan sorular başka yöne çekilmeye elverişli hale gelir. Kompozisyon sınavlarında da, kompozisyon konusu yeterince sınırlandırılarak tanımlanmalıdır.
4. Puanlama yaparken, her sorunun cevabına kaç puan verileceğini gösteren analitik bir puanlama anahtarı hazırlanmalıdır (Ebel 1965; Micheels ve Karnes 1968; Wrightstone, Justman ve Robbins 1956). Kompozisyon sınavları için de kompozisyonda nelerin ölçüleceğini ve bu özelliklerin kaç

puan olduğunu gösteren analitik bir puanlama anahtarı hazırlamak bu sınavların güvenilirliğini ve geçerliğini artırıcı yönde etki yapar.

Yazılı yoklamaları puanlamada farklı yöntemler kullanılmaktadır. Turgut (1997) bu puanlama yöntemlerini genel izlenimle puanlama, sınıflama, sıralama ve anahtarla puanlama olarak sınıflamıştır. Genel izlenimle puanlamada puanlayıcı bir cevap kağıdını baştan sona okur ve edindiği izlenime göre bir puan tayin eder. Sınıflamada puanlayıcı okuduğu kağıdı “iyi – orta – fena” sınıflarından birine yerleştirir ve arkasından her sınıf içinde bir sınıflama daha yapılır. Oluşan dokuz alt sınıfın ardışık sınıflarındaki en alt ve en üstler karşılaştırılarak gerekli deęiřtirmeler yapılır ve dokuz alt sınıf şeklinde öğrencilerin seviyeleri belirlenir. Sıralamada ise bütün kağıtlar birbiriyle karşılaştırılarak bir sıraya konur. Birbirlerine bağımlı olarak iyiden kötüye doğru veya kötüden iyiye doğru bir sıralama yapılarak puanlanır. Anahtarla puanlamada ise iki yöntem göze çarpmaktadır. Bunlar nokta puan (point) metodu ve dereceleme (rating) metodu olarak isimlendirilebilmektedir. Nokta puan metodunda her cevap noktasına bir puan tayin edilerek yapılır. Dereceleme metodunda ise “çok iyi”, “iyi”, “orta”, “zayıf” ve “çok zayıf” gibi dereceleri gösteren puanlama anahtarları hazırlanarak yapılabilir.

Yazılı yoklamalarda puanların güvenilirlik ve geçerlięi objektif testlere göre daha düşüktür. Puanlamada objektiflięi sağlamak, geçerlik ve güvenilirlięi artırmak için bazı önlemler alabiliriz. Bu önlemleri Gronlund (1967) ařaęıdaki şekilde sıralamıştır:

- a) Puanlayıcı puanlamadan önce muhtemel cevaplara göre bir anahtar hazırlamalıdır. Nasıl bir anahtar kullanacağını belirlemelidir.
- b) Amaca en uygun puanlama metodu kullanılmalıdır.
- c) Ölçülmek istenen öğrenme ürünleriyle doğrudan ilişkili olmayan faktörleri önceden belirlenmelidir.

- d) Cevaplar öğrenciden öğrenciye değil de, sorudan soruya puanlanmalıdır. Tüm öğrencilerin aynı cevapları puanladıktan sonra bir diğer soruya geçilmelidir.
- e) Öğrencilerin adlarına bakılmadan puanlama yapılmalıdır. Böylece öğrencilerin tanınması ile ilgili yanlışlıklar puanlamaya karışmamış olur.
- f) Sonuçları etkileyecek önemli kararları verirken iki veya daha fazla bağımsız puanlayıcı kullanılmalıdır. Bu puanlayıcıların verdikleri puanların ortalamalarına göre karar verilebilir.

Eğitimin bir amacı da bireylerin duygularını, düşüncelerini, görüş ve hayallerini yazılı olarak anlatabilme becerisi kazandırmaktır. Bu amaçla öğrencilere yazılı anlatım becerilerini geliştirmeye yönelik çalışmalar yaptırılır. Bireylerin yazılı anlatım becerilerindeki gelişmeleri belirleyebilmek amacıyla ölçmeler yapılır.

Bireylerin yazılı anlatım becerilerini ölçmenin en uygun yolu kompozisyon sınavı uygulamaktır. Yazılı anlatım becerilerinin daha objektif testlerle yoklanması geçerlik açısından sakıncalar doğurur. Çünkü yazılı anlatımın doğrudan yoklanması ancak kompozisyon yazdırma uygulamasıyla yapılabilir. Çoktan seçmeli veya kısa cevaplı testlerle ancak dilbilgisi yeterlilikleri ölçülebilir. Oysa yazılı anlatımda ölçmek istenilen beceri yalnızca dilbilgisi yeterliliği değil, aynı zamanda düşünceyi etkili bir şekilde ifade etme yeteneğidir. Enginarlar (1991) Perkins'in yazım becerilerini dolaylı yollarla yani çoktan seçmeli testlerle ölçen TWSE (Test of Standart Written English) test sonuçları ile sözü edilen becerilerin kompozisyon sınavı sonuçları arasında 0,45 düzeyinde bir ilişki bulunduğunu belirtmektedir.

Yazılı anlatım becerilerini ölçmeye yönelik olarak yapılan kompozisyon sınavlarının güvenilirliği ve geçerliği önemli bir konudur. Yazılı anlatım becerilerinin nesnel öğeleri fazlaca içerdiğinden dolayı, güvenilir olarak puanlanması oldukça güçtür. Çünkü bu tür sınavlarda soru cevaplandırma değil öğrencilerin duygu, düşünce ve

hayallerini dilin kurallarına uygun bir biçimde etkili bir anlatım kullanarak organize etmeleri istenir. Bu nedenle bir doğru cevap bulunmaz ve doğru cevap anahtarı hazırlanmaz. Bunun yerine içerik, organizasyon, dilbilgisi, anlatım yapısı, sözcük dağarcığı, noktalama ve yazım kuralları gibi ana başlıklar altında uygunluk ve yaratıcılıkları kapsayan bir anahtar hazırlanır.

Öğretmenler sınıfta uyguladıkları kompozisyon sınavlarında genelde bir puanlama anahtarı hazırlamazlar. Bunun yerine kendi uzmanlıklarını temel alarak genel izlenimlerine göre bir puanlama yaparlar. Bu tür bir puanlama geçerlik ve güvenilirlik konusunda sakıncalar içerebilir. Çünkü bazı öğretmenler içeriğe, bazıları düşüncelerin organizasyonuna, bazıları ise dilbilgisi ve noktalamaya ağırlık verdiği için sınavın geçerliği düşebilir. Ayrıca farklı bölümlere verilen ağırlıklar değiştiğinden puanlayıcı güvenirliliği de düşecektir.

Enginarlar (1991) kompozisyon puanlamada temel olarak üç puanlama yöntemi olduğunu söylemekte, diğer yöntemlerin bu temel yöntemlerin farklı bir uygulaması olduğunu belirtmektedir. Bu üç yöntem hata sayma, bütünsel (holistic) anahtar kullanarak puanlama ve analitik anahtar ile puanlama yöntemleridir.

Hata sayma; içerik, örgütleme, dilbilgisi, sözcük dağarcığı, noktalama ve yazım kuralları ana başlıklarına göre yapılan hatalara bakılarak hatanın önemine göre puan kesme şeklinde yapılan bir puanlama metodudur. Enginarlar (1991)' dan alınan örnek bir anahtar aşağıya çıkarılmıştır:

ÖLÇÜT	PUAN	ÖNERİLER VE NOTLAR
İÇERİK	6 Puan	Hatanın önemine göre 1 puan kesilebilir.
ÖRGÜTLEME	6 Puan	Hatanın önemine göre 0.5 veya 1 puan kesilebilir.
DİLBİLGİSİ	6 Puan	Hatanın önemine göre 0.25 veya 0.5 puan kesilebilir.
SÖZCÜK DAĞARCIĞI	4 Puan	Hatanın önemine göre 0.25 veya 0.5 puan kesilebilir.
NOKTALAMA VE İMLA	3 Puan	Noktalama, büyük harf, yazım hatası ve paragraf düzeni için her hatada 0.25 puan kesilebilir.

Bütünsel (Holistic) puanlama metodunda da 5 ya da 6 sınıflama belirlenmekte, bunlar: “Çok İyi”, “İyi”, “Yeterli”, “Yeterlik Altı” ve “Zayıf” gibi sınıflamalardan

oluşmakta, bu sınıfların niteliklerini taşıyan kompozisyonların özelliklerini belirten bir anahtar hazırlanmaktadır. Kompozisyon hangi sınıfın özelliklerini taşırsa o sınıfın puanını almaktadır. Bütünsel puanlamaya örnek oluşturabilmek üzere ilgili kaynaklar incelenmiştir. Elde edilen kaynaklarda yer alan bilgiler sentezlenerek bütünsel puanlama anahtarı hazırlanmıştır. Bu puanlama anahtarında yer alabilecek ölçüt listesine ilişkin bir örnek aşağıda sunulmuştur.

Bütünsel (Holistic) Kompozisyon Puanlama Anahtarı

Çok İyi (5 Puan)

- Düşünceler açık ifade edilmiş, düşünceler bütünlük içerisinde ve ilgi çekici, detaylar düşünceleri açıklamaya yardım etmektedir.
- İyi bir giriş, gelişme ve sonuç organizasyonu var, fikirler akıcı olarak anlatılmış, fikirlerin açıklanmasında çok düzenli bir sıra izlenmiş.
- İlgi çekici bir çok kelime kullanılmış, farklı cümle yapılarına yer verilmiş.
- Kelimeler doğru olarak yazılmış, büyük harfler, noktalama işaretleri doğru ve yerinde kullanılmıştır, paragraf düzenlemeleri yerinde yapılmış, harfler ve boşluklar kurallarına uygundur.

İyi (4 Puan)

- Düşünceler açık ifade edilmiş, fikirler basit görünümlü, detaylar ana fikri desteklemektedir.
- Giriş, gelişme ve sonuç var, fikirlerin açıklanmasında bir sıra var.
- Biraz ilgi çekici kelimeler kullanılmış, cümleler basit yapılandırılmış, fakat doğru.
- Çoğu kelime doğru yazılmış, büyük harfler doğru kullanılmış fakat bazı noktalama hataları bulunabilir, paragraf düzenlemeleri genelde yerinde yapılmış, harfler ve boşluklar kurallarına uygun.

Yeterli (3 Puan)

- Ana fikir anlaşılır, az miktarda destekleyici detaylara inilmiş.
- Giriş veya gelişme bölümleri bulunmayabilir, fikirlerde sıralama orta düzeyde.
- Kelimeler doğru fakat ilgi çekici kelimeler kullanılmamış, basit cümleler kullanılmış.
- Kelimelerde yazım yanlışları fazlaca yapılmış, büyük harfler ve nokta işareti doğru kullanılmış.

Yeterlilik Altı (2 Puan)

- Düşünceler açık ifade edilmemiş, fikirlerde tutarlılık yok, destekleyici detaylara inilmemiş.
- Açık bir giriş, gelişme ve sonuç yok, fikirlerde düzenli bir sıra yok.
- İlginç kelimeler hiç yok, eş anlamlı kelimelerde yanlışlar var, cümleler tam değil.
- Bir çok kelime yanlış yazılmış, büyük harfler ve noktalama işaretlerinde bir çok hata var, paragraflarda uygun girintiler konulmamış, dağınık bir yazım kullanılmış.

Zayıf (1 Puan)

- İçerikten bahsetmek güç, doldurulmak için yazılmış , mantıksız ifadeler var.
- Giriş, gelişme ve sonuç yok, düşüncelerde hiçbir bağlantı yok.
- Kelimeler gelişigüzel kullanılmış ve birçok kelime yanlış seçilmiş, cümle yapıları bozuk.
- Kompozisyon yazım hatalarıyla dolu, kullanılması gereken büyük harfler bazen kullanılmamış, paragraf girintisi yok.

Analitik puanlama metodunda ise ana başlıklar altında ayrıntılı olarak kompozisyonda bulunması gereken özellikleri gösteren liste hazırlanır. Bu listedeki özellikler ağırlıklandırılır. Bu özelliklerin bulunma derecesine göre bir puanlama yapılmaktadır. Aşağıda bu çalışmada kullanılan analitik puanlama anahtarı görülebilir.

KOMPOZİSYON PUANLAMA ANAHTARI

	<u>Puan</u>
I. Başlık	
a) Varlığı	1
b) Konuyla ilişkisi	2
c) Kısaldığı, anlatıcılığı, etkileyciliği	3
II. Anlatım Düzeni	
a) Giriş	
1) Varlığı	1
2) Konuyla ilişkisi	3
3) Sunuş açıklığı	3
4) Etkileyciliği	3
b) Gelişme	
1) Ana fikrin varlığı	2
2) Ana fikrin belirginliği	3
3) Ana fikre ulaşmadaki düşünce zincirinin sağlamlığı, tutarlılığı	5
4) Yardımcı fikirlerin varlığı	2
5) Yardımcı fikirlerin ana fikri destekleme ve tutarlılık dereceleri	5
c) Sonuç	
1) Varlığı	2
2) Ana fikirle tutarlılığı	3
3) Yan fikirlerle desteklenişi	2
4) Kısaldığı, açıklığı, kesinliği	5
III. İfade Zenginliği	
a) Kullanılan kelimelerin çeşitliliği	3
b) Kelimelerin yerindeliği	6
c) Cümle yapısının maksada ve okuyucuya uygunluğu	3
d) Cümle uzunluğunda uygunluk ve çeşitlilik	3
IV. Anlatım Tarza	
a) Dilin doğallığı ve akıcılığı	5
b) Vurgulamaların uygunluğu ve yerindeliği	5
c) Yenilik, kendine özgülük	8
d) Dilin etkili kullanımı	7
V. Yazım Kurallarına Uygunluk	
a) İmla	3
b) Söz dizimi (Gramer)	5
c) Noktalama	3
d) Paragraf organizasyonunun uygunluğu	4
Toplam 100	

Turgut (1997)'dan alınarak geliştirilmiştir.

Yazılı anlatım becerilerinin çoktan seçmeli testlerle dolaylı olarak yoklanması geçerlik açısından sakıncalar içerdiğinden doğrudan kompozisyon sınavlarıyla yoklanmasının gerekliliği konusunda giderek artan bir fikirbirliği oluşmaktadır (Enginarlar 1991). Ancak kompozisyonların uzmanlar tarafından elle puanlanması hem maliyet açısından pahalı hem de zaman alıcı olduğundan kullanışsız yönleri vardır.

Kompozisyon sınavlarından elde edilen puanların güvenilirliğini sağlamak kişisel kanı gerektirdiğinden oldukça zordur. Bu tip sınavların puanlanmasında birden fazla puanlayıcı kullanılmasının güvenilirliği artırıcı rolü olmaktadır. Ancak birden çok puanlayıcı kullanmak maliyet ve zaman açısından ekonomik değildir.

Bu nedenlerle, uzman puanlayıcıların puanlarını kestiricisi olabilecek biçimsel özellikler tespit edilebilirse bu özelliklerden uzman puanlarını yordayarak daha ekonomik bir puanlama yöntemi geliştirilmiş olacaktır. Kompozisyon puanlanırken kullanılan ölçütlerin göstergelerinin belirlenmesiyle uzmanların puanlarını yordamak mümkün olabilecektir.

Kompozisyon tipi sınavları puanlamada biçimsel özelliklerin analizi için puanlamada kullanılan ana özelliklerin belirlenmesine ihtiyaç vardır. Burstein ve arkadaşları (1998c) kompozisyonları puanlarken üç genel özellik kullandığını belirtmişlerdir. Bunlar cümlelerin yapıları (syntactic structure), anlatım düzeni yapısı (rhetorical structure) ve içerik (topical content) olarak belirtilmiştir. Bu üç ana başlığın kompozisyondaki muhtemel göstergelerinin belirlenmesi ve kompozisyondaki bulunuşluk derecelerinin tespit edilmesi ile uzman puanlarını kestirmek mümkün olabilecektir.

Cümle yapıları özelliklerinin göstergeleri olarak farklı cümle tür ve yapıları alınabilir (Burstein ve Chodorow 1999). Cümle ve kelime yapılarının yordayıcıları basit cümle sayıları, bileşik cümle sayıları, sıralı cümle sayıları, yardımcı cümlecik sayısı, master cümlecği sayısı, sıfat cümlecği sayısı, müktedirlik ve olasılık belirten cümlecik sayıları, gereklilik cümlecği sayısı, dilek-şart cümlecği sayısı, soru cümlecği sayısı,

ünlem cümlecığı sayısı, edat sayısı, kompozisyondaki cümle sayısı, ortalama cümle uzunluğu, kelime sayısı, ortalama kelime uzunluğu, yazım hatası sayısı olabilecektir.

Anlatım düzeninde uzmanların ölçütleri belirlemeleri oldukça zordur. Nolan (Akt: Burstein ve arkadaşları 1998a) birçok puanlama anahtarında geçen organizasyonla ilgili “inandırıcı”, “mantıklı”, “anamlı”, “iyi organize edilmiş” sözcüklerinin bulanık anlamlar içerdiğini belirtmektedir. Ayrıca Burstein ve arkadaşları (1998c) anlatım düzeni yapısının, birçok uzmanın bazı ipucu kelimelerle belirlenebileceği görüşünde birleştikleri belirtilmektedir.

Anlatım düzeni yapısında ise başlığın varlığı ve kelime sayısı, kompozisyonun uzunluğu, paragraf sayısı, ortalama paragraf uzunluğu, sonuç bölümünün belirteçleri olarak; “sonuç olarak”, “sonuçta”, “özetle”, “özetlemek gerekirse”, “özet olarak”, “kısaca”, “kısacası” gibi kelimeler kullanılabilir. Örnek verirken kullanılan; “örnek olarak”, “örneğin”, “bir örnek vermek gerekirse”, “mesela”, “misal olarak” kelimeleri örnekleme belirteçleri olarak alınabilir. Ayrıca “bence”, “fikrimce”, “kanımca”, “mümkündür”, “olabilir”, “sanırım” gibi kelimeler kişisel kanıları aktarırken kullanılan ifadeler olarak alınabilir. Bunun yanında “bu”, “bunlar” ve türevleri konunun değişmediğini gösteren ipuçlarıdır (Burstein ve arkadaşları 1998a). Master cümlecığı ile başlayan cümleler yeni bir konuya geçildiğini gösteren ipuçları olarak alınabilmektedir (Burstein ve arkadaşları 1998a).

Kompozisyonun içerik kısmının yordayıcıları olarak anahtar kelimelerin belirlenmesinde kompozisyon konusunun içeriğiyle ilgili yazılmış makaleler taranarak belirlenebilir. Ayrıca kompozisyonlar puanlandıktan sonra yüksek puan almış kompozisyonlardan çıkarılan kelimeler de belirlenerek puanların yordanmasında kullanılabilir. Belirlenen bu anahtar kelimeler çok fazla olacağından, bu konuda yapılmış çalışmalarda kullanılan ağırlıklandırılarak toplamlarının alınabileceği görülmüştür. Bu ağırlıklandırma hem sözcüklerin kompozisyon içindeki diğer sözcükler

bakımından bulunuşluk derecesine göre hem de bu sözcüklerin diğer kompozisyonlarda bulunma derecelerine göre yapılabilir.

Bu araştırmada içerik, anlatım düzeni ve cümle ile ilgili yapılar biçimsel özelliklerine göre analiz edilmiş, bu ana başlıklar etrafında alt yapılar belirlenerek bunların kompozisyonda bulunma dereceleri tespit edilmiş ve buna göre biçimsel özelliklere göre uzman puanları yordanarak bir puan elde edilmiştir. Ayrıca alan uzmanı iki öğretmene hem genel izlenimle hem de analitik bir anahtar ile kompozisyon puanlaması yaptırılmıştır.

Belirlenen yordayıcılarla geçerliği ve güvenilirliği yüksek bir puanlama metodu geliştirilebilirse hem ekonomik olarak hem de zaman açısından daha uygun bir puanlama yöntemi geliştirilmiş olacaktır. Ayrıca puanlamada kişisel kanı gerekmeceğinden daha objektif bir puanlama yapılabilecektir.

Problem Cümlesi

Kompozisyon tipi sınavlarda kompozisyonun biçimsel özellikleri, anahtarla ve genel izlenimle elde edilen puanları ne derece yordamaktadır? Anahtarla puanlamanın ve genel izlenime göre puanlamanın puanlama güvenilirlikleri nasıldır?

Bu problemin çözümü için aşağıdaki alt problemlere yanıt aranacaktır:

Alt Problemler

- 1) Kompozisyon tipi sınavlarda, kompozisyonun biçimsel özellikleri anahtara göre puanlamayı ne derecede yordamaktadır?
- 2) Kompozisyon tipi sınavlarda, kompozisyonun biçimsel özellikleri genel izlenime göre puanlamayı ne derecede yordamaktadır?
- 3) Kompozisyon tipi sınavlarda puanlamaların puanlayıcı güvenilirlikleri nasıldır?

- a) Kompozisyon tipi sınavlarda anahtarla puanlamanın puanlayıcı güvenilirliği nasıldır?
- b) Kompozisyon tipi sınavlarda genel izlenimle puanlamanın puanlayıcı güvenilirliği nasıldır?
- c) Kompozisyon tipi sınavlarda aynı puanlayıcının genel izlenimle ve anahtarla puanlaması arasındaki tutarlılığı nasıldır?

Sayıtlar

1. Sınav uygulanan bireylerin ciddiyetle kompozisyonları yazdıkları kabul edilmiştir.
2. Kompozisyonları puanlayan uzmanların puanlama konusunda yeterli bilgiye sahip oldukları kabul edilmiştir.

Sınırlılıklar

Bu araştırma, veri kaynağı olarak Hacettepe Üniversitesi Eğitim Fakültesi İlköğretim Bölümü Sınıf Öğretmenliği Anabilim Dalı birinci sınıf öğrencileriyle sınırlıdır.

Tanımlar

Yordama: Bilgileri elde bulunmayan ya da gelecekte olabilecek bir durumu bilgileri elde bulunan ve bu değişkenle ilişkili olduğu düşünülen bazı verileri kullanılarak kestirme işi.

Yordayıcı: Yordamada kullanılan değişkenler.

Araştırmanın Amacı ve Önemi

Bu araştırma ile; kompozisyon tipi sınavlarda uzman puanlarının yordayıcıları olan biçimsel özellikleri belirleme ve bu tür sınavlarda kullanılacak geçerliği ve güvenilirliği yüksek bir puanlama yöntemi geliştirebilme amaçlanmaktadır.

Kompozisyon sınavlarında uzman kanıları, objektif testlere göre daha fazla kullanıldığından puanlamaya öznel yargılar daha fazla karışabilmektedir. Bu nedenle öğrenci için önemli kararlar verilebilecek bir sınavda yalnızca bir puanlayıcı kullanmak güvenilirlik açısından sakıncalıdır. İki puanlayıcı kullanmak da büyük gruplarda ekonomik olmamaktadır. Bu sebeple kompozisyonun biçimsel özelliklerinden uzman puanlarının yordayıcıları tanımlanabilirse bu biçimsel özellikler daha sonra yapılabilecek bir bilgisayar programıyla belirlenerek ikinci bir puanlayıcı olarak kullanılabilir. Bu puanlama yöntemi ile hem zaman hem de maliyet açısından daha ekonomik bir puanlama yöntemi olacaktır.

Kompozisyon sınavlarının puanlanmasında biçimsel özelliklerden yordama yapmanın puanlama hatalarını düşüreceği düşünülmektedir. Geçerliği yüksek bir puanlama yöntemi geliştirilebildiği takdirde önemli derecede avantaj sağlanacaktır. Buna benzer bir araştırmanın daha önce ülkemizde yapılmamış olması da bu konuda bir araştırma yapma ihtiyacını ortaya çıkarmıştır.

İlgili Araştırmalar

Page ve Peterson (1995) yaptıkları bir çalışmada kompozisyon tipi sınavları otomatik olarak puanlamışlardır. Yaptıkları çalışmada yazılan kompozisyonların sayısallaştırılabilir yönlerini, örneğin ortalama cümle uzunluğu, paragraf sayısı, kelime olarak kompozisyon uzunluğu, virgül sayısı, parantez sayısı, tamlama bağlaçları sayısı ve ortalama kelime uzunluğu gibi verileri belirleyerek bunları uzman puanlarıyla çoklu regresyona tabi tutmuşlar ve yordamada ağırlıklarını belirleyerek otomatik bir puanlama yöntemi geliştirmişlerdir. Bu çalışmada belirlenen yordayıcılar manidar bulunmamıştır.

Burstein ve Chodorow (1999) tarafından yapılan bir çalışmada iki uzman puanlayıcı ve *e-rater* olarak adlandırılan otomatik puanlama yöntemi kullanılmıştır. Araştırmada anadili İngilizce olmayan bireylere uygulanan İngilizce Yazım Testi (Test of Written English) kullanılmıştır. Bu çalışmada uzman puanlarının yordayıcısı olarak cümle yapıları değişkenleri, anlatım düzeni değişkenleri ve içerik olarak anahtar kelimelerin bulunuşu kullanılmıştır. Sonuç olarak iki uzman puanı arasındaki ilişki 0,75, e-rater puanı ile bir uzman puanlayıcı arasındaki ilişki ise 0,73 olarak bulunmuştur.

Yine Burstein ve arkadaşları (1998a) tarafından GMAT kompozisyon sınavında otomatik puanlama yöntemi kullanılmıştır. Bu sınav çerçevesinde birçok uygulama yapılmış ve uzman puanlarıyla otomatik puanlamadan elde edilen puanlar arasında %87 ile %94 arasında değişen tutarlılıklar gözlenmiştir.

Erickson (2000) yaptığı çalışmada kompozisyon puanlamada anahtar kelimelerden uzmanların bütünsel puanlama anahtarı ile elde ettikleri puanları kestirilmiştir. Bu kestirimler yapılırken sadece anahtar kelimelerin bulunuşluk dereceleri bilgisayar programı kullanarak belirlenmiş, daha sonra bunlar puanların yordayıcıları olarak kullanılmıştır. 100 öğrenci üzerinde yapılan ilk çalışmada kestirilen puanlarla uzman puanları arasında 0,68 ilişki hesaplanmıştır. Daha sonra 127 öğrenciye yapılan kompozisyon sınavı uygulaması yine anahtar kelimelere dayalı bilgisayarla puanlama sistemi kullanılarak puanlanmıştır. Kestirilen puanlar ile uzman puanları arasında 0,61 ilişki hesaplanmıştır.

Karataylı'nın (1996) yaptığı çalışmada öğrencilerin yazma becerilerini ölçmek için hazırlanmış iki farklı anahtarın puanlama güvenilirliklerini incelenmiştir. Uygulanan İngilizce yazılı anlatım (kompozisyon) sınavında iki puanlayıcıya bütünsel (holistic) anahtar ile puanlama yaptırılmış, iki puanlayıcıya da analitik bir anahtarla puanlama yaptırılmıştır. Araştırmada bütünsel anahtarla puanlayan puanlayıcılar arası güvenilirlik katsayısı 0,74, analitik anahtarla puanlayan puanlayıcılar arası güvenilirlik katsayısı da

0.86 bulunmuştur. Bu iki korelasyon katsayısının arasında 0.05 düzeyinde anlamlı bir fark görülmüştür.

Enginarlar (1991)' in yaptığı araştırmada hata sayma, holistic ve analitik puanlama metotlarının puanlayıcı güvenilirlikleri ve aylık uygulanan 5 testin ortalamasıyla uyum geçerlikleri incelenmiştir. Sonuç olarak puanlayıcılar arası güvenilirlik katsayısı hata sayma metodunda 0,66, holistic puanlama metodunda 0,78, analitik puanlama metodunda ise 0,77 bulunmuştur. Uyum geçerliği ise hata sayma metodunda 0,64, holistic metotta 0,73, analitik metotta da 0,71 olarak hesaplanmıştır.



BÖLÜM II

YÖNTEM

Bu bölümde araştırmanın türü, araştırmanın yapıldığı grup, verilerin toplanması ve veri toplama araçları ile verilerin çözümlenmesi konuları üzerinde durulmuştur.

Araştırmanın Türü

Karasar (1991) araştırmaları geleneksel olarak iki kategoriye ayırmıştır. Bunlar temel araştırmalar ve uygulamalı araştırmalardır. Temel araştırmalar varolan bilgilere yenilerini katmayı, uygulamalı araştırmalar ise üretilen bilgilerin değerlendirilmesi ile problemin fiilen çözümünü gerçekleştirmeyi, bilimin olayları denetim altına alma işlevini gerçekleştirmeyi amaçlayan araştırmalardır.

Bu araştırma, öğrencilerin yazılı anlatım becerilerini ölçmede kullanılan kompozisyon türü sınavları puanlamada, kompozisyonun biçimsel özelliklerinden uzman puanlarının yordayıcılarını belirlemeyi ve varolan bilgilere yenilerini katmayı amaçladığından temel araştırma niteliği taşımaktadır.

Bu araştırmada kompozisyonların biçimsel özellikleri ile kompozisyonların uzmanlar tarafından anahtarla ve genel izlenime göre puanlanması arasındaki ilişkiler ve puanlama çeşitlerinin puanlama güvenilirlikleri inceleneceğinden araştırmanın modeli ilişkiyel tarama modelidir diyebiliriz.

Araştırmanın Yapıldığı Grup

Araştırmanın amacı gereği herhangi bir örnekleme yöntemiyle bir örneklem belirlenmemiştir. Araştırma, 2001-2002 öğretim yılında Hacettepe Üniversitesi Eğitim Fakültesi İlköğretim Bölümü Sınıf Öğretmenliği Anabilim Dalı'nın birinci sınıfında öğrenim gören 143 öğrenci üzerinde yürütülmüştür. 143 öğrencinin yazdığı kompozisyonlar Türk Dili ve Edebiyatı alanında uzman 2 öğretmene hem genel izlenimle, hem de puanlama anahtarı kullanılarak puanlatılmıştır. Birinci uzmanın

puanları regresyon analizi için kullanılmış, ikinci puanlayıcının puanları ise puanlayıcı güvenilirliğini hesaplamak için kullanılmıştır. Aynı puanlayıcının genel izlenimle puanlama ile anahtarla puanlama arasındaki tutarlılıklar hesaplanırken her iki puanlayıcının puanları da kullanılmıştır. İkinci puanlayıcı 1 tane kompozisyonu okuyamadığından puanlayıcı güvenilirliği ve ikinci puanlayıcının iki puanlama yöntemi arasındaki tutarlılığı hesaplanırken hesaplamalar 142 öğrenci üzerinden yapılmıştır.

Verilerin Toplanması ve Veri Toplama Araçları

Araştırmada öncelikle öğrencilere “Bazı İnsanlar Reklamların İnsanları Gerçekte İhtiyaç Olmayan Ürünleri Almaya Teşvik Ettiğini, Bazıları ise Reklamların İnsanlara Yaşamı Kolaylaştıracak Yeni Ürünleri Tanıttığını Düşünür. Hangi Görüşle Aynı Fikirdesiniz? Açık Sebepler ve Örneklerle Fikrinizi Destekleyecek Bir Kompozisyon Yazınız” konusunda bir kompozisyon sınavı uygulanarak yazılmış kompozisyonlar elde edilmiştir. Bu kompozisyonlar alan uzmanı iki tane Türk Dili ve Edebiyatı öğretmenine hem bir puanlama anahtarına göre hem de genel izlenime göre puanlatılmıştır.

Kompozisyonun biçimsel özellikleri belirlenirken içerik analizi tekniğinden yararlanılmıştır. İçerik analizinde biçimsel özelliklerin bir kısmı elle sayma şeklinde yapılmış, bir kısmı da bilgisayarda Microsoft Word programı ile belirlenmiştir. Bu biçimsel özellikler cümle yapısı, anlatım yapısı genel başlıkları olarak Tablo 1’de, anahtar kelimeler başlığı halinde de Tablo 2’de düzenlenerek verilmiştir.

Kompozisyonun cümle yapıları analizinde basit cümle sayısı, bileşik cümle sayısı, sıralı cümlecik sayısı, yardımcı cümlecik sayısı, master cümlecığı sayısı, sıfat cümlecığı sayısı, gereklilik cümlecığı sayısı, dilek-şart cümlecığı sayısı, olasılık ve müktedirlik belirten cümlecik sayıları kompozisyonlardan el ile sayılarak analiz edilmiş ve sayıları belirlenmiştir.

Kompozisyonun uzunluğu, cümle sayısı, kelime sayısı, paragraf sayısı, başlığın varlığı, kelime sayısı ve uzunluğu Microsoft Word programıyla belirlenmiştir. Ayrıca

kompozisyonda bulunan yazım hataları sayısı Microsoft Word programının yazım denetimi fonksiyonuyla belirlenmiştir.

Tablo 1: CÜMLE YAPISI VE ANLATIM YAPISI YORDAYICILARI

Cümle Yapısı Yordayıcıları	Anlatım Yapısı Yordayıcıları						
<ul style="list-style-type: none"> • Basit cümle sayısı • Bileşik cümle sayısı • Sıralı cümlecik sayısı • Yardımcı cümlecik sayısı • Master cümlecği sayısı • Sıfat cümlecği sayısı • Gereklilik cümlecği sayısı • Dilek-şart cümlecği sayısı • Olasılık ve müktedirlik belirten cümlecikler • Soru cümlecği sayısı • Ünlem cümlecği sayısı • Edat sayısı • Kompozisyondaki cümle sayısı • Ortalama cümle uzunluğu • Kelime sayısı • Ortalama kelime uzunluğu • Yazım hatası sayısı 	<ul style="list-style-type: none"> • Başlığın varlığı • Başlığın kelime sayısı • Başlığın uzunluğu • Kompozisyonun uzunluğu • Paragraf sayısı • Ortalama paragraf uzunluğu <table border="1"> <tr> <td>Örnek verirken kullanılan kelimeler</td> <td> <ul style="list-style-type: none"> • Örnek olarak • Örneğin • Bir örnek vermek gerekirse • Mesela • Misal olarak </td> </tr> <tr> <td>Kişisel kanıları aktarırken kullanılan kelimeler</td> <td> <ul style="list-style-type: none"> • Bence • Fikrimce • Kanımca • Mümkündür • Olabilir • Sanırım </td> </tr> <tr> <td>Sonuç bölümünün belirteçleri</td> <td> <ul style="list-style-type: none"> • Sonuç olarak • Sonuçta • Özetle • Özetlemek gerekirse • Özet olarak • Kısaca • Kısacası </td> </tr> </table> <ul style="list-style-type: none"> • Konunun değişmediğini gösteren ipuçları “bu”, “bunlar” ve türevleri • Master cümlecği yeni konuya geçildiğini gösteren ipucu 	Örnek verirken kullanılan kelimeler	<ul style="list-style-type: none"> • Örnek olarak • Örneğin • Bir örnek vermek gerekirse • Mesela • Misal olarak 	Kişisel kanıları aktarırken kullanılan kelimeler	<ul style="list-style-type: none"> • Bence • Fikrimce • Kanımca • Mümkündür • Olabilir • Sanırım 	Sonuç bölümünün belirteçleri	<ul style="list-style-type: none"> • Sonuç olarak • Sonuçta • Özetle • Özetlemek gerekirse • Özet olarak • Kısaca • Kısacası
Örnek verirken kullanılan kelimeler	<ul style="list-style-type: none"> • Örnek olarak • Örneğin • Bir örnek vermek gerekirse • Mesela • Misal olarak 						
Kişisel kanıları aktarırken kullanılan kelimeler	<ul style="list-style-type: none"> • Bence • Fikrimce • Kanımca • Mümkündür • Olabilir • Sanırım 						
Sonuç bölümünün belirteçleri	<ul style="list-style-type: none"> • Sonuç olarak • Sonuçta • Özetle • Özetlemek gerekirse • Özet olarak • Kısaca • Kısacası 						

Tablo 2: ANAHTAR KELİMELER

İçeriğin Yordayıcıları Olan Anahtar Kelimeler					
Puanlamadan Önce Yayımlanmış Makalelerden Çıkarılan Anahtar Kelimeler					Puanlamadan Sonra Yüksek Puan Almış Kompozisyonlardan Çıkarılan Anahtar Kelimeler
Abartmak	Dikkat	Kalite	Popüler	Televizyon	İthal
Afiş	Dinlemek	Kampanya	Psikolojik	Tercih	Çağımız
Albeni	Doğru	Kamuoyu	Rakip	Teşvik	Empoze
Aldatmak	Eğitim	Kapitalizm	Rekabet	Ticari	Slogan
Alıcı	Eğlendirici	Kaynak	Reklam ajansı	Tiraj	Sorumluluk
Anlaşılır	Ekonomi	Kitle	Reklamcılık	Tutum	Tatmin
Araştırma	Eleman	Kredi Kartı	Resim	Tüketim	Vitrin
Arz	Geleneksel	Kullanmak	Reyting	Tüketim Çılgınlığı	Yenilik
Basın	Gelişmekte Olan	Kurum	Rol	Tüketim Toplumu	Kalkınma
Başarı	Gereksinim	Küresel	Sağlık	Uygar	
Beğendirme	Görsel	Lüks	Satın Alma	Uyumlu	
Beğeni	Güdüleme	Maliyet	Satış	Uzman	
Benimsetme	Hatırlanma	MarkaBağımlılığı	Satma	Ücret	
Bilboard	Hedef	Medya	Savaş	Üretim	
Bilgi	Hizmet	Mesaj	Seçme	Üretici	
Bilgilendirme	İhtiyaç	Modern	Sektör	Ürün	
Bilgisayar	İkna	Müşteri	Sembol	Verimli	
Bilimsel	İlan	Müzik	Serbest Piyasa	Yanıltma	
Bilişim	İletişim	Odak	Sinir Bozucu	Yaratıcı	
Birden Fazla	İlgi	Olumlu	Sorgu	Yaygınlaşma	
Çeşitlilik	İlkel	Olumsuz	Sömürme	Yazı	
Çıkar	İmaj	Özellik	Söz	Yönelme	
Çocuk	İnandırıcı	Özendirme	Sponsor	Zengin	
Davranış	İndirim	Pay	Standart		
Değerlendirme	İnternet	Pazar	Şirket		
Değişik	İstek	Piyasa	Talep		
Denetim	İzlenme	Planlama	Tanıtım		
Dergi	Kadın		Teknik		

Kompozisyonun anlatım yapısının yordayıcıları olarak belirlenen örnek verirken kullanılan ipucu kelimeleri, konunun devam ettiğini gösteren ipucu kelimeleri, kişisel

kanaatleri belirtirken kullanılan ipucu kelimeleri ve sonuç bölümünde kullanılan ipucu kelimeleri Microsoft Word programıyla belirlenmiştir.

Anahtar kelimelerin bulunuşluk dereceleri de yine Microsoft Word programının “Bul” ve “Değiştir” komutlarıyla belirlenmiş ve sayısallaştırılmıştır. Dergilerde yayınlanmış makalelerden çıkarılan anahtar kelimeler her kompozisyona göre Burstein ve Marcu (2000)’dan alınan yöntemle göre ağırlıklandırılmıştır. Ağırlıklandırma aşağıdaki formül kullanılarak yapılmıştır.

$$\text{ağırlık}_{wj} = (\text{freq}_{wj} / \text{maxfreq}_j) \times \log(\text{nkomp} / \text{komp}_w)$$

Bu formülde,

ağırlık_{wj} : w kelimesinin j kompozisyonundaki ağırlığı

freq_{wj} : w kelimesinin j kompozisyonundaki frekansı

maxfreq_j : j kompozisyonundaki maksimum frekanslı kelimenin frekansı

nkomp : regresyon analizinde kullanılacak toplam kompozisyon sayısı

komp_w : w kelimesinin bulunduğu toplam kompozisyon sayısı

Yukarıdaki formüle göre belirlenen kelimeler her bir kompozisyondaki ve tüm kompozisyonlardaki bulunuşluğuna göre ağırlıklandırmaları yapılmıştır. Daha sonra her kompozisyondaki kelimeler için hesaplanan ağırlıkların toplamları alınarak tüm anahtar kelimeler için her bir öğrencinin bir tek ağırlıklı anahtar kelime puanı elde edilmiştir.

Kompozisyonlar puanlandıktan sonra yüksek puan almış kompozisyonlardan daha önceki anahtar kelimelerde olmayan yeni anahtar kelimeler belirlenmiştir. Belirlenen anahtar kelimelerin sayıları toplanarak ayrı bir anahtar kelime puanı elde edilmiştir. Belirlenen bu anahtar kelimelerin seçilme yönteminden dolayı kompozisyon puanlarıyla doğrudan ilişkili olduğu düşünülerek ağırlıklandırmaya gidilmemiştir.

Verilerin Analizi

Kompozisyonların olası yordayıcıları uzmanların anahtara göre ve genel izlenime göre verdikleri puanlarla regresyon analizine tabi tutularak yordayıcılık güçleri belirlenmiştir.

1. Birinci alt problemin araştırılması için kompozisyon sınavını puanlamada anahtara göre verilen puanlar bağımlı değişken olarak alındığında kompozisyonun belirlenen biçimsel özelliklerine göre puanlanmasında hangi biçimsel özelliklerin anlamlı birer yordayıcı olduğunu belirlemek için biçimsel özellikler bağımsız değişken olarak alınıp regresyon analizi yapılmıştır. Birden çok bağımsız değişken bulunduğu için çoklu regresyon analizi kullanılmıştır. Regresyon katsayılarının anlamlılığının test edilmesinde t testi, regresyon denkleminin önemliliğinin test edilmesinde ise F testi kullanılmıştır. Regresyon analizi SPSS programı yardımıyla yapılmıştır.
2. İkinci alt problemin araştırılmasında kompozisyon sınavını puanlamada genel izlenime göre verilen puanlar bağımlı değişken olarak alındığında kompozisyonun belirlenen biçimsel özelliklerine göre puanlanmasında hangi biçimsel özelliklerin anlamlı birer yordayıcı olduğunu belirlemek için biçimsel özellikler bağımsız değişken olarak alınıp regresyon analizi yapılmıştır. Birden çok bağımsız değişken bulunduğu için çoklu regresyon analizi kullanılmıştır. Regresyon katsayılarının anlamlılığının test edilmesinde t testi, regresyon denkleminin önemliliğinin test edilmesinde ise F testi kullanılmıştır. Regresyon analizi SPSS programı yardımıyla yapılmıştır.

Bu araştırmada uzmanların verdikleri puanları yordamak için aşağıdaki formülden yararlanarak her iki puanlama türünün kestiricisi puanlar elde edilmiştir.

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon \quad \text{Ünver ve Gamgam (1986)}$$

β parametrelerini kestirmek için kullanılacak formül matrisler çarpımı şeklinde aşağıdaki gibi ifade edilebilir:

$$\underline{b} = (X'X)^{-1}(X'Y) \quad \text{Ünver ve Gamgam (1986)}$$

Matris olarak gösterirsek,

$$X = \begin{bmatrix} 1 & X_{11} & X_{12} & X_{13} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & X_{23} & \dots & X_{2k} \\ 1 & X_{31} & X_{32} & X_{33} & \dots & X_{3k} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & X_{n3} & \dots & X_{nk} \end{bmatrix} \quad Y = \begin{bmatrix} Y_{A1} \\ Y_{A2} \\ Y_{A3} \\ \vdots \\ Y_{An} \end{bmatrix} \quad \underline{b} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}$$

(Bağımsız değişkenler)

(Bağımlı değişken)

Katsayılar

3. Üçüncü alt problemin araştırılması için,

- a) İki puanlayıcının genel izlenimle puanlaması sonucu elde edilen puanları arasındaki puanlayıcı güvenilirliği hesaplanmıştır. İki puanlayıcı arasındaki puanlayıcı güvenilirliği, Pearson Momentler Çarpımı Korelasyon Katsayısı kullanılarak hesaplanmıştır.
- b) İki puanlayıcının puanlama anahtarı kullanarak puanlaması sonucu elde edilen puanları arasındaki puanlayıcı güvenilirliği hesaplanmıştır. İki puanlayıcı arasındaki puanlayıcı güvenilirliği yine Pearson Momentler Çarpımı Korelasyon Katsayısı kullanılarak hesaplanmıştır.
- c) Aynı puanlayıcının genel izlenimle ve anahtarla puanlaması arasındaki tutarlılığı hesaplamak için her iki puanlayıcı için genel izlenimle

puanlamadan elde edilen puanlarla anahtarla elde edilen puanlar arasındaki tutarlılık, Pearson Momentler Çarpımı Korelasyon Katsayısı ile hesaplanmıştır.



BÖLÜM III

BULGULAR VE YORUM

Bu bölümde, araştırmada incelenen problemin ve alt problemlerin çözümü için, verilerin çözümlenmesi sonucunda elde edilen bulgulara ve bu bulgulara ilişkin yorumlara yer verilmiştir.

Verilerin analizinde kullanılan iki puanlayıcının anahtarla ve genel izlenimle elde edilen puanlarına ilişkin betimleyici istatistikler Tablo 3'te verilmiştir.

Tablo 3: Puanlamalara Ait Betimleyici İstatistikler

		\bar{X}	S_x	N
1. Puanlayıcı	Anahtar Puanı	57,66	10,68	143
	Genel İzlenim Puanı	59,44	9,28	143
2. Puanlayıcı	Anahtar Puanı	38,49	16,10	142
	Genel İzlenim Puanı	40,42	13,08	142

Regresyon analizinde kullanılan 1. puanlayıcının verdiği puanlar için yapılan Kolmogrov-Simironov normallik testinde değerler anahtarla puanlamada $K-S_{\text{anah}}=1,073$ $p=0,20$ ve genel izlenimde $K-S_{g.iz}=1,594$ $p=0,012$ bulunmuştur. Bu değerler 0,01 manidarlık düzeyinde normal dağılım gösterdiklerini belirlemektedir.

Verilerin analizi sonucu elde edilen bulgular ve yorumlar alt problemlerin sırasına göre aşağıda sunulmuştur.

1. Alt Probleme İlişkin Bulgular ve Yorum

Kompozisyon tipi sınavlarda, kompozisyonun biçimsel özellikleri anahtara göre puanlamayı ne derecede yordamaktadır?

Bu alt probleme yanıt bulabilmek için kompozisyonlardan çıkarılan biçimsel özelliklere ait veriler ile kompozisyonun anahtarla puanlanması sonucu elde edilen puanlar arasında çoklu regresyon analizi yapılmıştır. Bu regresyon analizi sonucunda

elde edilen regresyon denkleminin önemliliği test edilmiştir. Regresyon denkleminin önemliliğinin test edilmesinde F testi kullanılmıştır. Regresyon denkleminin ilgili ANOVA tablosu Tablo 4’te verilmiştir.

Tablo 4: Anahtarla Puanlama ile Kompozisyonun Biçimsel Özelliklerinin Regresyon Analizine İlişkin ANOVA Tablosu

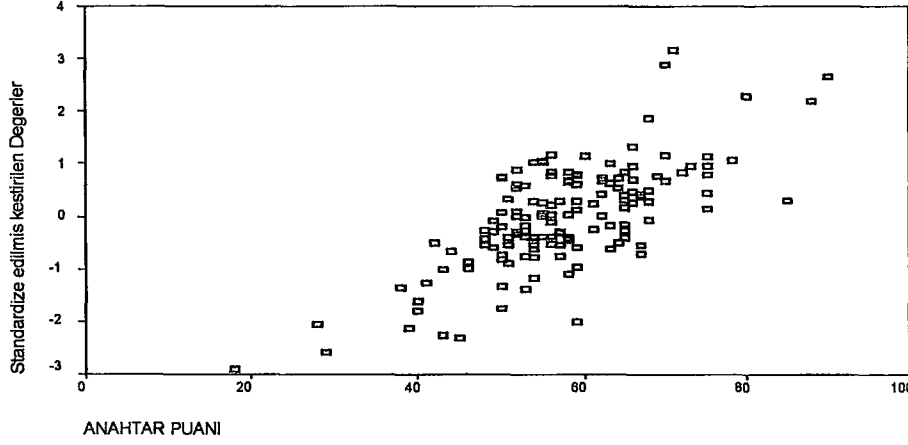
Varyansın Kaynağı	Kareler Toplamı	Serbestlik Derecesi	Kareler Ortalaması	F	Manidarlık
Regresyon	8334,928	28	297,676	4,324	,000
Hata	7847,282	114	68,836		
Toplam	16182,210	142			

Tablo 4’te görüldüğü gibi oluşturulan regresyon denkleminin 0,05 manidarlık düzeyinde önemlidir. Elde edilen bulgu, kestirilen kısmi regresyon katsayılarıyla oluşturulacak denklemin kullanılarak puanlayıcının anahtarla yaptığı puanlamanın kestirebileceği şeklinde yorumlanabilir.

Regresyon analizi sonucunda elde edilen regresyon denkleminde bağımsız değişkenler, bağımlı değişken olan kompozisyonun anahtarla puanlanmasından elde edilen puanlardaki değişimin yüzde kaçını açıklayabildiğini belirlemek için çoklu korelasyon katsayısının karesi alınarak belirtme katsayısı hesaplanmıştır.

Regresyon analizi sonucunda kullanılan bağımsız değişkenlerle bağımlı değişken arasındaki çoklu korelasyon katsayısı 0,72, belirtme katsayısı ise 0,52, kestirmenin standart hatası da 8,30 olarak hesaplanmıştır. Buradan bağımsız değişkenlerin bağımlı değişkendeki değişkenliğin yüzde 52’sini açıklayabildiğini söylenebilir.

Kompozisyonun anahtarla puanlanması sonucu elde edilen puanların gözlenen değerleri ile kestirilen değerlerinin standartlaştırılmış değerlerine ait saçılma grafiği Şekil 1’de görülmektedir.



Şekil 1: Anahtarla puanlama Gözlenen Değerler ve Standartlaştırılmış Kestirilen Değerler Saçılma Grafiği

Grafikten anlaşılacağı gibi kestirmelerin hataları düşük puanlarda daha az, yüksek puanlarda ise daha fazla olduğu gözükmektedir.

Regresyon analizi sonucunda elde edilen kısmi regresyon katsayılarının anlamlılığı t testi ile test edilmiştir. t testi sonucunda kompozisyonun anahtarla puanlanmasının en iyi yordayıcıları belirlenmiştir. Ayrıca tüm bağımsız değişkenler yordayıcılık güçlerine göre sıralanmıştır.

Regresyon analizi sonucunda elde edilen kısmi regresyon katsayıları, bunlara ait standart sapmalar, t değerleri ve manidarlık düzeyi Tablo 5'te verilmiştir.

Tablo 5'te görüldüğü gibi uzman puanlayıcıların puanlama anahtarı kullanarak verdiği puanların en iyi yordayıcıları olarak "*Yazım Hatası Sayısı*", "*Ortalama Cümle Uzunluğu*", "*Yayınlanmış Makalelerden Çıkarılmış ve Ağırlıklandırılmış Anahtar Kelimeler*", "*Başlığın Var Olup Olmaması*" ve "*Basit Cümle Sayısı*" gibi

kompozisyonların biçimsel özelliklerine ilişkin değişkenler belirlenmiştir. Bu yordayıcıların t değerleri 0,05 düzeyinde manidar bulunmuştur.

Tablo 5: Kısmi Regresyon Katsayıları, Standart Hataları, t değerleri ve Manidarlık Düzeyleri

Bağımsız Değişkenler	Katsayılar (β)	Standart Hata	t	p
(Sabit)	44,365	37,499	1,183	,239
Başlığın Varlığı	9,141	3,562	2,566	,012
Başlığın Kelime Sayısı	0,956	1,063	0,900	,370
Başlığın Uzunluğu	-0,077	0,152	-0,511	,611
Kompozisyonun Uzunluğu	0,004	0,024	0,146	,884
Kelime Sayısı	-0,063	0,172	-0,366	,715
Ortalama Kelime Uzunlu.	-4,034	5,470	-0,737	,462
Ortalama Cümle Uzunluğu	0,220	0,077	2,859	,005
Paragraf Sayısı	-0,081	1,026	-0,079	,937
Ortalama Paragraf Uzun.	-0,009	0,011	-0,854	,395
Yazım Hatası Sayısı	-1,340	0,404	-3,316	,001
Basit Cümle Sayısı	0,998	0,394	2,535	,013
Bileşik Cümle Sayısı	2,951	1,779	1,659	,100
Sıralı Cümle Sayısı	0,516	0,637	0,810	,419
Yardımcı Cümlecik Sayısı	-0,820	1,401	-0,586	,559
Mastar Cümlecığı Sayısı	0,035	0,280	0,124	,901
Sıfat Cümlecığı Sayısı	0,266	0,330	0,806	,422
Olasılık ve Muktedirl. Bel. Cüm. Say.	0,034	0,366	0,092	,927
Gereklilik Cümlecığı Sayısı	0,010	0,477	0,021	,983
Dilek-Şart Cümlecığı Sayısı	-0,509	0,783	-0,650	,517
Edat Sayısı	0,013	0,169	0,079	,937
Soru Cümlesi Sayısı	0,967	0,561	1,724	,087
Ünlem Cümlesi Sayısı	0,071	1,189	0,060	,952
Kişisel Kanıları Bel.Kul.Kel	1,275	0,742	1,718	,088
Örnek Vermede Kullanılan Kelimeler	-1,245	1,615	-0,771	,442
Bu, Bunlar ve türevleri	0,346	0,274	1,264	,209
Sonuç Kısmının Belirteci Kelimeler	-0,989	1,611	-0,614	,540
Ağırlıklandırılmış Anahtar Kelimeler	1,312	0,492	2,668	,009
Pualamadan Sonra Çıkarıl. Anah.Kel	1,780	1,168	1,524	,130

Diğer yordayıcılar 0,05 düzeyine göre manidar bulunmamıştır ama yordayıcılık güçleri bakımından sırayla şu şekilde sıralanmıştır: “Soru Cümlesi Sayısı”, “Kişisel

Kanı Belirtirken Kullanılan Kelimelerin Sayısı, *“Bileşik Cümle Sayısı”*, *“Kompozisyon Puanlandıktan Sonra Çıkarılan Anahtar Kelimeler”*, *“Bu, Bunlar ve Türevleri”*, *“Başlığın Kelime Sayısı”*, *“Ortalama Paragraf Uzunluğu”*, *“Sıralı Cümle Sayısı”*, *“Sıfat Cümlesi Sayısı”*, *“Örnek Verirken Kullanılan Kelimelerin Varlığı”*, *“Ortalama Kelime Uzunluğu”*, *“Dilek-Şart Cümlesi Sayısı”*, *“Sonuç Bölümünün Belirteçleri Kelimeler”*, *“Yardımcı Cümlecik Sayısı”*, *“Başlığın Uzunluğu”*, *“Kelime Sayısı”*, *“Kompozisyonun Uzunluğu”*, *“Mastar Cümleciği Sayısı”*, *“Muktedirlik ve Olasılık Bildiren Cümlecik Sayısı”*, *“Edat Sayısı”*, *“Paragraf Sayısı”*, *“Ünlem Cümlesi Sayısı”*, *“Gereklilik Cümlesi Sayısı”* şeklinde sıralanmıştır. Ayrıca *“Cümle Sayısı”* ile ilgili veriler SPSS programı tarafından tolerans sınırı aşıldığından analizden çıkarılmıştır.

2. Alt Probleme İlişkin Bulgular ve Yorum

Kompozisyon tipi sınavlarda, kompozisyonun biçimsel özellikleri genel izlenime göre puanlamayı ne derecede yordamaktadır?

Bu alt probleme yanıt bulabilmek için kompozisyonlardan çıkarılan biçimsel özelliklere ait veriler ile kompozisyonun genel izlenimle puanlanmasından elde edilen puanlar arasında yine çoklu regresyon analizi yapılmıştır. Bu regresyon analizi sonucunda elde edilen regresyon denkleminin önemliliği test edilmiştir. Regresyon denkleminin önemliliğinin test edilmesinde F testi kullanılmıştır. Regresyon denkleminin ilgili ANOVA tablosu Tablo 6’da verilmiştir.

Tablo 6: Genel İzlenimle Puanlama ile Kompozisyonun Biçimsel Özelliklerinin Regresyon Analizine İlişkin ANOVA Tablosu

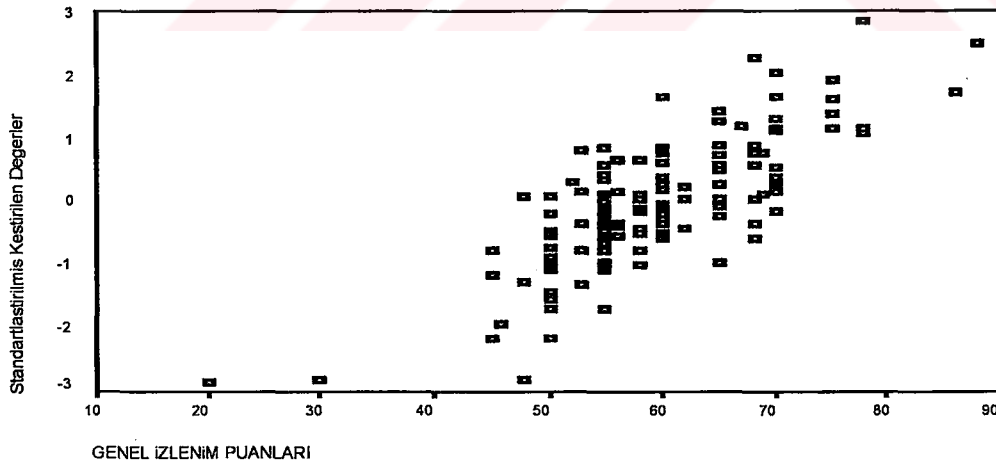
Varyansın Kaynağı	Kareler Toplamı	Serbestlik Derecesi	Kareler Ortalaması	F	Manidarlık
Regresyon	7310,217	28	261,079	6,051	,000
Hata	4919,028	114	43,149		
Toplam	12229,245	142			

Tablo 6’da görüldüğü gibi oluşturulan regresyon denklemi 0,05 manidarlık düzeyinde önemlidir. Elde edilen bulgu, bu regresyon denkleminin kompozisyon puanlarını kestirmek için kullanılabilceği şeklinde yorumlanabilir.

Regresyon analizi sonucunda elde edilen regresyon denkleminde bağımsız değişkenler bağımlı değişken olan kompozisyonun genel izlenimle puanlanmasından elde edilen puanlardaki değişimin yüzde kaçını açıklayabildiğini belirlemek için çoklu korelasyon katsayısının karesi alınarak belirtme katsayısı hesaplanmıştır.

Regresyon analizinde kullanılan bağımsız değişkenler ile bağımlı değişken arasındaki çoklu korelasyon katsayısı 0,77 olarak, belirtme katsayısı ise 0,60 olarak, kestirmenin standart hatası ise 6,57 olarak hesaplanmıştır. Buradan bağımsız değişkenlerin bağımlı değişkendeki varyansın yüzde 60’ını açıklayabildiğini söyleyebiliriz.

Kompozisyonun genel izlenimle puanlanması sonucu elde edilen puanların gözlenen değerleri ile standartlaştırılmış kestirilen değerlerine ait saçılma grafiği Şekil 2’de verilmiştir.



Şekil 2: Genel İzlenimle Puanlamanın Gözlenen Değerleri ile Standartlaştırılmış Kestirilen Değerlerine Ait Saçılma Grafiği

Grafikte görüldüğü gibi çok düşük ve yüksek puanlarda kestirilen puanlarla gözlenen puanlar arasında daha büyük farklar gözlenirken ortadaki puanlarda daha fazla tutarlılık gözlenmektedir.

Regresyon analizi sonucunda elde edilen kısmi regresyon katsayılarının anlamlılığı t testi ile test edilmiştir. t testi sonucunda kompozisyonun genel izlenimle puanlanmasının en iyi kestiricileri belirlenmiştir.

Regresyon analizi sonucunda elde edilen kısmi regresyon katsayıları, bunların standart sapmaları, t değerleri ve manidarlık düzeyleri Tablo 7’de verilmiştir. “Cümle Sayısı” değişkeni SPSS programı tarafından tolerans sınırı aşıldığından analizden çıkarılmıştır.

Tablo 7’de görüldüğü gibi uzman puanlayıcıların genel izlenimle verdiği puanların en iyi yordayıcıları olarak “*Başlığın Var Olup Olmaması*”, “*Örnek Verirken Kullanılan Kelimelerin Varlığı*”, “*Kompozisyon Puanlandıktan Sonra Çıkarılan Anahtar Kelimeler*”, “*Soru Cümlesi Sayısı*” ve “*Yayınlanmış Makalelerden Çıkarılmış ve Ağırlıklandırılmış Anahtar Kelimeler*” gibi biçimsel özellik değişkenleri belirlenmiştir. Bu yordayıcılar, 0,05 manidarlık düzeyinde anlamlı bulunan yordayıcılardır.

Diğer yordayıcılar 0,05 manidarlık düzeyine göre önemli bulunmamıştır. Bu değişkenler yordayıcılık güçleri bakımından sırayla şu şekilde sıralanmıştır: “*Ortalama Cümle Uzunluğu*”, “*Sonuç Bölümünün Belirteçleri Olan Kelimeler*”, “*Ünlem Cümlesi Sayısı*”, “*Başlığın Kelime Sayısı*”, “*Yazım Hatası Sayısı*”, “*Muktedirlik ve Olasılık Bildiren Cümlecik Sayısı*”, “*Başlığın Uzunluğu*”, “*Basit Cümle Sayısı*”, “*Dilek-Şart Cümlesi Sayısı*”, “*Sıfat Cümlesi Sayısı*”, “*Edat Sayısı*”, “*Paragraf Sayısı*”, “*Ortalama Paragraf Uzunluğu*”, “*Gereklilik Cümlesi Sayısı*”, “*Kelime Sayısı*”, “*Kompozisyonun Uzunluğu*”, “*Yardımcı Cümlecik Sayısı*”, “*Bileşik Cümle Sayısı*”, “*Kişisel Kanı Belirtirken Kullanılan Kelimelerin Sayısı*”, “*Sıralı Cümle Sayısı*”, “*Mastar Cümlecği*”

Sayısı”, “Bu, Bunlar ve Türevleri”, “Ortalama Kelime Uzunluğu”, şeklinde sıralanmıştır..

Tablo 7: Kısmi Regresyon Katsayıları, Standart Hataları, t değerleri ve Manidarlık Düzeyleri

Bağımsız Değişkenler	Katsayılar (β)	Standart Hata	t	p
(Sabit)	28,228	29,689	0,951	,344
Başlığın Varlığı	10,545	2,820	3,740	,000
Başlığın Kelime Sayısı	1,121	0,842	1,332	,186
Başlığın Uzunluğu	-0,127	0,120	-1,060	,292
Kompozisyonun Uzunluğu	-0,011	0,019	-0,570	,570
Kelime Sayısı	0,086	0,136	0,627	,532
Ortalama Kelime Uzunlu.	-0,348	4,331	-0,080	,936
Ortalama Cümle Uzunluğu	0,117	0,061	1,917	,058
Paragraf Sayısı	0,581	0,812	0,715	,476
Ortalama Paragraf Uzun.	0,006	0,009	0,710	,479
Yazım Hatası Sayısı	-0,382	0,320	-1,194	,235
Basit Cümle Sayısı	0,317	0,312	1,016	,312
Bileşik Cümle Sayısı	0,565	1,409	0,401	,689
Sıralı Cümle Sayısı	-0,153	0,505	-0,303	,763
Yardımcı Cümlecik Sayısı	-0,587	1,109	-0,529	,598
Mastar Cümlecği Sayısı	-0,063	0,221	-0,284	,777
Sıfat Cümlecği Sayısı	0,222	0,262	0,848	,398
Olasılık ve Muktedirl. Bel. Cüm. Say.	0,346	0,290	1,195	,235
Gereklilik Cümlecği Sayısı	0,264	0,378	0,698	,487
Dilek-Şart Cümlecği Sayısı	0,553	0,620	0,892	,374
Edat Sayısı	-0,100	0,134	-0,749	,455
Soru Cümlesi Sayısı	1,234	0,444	2,779	,006
Ünlem Cümlesi Sayısı	-1,310	0,942	-1,391	,167
Kişisel Kanıları Bel.Kul.Kel	0,201	0,588	0,342	,733
Örnek Vermede Kullanılan Kelimeler	-3,747	1,279	-2,930	,004
Bu, Bunlar ve türevleri	0,048	0,217	0,217	,828
Sonuç Kısmının Belirteç Kelimeleri	-2,208	1,276	-1,731	,086
Ağırlıklandırılmış Anahtar Kelimeler	2,669	0,924	2,887	,005
Pualamadan Sonra Çıkarıl. Anah.Kel	0,902	0,389	2,318	,022

Anahtarla puanlama ve genel izlenimle puanlamanın ortak kestiricileri olarak başlığın var olup olmaması ve ağırlıklandırılmış anahatar kelimeler belirlenmiştir. Buradan

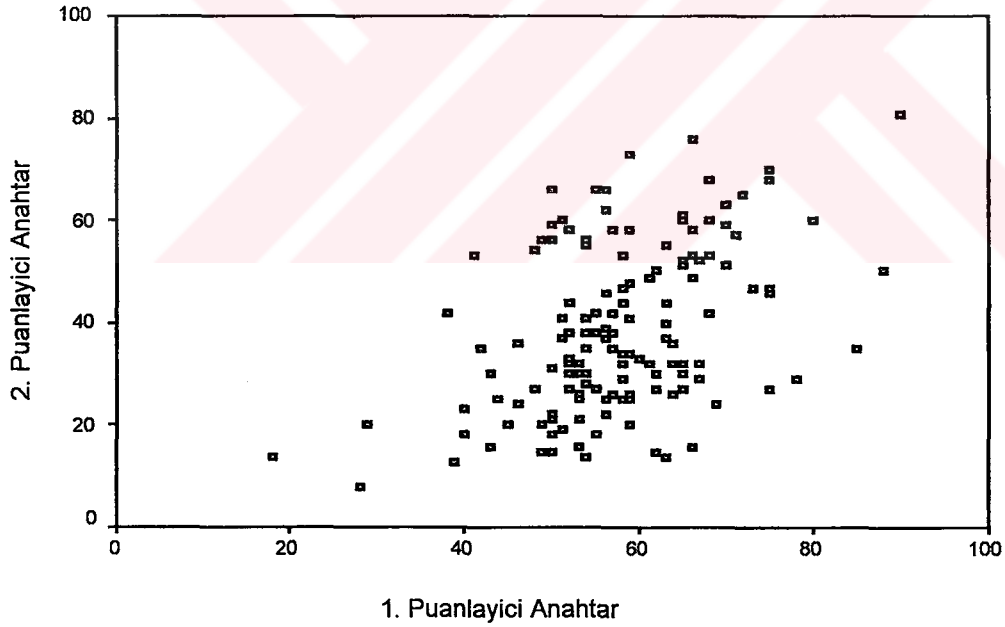
puanlayıcının puanlama türüne bağımlı olmadan başlığın var olup olmaması ve bazı kelimelerin kullanılmasından etkilendiğini söyleyebiliriz.

3. Alt Probleme İlişkin Bulgular ve Yorum

Kompozisyon tipi sınavlarda puanlamaların puanlayıcı güvenilirlikleri nasıldır?

- a) Kompozisyonun biçimsel özelliklerinden anahtarla puanlamanın puanlayıcı güvenilirliği nasıldır?

Bu alt probleme çözüm bulmak için, birinci puanlayıcı ile ikinci puanlayıcının puanlama anahtarına göre elde edilen puanları arasında Pearson Momentler Çarpımı korelasyon katsayısı hesaplanmıştır. Bu korelasyon katsayısı puanlayıcı güvenilirliği olarak kullanılmıştır.

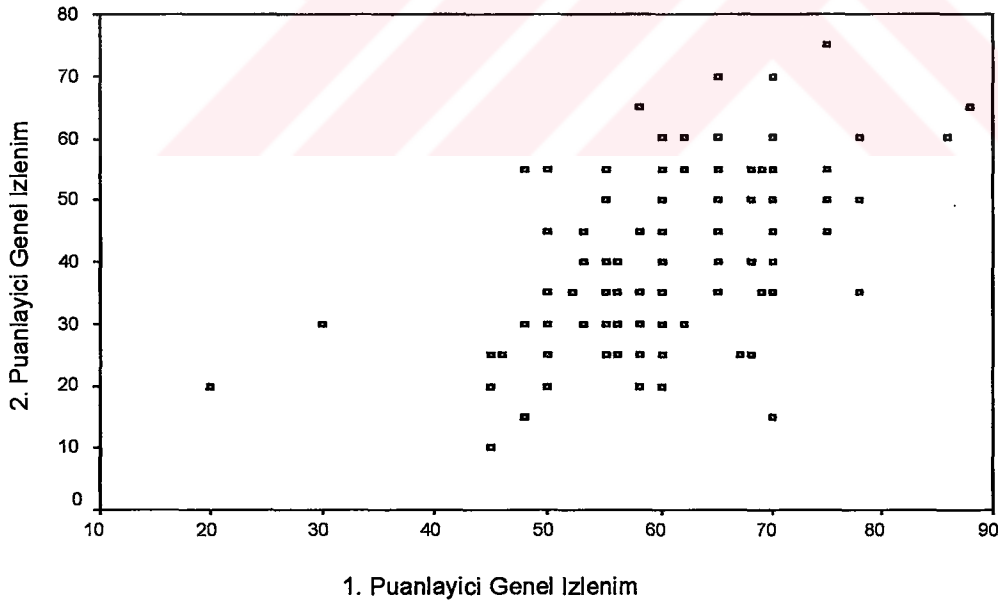


Şekil 3: Anahtarla Puanlamada İki Puanlayıcının Verdiği Puanlar Arasındaki İlişkiye Ait Grafik

Hesaplanan korelasyon katsayısı sonucunda güvenilirlik katsayısı $r = 0,45$ bulunmuştur. Puanlayıcılar arası tutarlılık anlamındaki bu güvenilirlik katsayısı yeterince yüksek değildir. Şekil 3'te iki puanlayıcının anahtarla verdiği puanlara ait ilişki grafiğinde görüldüğü gibi orta düzeyde bir ilişki görülmektedir. Bunun sebebi yazılı anlatım gibi kişisel kanıların ön plana fazlaca çıkaran bir konu olduğundan kaynaklanıyor olabilir. Puanlama anahtarı kullanarak yapılan bir puanlamada tutarlılığın düşük olması puanlamaya esas olan özelliklerin yeterince net olmamasından da kaynaklanıyor olabilir.

- b) Kompozisyonun biçimsel özelliklerinden genel izlenimle puanlamanın puanlayıcı güvenirligi nasıldır?

Bu alt probleme çözüm bulmak için, birinci puanlayıcı ile ikinci puanlayıcının genel izlenimle verdikleri puanlar arasında Pearson Momentler Çarpımı Korelasyon Katsayısı hesaplanmıştır.



Şekil 4: Genel İzlenimle Puanlamada İki Puanlayıcının Verdiği Puanlar Arasındaki İlişkiye Ait Grafik

Hesaplanan korelasyon katsayısı sonucunda güvenilirlik katsayısı $r = 0,53$ bulunmuştur. Şekil 4'te de görülebileceği gibi puanlayıcılar arası tutarlılık anlamındaki bu güvenilirlik katsayısı orta düzeyde diyebiliriz.

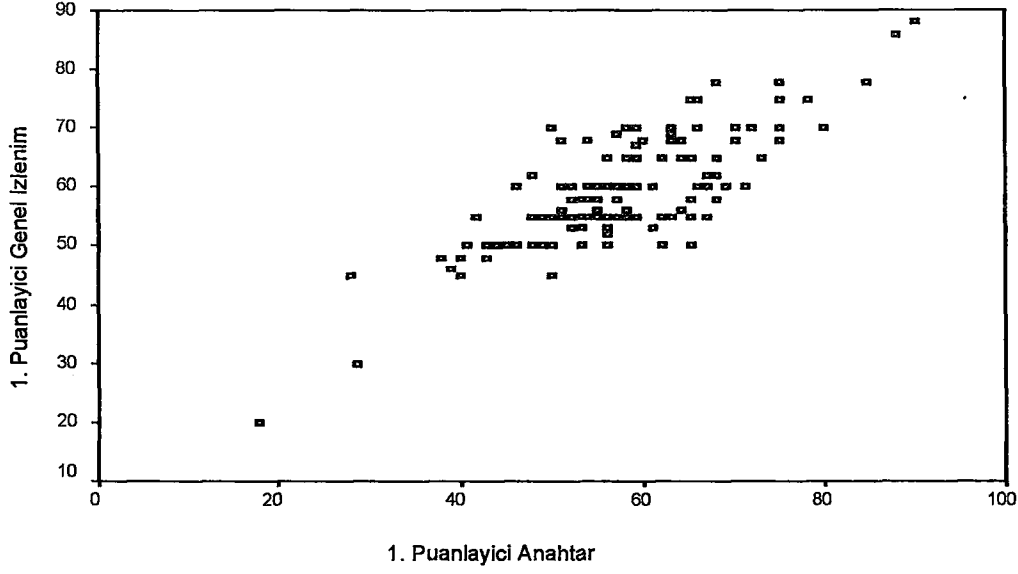
Genel izlenimle yapılan bu puanlamaların arasındaki tutarlılığın anahtarla yapılan puanlamalar arasındaki tutarlılıktan fazla olmasını beklenmediği halde daha yüksek çıkması bu iki korelasyon sayısı arasında fark olup olmadığını test etmeye yöneltmiştir. Korelasyon katsayılarının test edilmesi sonucunda 0,05 düzeyinde aralarında manidar bir fark bulunmamıştır. İki puanlama türünün puanlayıcılar arası tutarlılık anlamındaki güvenilirliğinin farklı olmadığını söyleyebiliriz.

Puanlayıcı güvenilirliklerinin tam olmamasından dolayı birinci ve ikinci alt problemde manidar çıkmayan bazı değişkenlerin yordayıcılık düzeyleri ortaya çıkmamış olabilir. Bu nedenle puanlayıcı güvenliği daha yüksek puanlarla çalışma yapıldığında daha net bir yorum yapılabilir.

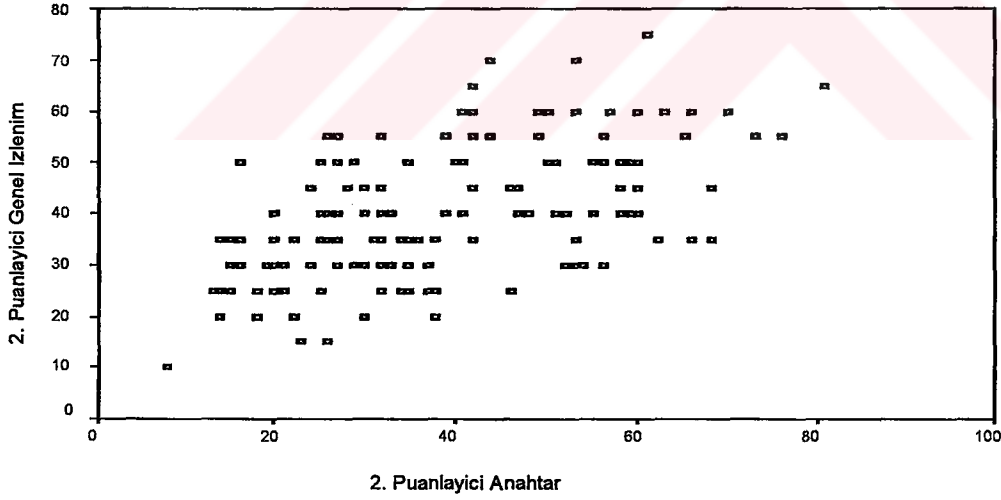
- c) Kompozisyon tipi sınavlarda aynı puanlayıcının genel izlenimle ve anahtarla puanlaması arasındaki tutarlılığı nasıldır?

Bu alt problemi çözmek için, her puanlayıcının anahtarla ve genel izlenimle verdikleri puanlar arasında Pearson Momentler Çarpımı korelasyon katsayısı hesaplanarak iki puanlama metodu arasındaki tutarlılığa bakılmıştır.

Hesaplanan korelasyon katsayısı sonucunda birinci puanlayıcının anahtarla ve genel izlenimle verdikleri puanlar arasında 0,78 gibi bir tutarlılık gözlenirken ikinci puanlayıcının her iki puanı arasında 0,55 gibi bir tutarlılık gözlenmiştir. Birinci puanlayıcıda yüksek bir tutarlılık gözlenirken ikinci puanlayıcıda daha düşük bir tutarlılık görülmektedir. İki korelasyon katsayısının test edilmesi sonucunda 0,05 düzeyinde farklı oldukları sonucuna varılmıştır.



Şekil 5: Birinci Puanlayıcının Anahtarla ve Genel İzlenimle Verdiği Puanlara Ait İlişki Grafiği



Şekil 6: İkinci Puanlayıcının Anahtarla ve Genel İzlenimle Verdiği Puanlara Ait İlişki Grafiği

Şekil 5'te birinci puanlayıcının verdiği puanlar arasındaki yüksek tutarlılık açıkça görülmektedir. Şekil 6'da ise ikinci puanlayıcının puanlarında daha düşük bir tutarlılık olduğu görülmektedir.

BÖLÜM IV

SONUÇ VE ÖNERİLER

Bu araştırmanın amacı problem cümlesi ve alt problemlerde belirtildiği gibi (1) Kompozisyon tipi sınavlarda kompozisyonun biçimsel özelliklerinin anahtarla puanlamayı ne derece yordayabildiğini ortaya çıkarmak, (2) Kompozisyon tipi sınavlarda kompozisyonun biçimsel özelliklerinin genel izlenimle puanlamayı ne derece yordayabildiğini ortaya çıkarmak, (3) Kompozisyon tipi sınavlarda iki puanlama yönteminin puanlayıcı güvenilirliklerini belirlemek ve aynı puanlayıcının anahtarla ve genel izlenimle verdiği puanlar arasındaki tutarlılığı belirlemektir.

Bu bölümde amaçlar doğrultusunda elde edilen bulgulara bağlantılı olarak sonuçlara ve daha sonra yapılacak araştırmalara yön göstermesi açısından önerilere yer verilmiştir.

SONUÇLAR

- 1) Kompozisyon tipi sınavlarda kompozisyonun biçimsel özelliklerinden anahtarla puanlamanın ne derece yordayabildiğini belirlemek için yapılan regresyon analizi sonucunda elde edilen denklem önemli bulunmuştur. Buradan kompozisyonun biçimsel özelliklerinden anahtarla puanlamanın kestirebileceği sonucuna varılmıştır.
- 2) Kompozisyonun biçimsel özelliklerine ilişkin değişkenlerin anahtarla puanlamadaki değişkenliğin % 52'sini açıklayabileceği sonucuna varılmıştır.
- 3) Kompozisyonun anahtarla puanlanmasının en iyi yordayıcıları olarak “*Yazım Hatası Sayısı*”, “*Ortalama Cümle Uzunluğu*”, “*Yayınlanmış Makalelerden Çıkarılmış ve Ağırlıklandırılmış Anahtar Kelimeler*”, “*Başlığın Var Olup Olmaması*” ve “*Basit Cümle Sayısı*” olabileceği belirlenmiştir.

- 4) Kompozisyon tipi sınavlarda kompozisyonun biçimsel özelliklerinden genel izlenimle puanlamanın ne derece yordanabildiğini belirlemek için yapılan regresyon analizi sonucunda elde edilen denklem önemli bulunmuştur. Buradan kompozisyonun biçimsel özelliklerinden genel izlenimle puanlamanın kestirebileceği sonucuna varılmıştır.
- 5) Kompozisyonun biçimsel özelliklerine ilişkin değişkenlerin genel izlenimle puanlamadaki değişkenliğin % 60'ını açıklayabileceği sonucuna varılmıştır.
- 6) Kompozisyonun genel izlenimle puanlanmasının en iyi yordayıcıları olarak *“Başlığın Var Olup Olmaması”*, *“Örnek Verirken Kullanılan Kelimelerin Varlığı”*, *“Kompozisyon Puanlandıktan Sonra Çıkarılan Anahtar Kelimeler”*, *“Soru Cümlesi Sayısı”* ve *“Yayınlanmış Makalelerden Çıkarılmış ve Ağırlıklandırılmış Anahtar Kelimeler”* olabileceği belirlenmiştir.
- 7) Anahtarla puanlamada puanlayıcı güvenirliliği, genel izlenimle puanlamanın puanlayıcı güvenirliliğinden daha düşük bulunmuştur. Ama bunun iki tane puanlayıcı kullanıldığından kaynaklanabileceği sonucuna varılmıştır.
- 8) Kompozisyonları puanlayan uzmanların her ne kadar alanda uzman olsa da puanlama konusunda yeterlilikleri olmadığı anahtarla puanlamada puanlamanın güvenirliliğinin düşük olduğundan çıkarılabilir.

ÖNERİLER

- 1) Bu araştırmada biçimsel özelliklerin belirlenmesinde bir bilgisayar programı geliştirilerek yapılması düşünülmüş fakat imkanların yeterli olmamasından dolayı bu amaçların dışında tutulmuştur. Bu konuda yapılacak yeni bir

çalışmada biçimsel özelliklerin sayılarını belirlemede kullanılacak bir bilgisayar programı geliştirilebilir.

- 2) Regresyon analizine katılan biçimsel özelliklere ait veriler artırılabilir. Cümlecik türlerinden farklı cümlecik türleri (isim cümlecikleri, zarf cümlecikleri vb.) ve cümlelerin farklı öğeleri de (zamir sayıları, ilgi zamirleri, sıfat tamlamaları, isim tamlamaları vb.) belirlenerek analize katılabilir.
- 3) Genel izlenimle puanlama, anahtarla puanlamaya göre puanlayıcılar arası tutarlılık anlamındaki güvenirlilik katsayısında daha olumlu sonuçlar vermiştir. Ancak, araştırmanın sınırlılığı çerçevesinde 2 puanlayıcı kullanılması bu sonuca ulaşılmasında büyük bir etken olabilir. Bu nedenle fazla puanlayıcı kullanılan bir araştırma desenlenmesi önerilmektedir.
- 4) Kompozisyonları puanlayan uzmanların anahtarla puanlamada puanlamanın puanlayıcı güvenirliliğini artırmak için anahtar kullanma ve puanlama konusunda ön bir eğitim verilerek bir araştırma yapılması önerilmektedir.
- 5) Bu çalışmada iki farklı puanlama metodu kullanılmıştır. Bu puanlama türlerine bütünsel (holistic) puanlama metodu ve hata sayma metodu da eklenerek dört puanlama metodunun karşılaştırıldığı bir araştırma desenlenmesi önerilmektedir.

KAYNAKÇA

- BAYKUL, Yaşar. *Eğitimde Ve Psikolojide Ölçme*. ÖSYM yayınları. Birinci baskı. Ankara, 2000.
- BURSTEIN, J. ve D. MARCU. "Toward Using Text Summarization for Essay-Based Feedback." *Conférence TALN 2000, Lausanne, 16-18 octobre 2000*. 20 Eylül 2001 <<http://www.ets.org/research/talnsum.doc>>.
- BURSTEIN, J., R. KAPLAN, S. WOLFF ve C. LU (1996). "Using Lexical Semantic Techniques to Classify Free-Responses." 20 Eylül 2001 <<http://www.ets.org/research/bursteinfin.pdf>>.
- BURSTEIN, J., K. KUKICH, S. WOLFF, C. LU, ve M. CHODOROW (1998b). "Enriching Automated Essay Scoring Using Discourse Marking." 20 Eylül 2001 <<http://www.ets.org/research/dscrfinl.pdf>>.
- BURSTEIN, J., K. KUKICH, S. WOLFF, C. LU, M. CHODOROW, L. B. HARDER, ve M. D. HARRIS (1998c). "Automated Scoring Using A Hybrid Feature Identification Technique." 20 Eylül 2001 <<http://www.ets.org/research/aclfinal.pdf>>.
- BURSTEIN, Jill; Martin CHODOROW (1999). "Automated Essay Scoring for Nonnative English Speakers." 20 Eylül 2001 <<http://www.ets.org/research/acl99rev.pdf>>.
- BURSTEIN, J., K. KUKICH, S. WOLFF, C. LU, ve M. CHODOROW (1998a). "Computer Analysis of Essays." NCME Symposium on Automated Scoring 20 Eylül 2001 <<http://www.ets.org/research/ncmefinal.pdf>>.
- CLAUSER, Brian E. "Recurrent Issues and Recent Advances in Scoring Performance Assessments." *Applied Psychological Measurement*, Vol. 24 No:4, 2000:310-324.

- EBEL, Robert L. *Measuring Educational Achievement*. Sekizinci Baskı. Englewood Cliffs, N. Jersey: Prentice Hall inc. 1965.
- ENGİNARLAR, Hüsnü. "A Quantitative And Qualitative Comparison of Three Techniques of Grading ESL/EFL Essays." *Journal of Human Sciences*, X, sayı 1, 1991: 23-45.
- ERICKSON, Jon Dale. "Using Keywords and Computers to Assess Student Writing." (Yayımlanmamış Doktora Tezi), Washington State University, 2000.
- GRONLUND, Norman E. *Measurement and Evaluation in Teaching*. The Macmillan Company. Newyork, 1967.
- KARASAR, Niyazi. *Bilimsel Araştırma Yöntemi*. Sanem Matbaacılık. Ankara , 1991.
- KARATAYLI, Tuba. "Comparison of Two Essay Scoring Methods A Case Study." (Yayımlanmamış Yüksek Lisans Tezi), Ankara: Ortadoğu Teknik Üniversitesi, 1996.
- MICHEELS, W. J. ve M. R. KARNES. *Eğitimde Başarının Ölçülmesi*, (1950) (Çeviren İbrahim Yurt), Ankara: Ajans-Türk Matbaası, 1968.
- MILLER, M. D. ve R. L.LINN. "Validation of Performance-Based Assessments." *Applied Psychological Measurement*, Vol. 24 No:4, 2000.
- TRAVERS, Robert M. W. *Educational Measurement*. The Macmillan Company. Newyork, 1955.
- TURGUT, M. Fuat. *Eğitimde Ölçme ve Değerlendirme*. Onuncu baskı, Yargıcı Matbaası, Ankara, 1997.
- ÜNVER, Özkan ve Hamza GAMGAM *Uygulamalı İstatistik Yöntemler*. Ankara, 1986.

WHITTINGTON, D. ve H. HUNT. "Approaches to the Computerized Assessment of Free Text Responses" in Proceedings of Third Annual Computer Assisted Assessment Conference, University of Loughborough, 1999. 19 Aralık 2001 <<http://cvu.strath.ac.uk/dave/publications/caa99.html>>

WRIGHTSTONE, J. W., J. JUSTMAN ve I. ROBBINS. *Evaluation in Modern Education*. American Book Company. Newyork, 1956.

