

Orijinal Makale Başlığı:

İlkokul 2. sınıf İngilizce dersine yönelik dinleme ve konuşma başarı testinin geliştirilmesi

Makalenin İngilizce Başlığı:

Development of a listening and speaking achievement test for primary school 2nd grade English course

Yazar(lar):

Fatma ÖZÜDOĞRU, Oktay Cem ADIGÜZEL

Kaynak Gösterimi İçin:

Özüdoğru, F., & Adıgüzel, O.C. (2015). İlkokul 2. sınıf İngilizce dersine yönelik dinleme ve konuşma başarı testinin geliştirilmesi. *Pegem Eğitim ve Öğretim Dergisi*, 5(4), 375-396, <http://dx.doi.org/10.14527/pegegog.2015.021>.

Original Title of Article:

İlkokul 2. sınıf İngilizce dersine yönelik dinleme ve konuşma başarı testinin geliştirilmesi

English Title of Article:

Development of a listening and speaking achievement test for primary school 2nd grade English course

Author(s):

Fatma ÖZÜDOĞRU, Oktay Cem ADIGÜZEL

For Cite in:

Özüdoğru, F., & Adıgüzel, O.C. (2015) İlkokul 2. sınıf İngilizce dersine yönelik dinleme ve konuşma başarı testinin geliştirilmesi. *Pegem Eğitim ve Öğretim Dergisi*, 5(4), 375-396, <http://dx.doi.org/10.14527/pegegog.2015.021>.

İlkokul 2. Sınıf İngilizce Dersine Yönelik Dinleme ve Konuşma Başarı Testinin Geliştirilmesi*

Fatma ÖZÜDOĞRU^{a†}, Oktay Cem ADIGÜZEL^b

^aUşak Üniversitesi, Yabancı Diller Yüksekokulu, Uşak/Türkiye

^bAnadolu Üniversitesi, Eğitim Fakültesi, Eskişehir/Türkiye



Makale Bilgisi

DOI: 10.14527/pegegog.2015.021

Makale Geçmişi:

Geliş 04 Haziran 2014
Düzeltilme 05 Mayıs 2015
Kabul 18 Ağustos 2015

Anahtar Kelimeler:

Dinleme;
Konuşma;
Başarı testi geliştirme.

Öz

Bu çalışmada ilkökul 2. sınıf İngilizce öğretim programına devam etmekte olan öğrencilerde öngörülen amaçların gerçekleştirme durumunu ortaya çıkarmak amacıyla dinleme ve konuşma becerilerine odaklanan başarı testi geliştirilmesi amaçlanmıştır. Hazırlanan taslak başarı testinin geçerliliğini sağlamak için üç yabancı dil alan uzmanının ve üç 2. sınıf İngilizce öğretmeninin görüşleri alınarak taslak başarı testinde yer alacak madde sayısı belirlenmiştir ve sonrasında madde ayırt edicilik ve madde güçlük analizleri hesaplanmıştır. Hazırlanan taslak başarı testinin dinleme bölümü, 2013-2014 eğitim öğretim yılı bahar döneminde 202, başarı testinin konuşma bölümü ise 125 ilkökul 2. sınıf öğrencisine uygulanmıştır. Dinleme bölümünde analizler sonucunda 18 madde çıkarılmıştır. Kalan 22 maddeden puanlama kolaylığı sağlama açısından maddelerin güçlük dağılımları arasındaki denge de gözeticilerle 2 kolay madde daha çıkarılmış nihai testte 20 madde yer almıştır. Konuşma bölümünde ise sadece madde güçlük analizi sonrasında 4 madde çıkarılmış ve puanlama kolaylığı açısından bir maddenin ölçme aracından çıkarılmasına karar verilmiştir ve nihai testte 20 madde yer almıştır. Testin iç tutarlılığı için KR-20 değerine bakılmış ve dinleme testinin .70, konuşma testinin .81 olarak yeterli bir güvenilirliğe sahip olduğu sonucuna varılmıştır.

Development of a Listening and Speaking Achievement Test for Primary School 2nd Grade English Course

Article Info

DOI: 10.14527/pegegog.2015.021

Article history:

Received 04 June 2014
Revised 05 May 2015
Accepted 18 August 2015

Keywords:

Listening;
Speaking;
Achievement test development.

Abstract

In this research, it is aimed to develop an achievement test focusing on listening and speaking skills in order to reveal the realization of the objectives set forth for students studying primary school 2nd grade English teaching curriculum. For validity purposes of the draft achievement test, three foreign language field experts and three 2nd grade English teachers were consulted and the number of items in the achievement test was determined and then item discrimination and item difficulty analyses were carried out. The listening part of the achievement test was applied to 202 students and the speaking part was applied to 125 2nd grade students in the spring semester of 2013-2014 academic years. The final listening test included 20 items after 18 items were excluded after analysis and 2 easy items were also excluded for grading easiness. For the speaking part, only 4 items were left out after item difficulty analysis and 1 item was excluded for grading easiness, so 20 items were included in the final speaking test. KR-20 value of the test was examined for the internal consistency and it was found that the listening test had .70 reliability and the speaking test had .81 reliability, which are adequate.

* Bu çalışma Anadolu Üniversitesi Bilimsel Araştırma Projeleri Birimi tarafından (Proje No: 1303E046) desteklenmiştir.

† Yazar: fatma.ozudogru@usak.edu.tr

Giriş

Erken yaşta yabancı dil eğitimi konusu dil eğitiminde önemli tartışma konularından birini oluşturmaktadır. Yabancı dil eğitimine başlama yaşı konusunda dilbilimcilerin yaklaşımları incelendiğinde yaklaşımlar arası farklılıklar dikkati çekmektedir. Çayıroğlu-İşgören (2005), bu yaklaşımları şu üç başlık altında toplamaktadır: Yabancı Dil Eğitiminin Erken Yaşta Başlaması Gerektiği Konusundaki Yaklaşımlar, Yabancı Dil Eğitiminin İyi Bir Anadili Eğitiminin Sonrasında Başlaması Gerektiği Konusundaki Yaklaşımlar ve Yabancı Dil Eğitiminin Her Yaşta Başlayabileceği Konusundaki Yaklaşımlar.

Avrupa Birliği (AB) ülkelerinde yabancı dil öğretimi gittikçe daha erken yaşta başlamaktadır. Türkiye'deki yabancı dil alanında program geliştirme çalışmaları incelendiğinde ise, AB ülkelerinde olduğu gibi daha erken yaşta dil öğretimine başlama eğilimi göze çarpmaktadır. 1997 yılında, ilköğretimin sekiz yıla çıkmasıyla, ilk defa ilköğretim 4. ve 5. sınıflarda İngilizce dersi zorunlu olarak okutulmaya başlanmış ve bu uygulama ile yabancı dil öğrenimine 10 yaşında başlanmasına olanak sağlanmıştır (Haznedar, 2010). Türkiye'de Milli Eğitim Bakanlığı (MEB) tarafından alınan yeni alınan karara göre ise 2012-2013 öğretim yılından itibaren uygulanmakta olan 8 yıllık kesintisiz zorunlu eğitim yerine, 12 yıllık zorunlu kademeli eğitim getirilmiş ve yabancı dil öğretiminin ilkökul 2. sınıftan itibaren başlamasına karar verilmiştir (MEB, 2012). Yeni sistemde İngilizce eğitiminin 4. sınıf yerine 2. sınıfta başlaması 2. ve 3. sınıflara yer veren yeni bir programın hazırlanmasını gerekli kılmıştır. Bu nedenle, genç öğrenenlerin ihtiyaçlarını dikkate alan yeni bir program ihtiyacı ortaya çıkmış ve bu doğrultuda programlar yenilenmiştir.

Mirici (2001), çocuklarla yetişkinlerin öğrenme-öğretme süreci içerisinde sahip oldukları hazır bulunuşluk düzeyleri, öğretme etkinliklerindeki tepkileri, öğrenme durumlarındaki motivasyonları vb. birbirinden farklı olduğu için her iki gruba yabancı dil öğretmek için izlenecek yolun da birbirinden farklı olması gerektiğini savunmaktadır. Bu nedenle, hazırlanan yeni programların çocukların öğrenme özelliklerine uygun kazanımlar, içerik, öğretim yöntem ve teknikleri ile ölçme ve değerlendirme tekniklerine sahip olması gerekmektedir.

Hazırlanan ilkökul 2. sınıf İngilizce öğretim programı 2013-2014 öğretim yılından itibaren uygulanmaktadır. ilkökul 2. sınıf İngilizce öğretim programı incelendiğinde erken yaşlardan itibaren dil öğrenmeye karşı olumlu tutum geliştirmenin programın temel amacı olduğu göze çarpmaktadır. İlkokul 2. sınıf İngilizce öğretim programında dinleme ve konuşma olarak iki beceri ağırlıklı olarak hedeflenmektedir (MEB, 2013).

İlkokul 2. sınıf İngilizce öğretim programının içeriğinde önceki konu ve etkinliklerin sık sık hatırlatılarak öğrencilerin bilgilerinin canlı tutulması esastır. Bunun yanında, ilkökul 2. sınıf İngilizce öğretim programı genç öğrenenlerin öğrenme süreci boyunca kendilerini rahat ve desteklenmiş hissetmelerini sağlayacak eğlenceli ve motive edici öğrenme ortamı öngörmektedir. Ayrıca, programda dili öğrenen/kullanıcıların eğlenceli görsel-işitsel araçlarla zenginleştirilmiş oyun temelli etkinlikler aracılığıyla öğrenmeleri planlanmıştır.

Hazırlanan bir öğretim programında öğrencilerin akademik gelişimlerinin programdan nasıl etkilendiğinin belirlenmesi de gerekmektedir (Parlett & Hamilton, 1977). Bu nedenle, ölçme ve değerlendirmenin tüm sınıf seviyelerinde etkili bir şekilde uygulanması önem taşımaktadır. İlkokul 2. sınıf İngilizce öğretim programının ölçme ve değerlendirme boyutunda ise proje ve portfolyo değerlendirme, öz ve akran değerlendirme ve öğretmen gözlemleri ve değerlendirmeleri önerilmektedir. Bunun yanında, öğrenenlerin başarısını nesnel değerlendirmek amacıyla dinleme ve konuşma becerilerini ölçen ölçme ve değerlendirme yöntemleri önerilmektedir (MEB, 2013). Bu nedenle, bu araştırmada ilkökul 2. sınıf İngilizce öğretim programına devam etmekte olan öğrencilerde öngörülen amaçların gerçekleşme durumunu ortaya çıkarmak amacıyla konuşma ve dinleme becerilerine odaklanan başarı testi geliştirilmiştir. İlkokul 2. sınıf İngilizce öğretim programında dinleme ve konuşma becerileri hedeflendiği için başarı testinde dinleme ve konuşma becerilerini ölçen maddeler yer almaktadır.

Yöntem

Araştırma Modeli

“Tarama modelleri, geçmişte ya da halen varolan bir durumu var olduğu şekliyle betimlemeyi amaçlayan araştırma yaklaşımlarıdır” (Karasar, 2012; s.77). Bu çalışmada, tarama modeli kullanılmıştır. Bir başarı testi geliştirilmiş olup, bu test yardımıyla mevcut durum betimlenebilir.

Katılımcılar

Araştırmanın evreni ilkökul 2. sınıf öğrencileridir. Hazırlanan taslak dinleme testinin güvenilirlik analizi için katılımcı grubunu 2013-2014 öğretim yılı bahar döneminde üç ilkökulda sekiz farklı sınıfta öğrenim gören 202 ikinci sınıf öğrencisi oluşturmuştur. Öğrencilerin 108’i kız, 94’ü erkek öğrencidir. Taslak konuşma testi ise üç ilkökulda sekiz farklı sınıfta öğrenim gören 125 ikinci sınıf öğrencisine uygulanmıştır. Bu öğrencilerin 69’u kız, 56’sı erkektir.

Birinci Çalışma: Dinleme ve Konuşma Başarı Testinin Hazırlanması ve Geçerlik Çalışması

Araştırma kapsamında öğrencilerde programdaki dinleme ve konuşma becerisi kazanımlarının gerçekleşme durumunu ortaya koymada program kapsamında yer alan üniteler incelenmiş ve ölçülebilir nitelikteki kazanımlar belirlenmiştir. Belirlenen ölçülebilir nitelikteki dinleme becerisi kazanımları Tablo 1’de sunulmuştur:

Tablo 1.

İlkökul 2. Sınıf İngilizce Öğretim Programı Dinleme Becerisi Kazanımları.

Üniteler	Dinleme Becerisi Kazanımları
1. Ünite Words	1. Üretilen seslerle yazılı harfleri eşleştirebilir.
3. Ünite In the Classroom	2. Sınıfta kullanılan kısa ve basit yönergeleri anlayabilir.
4. Ünite Numbers	3. 1’den 10’a kadar rakamları tanıyabilir.
5. Ünite Colors	4. Renkleri anlayabilir.
6. Ünite At the Playground	5. Oyun alanı ile ilgili kısa ve basit yönergeleri takip edebilir.
7. Ünite Body Parts	6. Vücudun bölümlerini anlayabilir.
8. Ünite Pets	7. Evcil hayvanları tanıyabilir.
9. Ünite Fruits	8. Meyve isimlerini tanıyabilir.
10. Ünite Animals	9. Hayvan isimlerini tanıyabilir.

Tablo 1’de görüldüğü gibi, ilkökul 2. sınıf İngilizce öğretim programında dinleme becerisine yönelik ölçülebilir dokuz kazanım yer almaktadır. Testin dinleme bölümü kapsamında her kazanıma ilişkin en az üç soru hazırlanmıştır. Dinleme soruları hazırlanırken öğrencilerin yaş ve gelişim düzeyleri dikkate alınmıştır. Bu bağlamda resimlerden, renkli kartlardan yararlanılmıştır.

Dinleme sınavının kapsam geçerliği için dinleme bölümü uzman görüşüne sunulmuştur. Bu bağlamda, üç İngilizce öğretimi alan uzmanı ve üç 2. sınıf İngilizce öğretmeninin görüşleri ve dönütleri sonrasında gerekli düzeltmeler yapılmıştır. Taslak dinleme testinde yer alacak madde sayısı 40 olarak belirlenmiştir. Geliştirilmesi hedeflenen dinleme testi, hem dinleme becerisinin doğası gereği hem de öğrencilerin içinde buldukları gelişim özelliklerinden dolayı, Bloom’un eğitimin hedefleri ile ilgili aşamalı sınıflamasının uygulama, analiz, sentez, değerlendirme düzeylerinde sorular üretilmesi mümkün olamamıştır. Dinleme testi için hazırlanan belirtke tablosu Tablo 2’de verilmiştir.

Tablo 2.
Dinleme Testi Belirtke Tablosu.

	Kavrama
Words	6
In the Classroom	4
Numbers	5
Colors	5
At the Playground	4
Body Parts	4
Pets	4
Fruits	4
Animals	4
Toplam	40

Başarı testinin pilot uygulaması öncesinde testte yer alan maddelerin anlaşılabilirliğini ve düzeye uygunluğunu saptamak amacıyla dinleme kayıtları ilkökul 2. sınıfta öğrenim gören beş öğrenciye dinletilmiş ve alınan dönütler çerçevesinde gerekli değişiklikler ve düzeltmeler yapılmıştır. Dinleme kayıtlarını anadili İngilizce olan bir İngilizce öğretmeni okumuştur. Program kapsamında ölçülebilir nitelikteki konuşma becerileri kazanımları Tablo 3'te gösterilmiştir:

Tablo 3.
İlkokul 2. Sınıf İngilizce Öğretim Programı Konuşma Becerisi Kazanımları.

Üniteler	Konuşma Becerisi Kazanımları
1. Ünite Words	1. Nesnelere, insanlar ya da hayvanlar için doğru kelimeyi kullanabilir.
2. Ünite Friends	2. Kendini tanıtabilir ve selamlaşabilir.
4. Ünite Numbers	3. Rakamları ve miktarları söyleyebilir.
5. Ünite Colors	4. Nesnelere renklerini söyleyebilir.
7. Ünite Body Parts	5. Vücudun bölümlerini gösterip söyleyebilir.
8. Ünite Pets	6. Evcil hayvanların isimlerini söyleyebilir.
9. Ünite Fruits	7. Meyvelerin isimlerini söyleyebilir.
10. Ünite Animals	8. Hayvanların isimlerini söyleyebilir.

Tablo 3'te görüldüğü gibi, ilkökul 2. sınıf İngilizce öğretim programında konuşma becerisine yönelik ölçülebilir sekiz kazanım yer almaktadır. Testin konuşma bölümü kapsamında her kazanıma ilişkin en az iki soru hazırlanmıştır. Konuşma soruları hazırlanırken öğrencilerin yaş ve gelişim düzeyleri dikkate alınmıştır. Bu bağlamda resimlerden, renkli kartlardan yararlanılmıştır.

Konuşma sınavının kapsam geçerliği için konuşma bölümü uzman görüşüne sunulmuştur. Bu bağlamda, üç İngilizce öğretimi alan uzman ve üç 2. sınıf İngilizce öğretmenin görüşleri ve dönütleri sonrasında gerekli düzeltmeler yapılmıştır. Taslak konuşma testinde yer alacak madde sayısı 25 olarak belirlenmiştir. Geliştirilmesi hedeflenen konuşma testinde öğrencilere daha çok kavrama ve uygulama düzeyinde sorular sorulmuştur. Öğrencilerin yaş özelliklerinden dolayı analiz, sentez, değerlendirme düzeylerinde sorular üretilmesi mümkün olamamıştır. Konuşma testi için hazırlanan belirtke tablosu Tablo 4'te sunulmuştur:

Tablo 4.
Konuşma Testi Belirtke Tablosu.

	Kavrama	Uygulama	Toplam
Words	6	-	6
Friends	-	2	2
Numbers	3	-	3
Colors	3	-	3
Body Parts	4	-	4
Pets	3	-	3
Fruits	2	-	2
Animals	2	-	2
Toplam	23	2	25

Konuşma bölümünün pilot uygulaması öncesinde testte yer alan maddelerin anlaşılabilirliğini ve düzeye uygunluğunu saptamak amacıyla ilkökul 2. sınıfta öğrenim gören beş öğrenciye sorulmuş ve alınan dönütler çerçevesinde gerekli değişiklikler ve düzeltmeler yapılmıştır.

İkinci Çalışma: Dinleme ve Konuşma Başarı Testinin Güvenirlik Analizi Uygulama ve Veri Analiz Sonuçları

Dinleme ve konuşma bölümlerinden oluşan taslak başarı testinin dinleme bölümü 2013-2014 bahar döneminde ilkökul 2. sınıf İngilizce öğretim programının tamamlandığı Mayıs ayında üç ilkökulda sekiz farklı sınıfta öğrenim gören 202 ilkökul ikinci sınıf öğrencisine, konuşma bölümü ise 125 ilkökul ikinci sınıf öğrencisine uygulanmış ve madde analizleri yapılmıştır. Madde analizleri kapsamında her maddenin ayırt edicilik ve madde güçlük indeksi hesaplanmıştır.

Madde ayırt edicilik indeksi (D), yoklanan davranışı öğrenmiş olan cevaplayıcıları bu davranışı öğrenmemiş olanlardan ayırma gücünü göstermektedir (Özçelik, 2013). Test geliştirmede, maddelerin, o madde ile ölçülen niteliğe sahip olanlarla olmayanları ayırmaları istenmektedir. Bu, ölçülmek istenen niteliğe sahip olanların maddeyi doğru cevaplamaları, olmayanların yanlış cevaplamaları ile eş anlamlıdır. Maddenin, ölçülen nitelik bakımından bireyleri ayırmasının bir ölçüsü olarak madde ayırt edicilik indeksi kullanılmaktadır (Baykul, 2010). Madde ayırt edicilik indeksi -1 ile +1 arasında değişebilir. Bu değer negatif olması, maddenin ölçülen özellik bakımından bireyleri ters ayırt ettiğini gösterir. Bu nedenle, bu tür maddeler testten çıkarılmalıdır (Büyüköztürk, Kılıç Çakmak, Akgün, Karadeniz, & Demirel, 2013).

Madde ayırt edicilik değeri, testin toplam puanlarına göre oluşturulan alt %27 ve üst %27'lik grupların madde ortalama puanları arasındaki farkların bağımsız t-testi kullanılarak test edilmesiyile bulunabilir. Bu bağlamda öncelikle testte yer alan maddelerin ayırt edicilik düzeylerine bakmak üzere alt ve üst %27'lik gruplar için bağımsız t-testi uygulanmıştır. Öğrencilerin başarı testinden almış oldukları puanlar büyükten küçüğe doğru sıralanmış ve alt ve üst %27'ye göre dinleme testinde 55 alt ve 55 üst grup, konuşma testinde alt ve üst %27'ye göre 34 alt ve 34 üst grup olmak üzere iki kategori oluşturulmuştur. Alt ve üst grubu oluşturan öğrencilerin vermiş oldukları cevaplara göre başarı testindeki her maddenin ayırt edicilik indeksi hesaplanmıştır. Statistical Package for Social Sciences (SPSS) veri çözümleme programı ile veriler incelenmiştir. Madde ayırt edicilik analizi aşamasında indeks değerleri 0,20 altında olan maddeler düşük ayırt edici olduğu kabul edilmiş ve testten çıkarılmıştır (Özçelik, 2013; Turgut & Baykul, 2012). Ayırt edicilik indeksi şu şekilde hesaplanabilir (Turgut & Baykul, 2012):

Üst grupta doğru cevaplayanların sayısı-Alt grupta doğru cevaplayanların sayısı

Grubun %27 si

Madde güçlük indeksi (P), bir maddeye doğru cevap verenlerin sayısının testi alanların sayısına oranıdır. Madde güçlük indeksi 0 ile 1 arasında değerler alır. Güçlük indeksinin 1'e yaklaşması maddenin kolaylaştığını, 0'a yaklaşması maddenin zorlaştığını göstermektedir (Turgut & Baykul, 2012). Başarı testi için bu genişliğin 0.20 ile 0.80 arasında olmasına dikkat edilmektedir (Özçelik, 1992). Madde güçlük indeksi aşağıdaki gibi hesaplanabilir (Turgut & Baykul, 2012):

Üst grupta doğru cevaplayanların sayısı+Alt grupta doğru cevaplayanların sayısı

Grubun %54'ü

Bu bağlamda, madde güçlük indeksi için frekans parametresinden maddelere doğru ve yanlış cevap verenlerin sayısı ve maddelerin doğru cevaplandırılma yüzdesi incelenmiştir. Analiz aşamasında madde güçlük indeksi (p) değerinin 0.20 ve 0.80 arasında olmasına dikkat edilmiştir. Zorluk derecesi 0.20'nin altında olan maddeler çok zor, 0.80'in üstünde olanlar ise çok kolay soru olarak nitelendirilmiş ve bu sorular başarı testine dâhil edilmemişlerdir. Madde güçlük indeksleri sınıflandırılırlarken: $0.20 \leq p \leq 0.40$ arası maddeler zor; $0.41 \leq p \leq 0.60$ arası maddeler orta; $0.61 \leq p \leq 0.80$ arası maddeler kolay olarak değerlendirilmiştir (Kutlu, 2004).

Bulgular

Dinleme Testine İlişkin Bulgular

Dinleme testinin madde ayırt ediciliğine ilişkin gerçekleştirilen alt ve üst %27'lik gruplar için bağımsız t-testi sonuçları Tablo 5'te gösterilmiştir.

Tablo 5.
T-Testine İlişkin Sonuçlar.

Madde No	Varyansların Eşitliği İçin Levene Testi		Varyansların Eşitliği İçin t-testi		Madde No	Varyansların Eşitliği İçin Levene Testi		Varyansların Eşitliği İçin t-testi	
	F	Sig.	T	Sig.		F	Sig.	T	Sig.
M.1	11.37	.00	2.37	.01	M.21	9.85	.00	1.63	.10
M.2	31.76	.00	2.97	.00	M.22	.00	1	2.52	.01
M.3	11.44	.00	1.62	.10	M.23	11.92	.00	1.72	.08
M.4	16.01	.00	-1.92	.05	M.24	48.08	.00	3.51	.00
M.5	1.23	.26	1.92	.05	M.25	16.45	.00	3.02	.00
M.6	.50	.48	.94	.34	M.26	3.32	.07	.97	.33
M.7	109.07	.00	5.09	.00	M.27	1.33	.25	2.74	.00
M.8	13.72	.00	2.02	.04	M.28	18.82	.00	4.15	.00
M.9	.13	.71	-.18	.85	M.29	12.61	.00	1.68	.09
M.10	3.70	.05	3.17	.00	M.30	66.02	.00	4.89	.00
M.11	82.37	.00	6.58	.00	M.31	228.38	.00	5.09	.00
M.12	44.51	.00	3.86	.00	M.32	15.27	.00	1.84	.06
M.13	18.37	.00	2.06	.04	M.33	85.59	.00	3.60	.00
M.14	63.44	.00	4.01	.00	M.34	175.92	.00	4.65	.00
M.15	11.85	.00	3.65	.00	M.35	12.34	.00	2	.04
M.16	5.95	.01	1.23	.21	M.36	2.15	.14	.77	.44
M.17	44.32	.00	3.05	.00	M.37	63.44	.00	4.01	.00
M.18	35.63	.00	2.89	.00	M.38	6.98	.00	1.31	.19
M.19	41.64	.00	3.44	.00	M.39	16.31	.00	2.08	.03
M.20	20.05	.00	3.26	.00	M.40	202.55	.00	6.33	.00

Bağımsız iki grubun karşılaştırıldığı t-testinin hesaplanmasında gruplar arasında varyans farkının olup olmadığına dikkat edilmesi gerekmektedir. Bu sebeple, bu çalışmada t-testi hesaplamasında öncelikle varyansların eşit olup olmadığına test edilmesine gerek duyulmuştur. Varyansların eşit olup olmadığını test eden testlerden biri de Levene testidir (Büyüköztürk, 2012). Levene testi anlamlılık düzeyi 0.05 değerinin altındaysa buradan iki grubun varyanslarının farklı olduğu anlaşılmaktadır. Farklı varyanslar olduğunda SPSS veri sayfasında “Equal variances not assumed (Varyansların eşitliğinin sağlanmadığı)” satırındaki değer dikkate alınmıştır.

Tablo 5’ten t değerine karşılık gelen anlamlılık değerleri incelendiğinde toplam 13 maddenin yeterince ayırt etmediği veya ters ayırt ettiği tespit edilmiştir. Bu nedenle M.3, M.4, M.5, M.6, M.9, M.16, M.21, M.23, M.26, M.29, M.32, M.36, M.38, taslak başarı testinden çıkarılmışlardır.

t testi değerlerine göre 13 soru maddesinin test içerisinden çıkarılması ile 40 soruluk test 27 soruluk bir teste dönüşmüştür. Madde ayırt edicilik indeksleri her soruya üst grupta doğru yanıt verenlerin sayısından alt grupta doğru cevap verenlerin sayısı çıkarılıp grubun %27’sine (55) bölünerek hesaplanmıştır. Madde ayırt edicilik indekslerinin yanında, madde güçlük indeksleri Tablo 6’da sunulmuştur.

Tablo 6.
Madde Ayırt Edicilik ve Güçlük Değerlerine İlişkin Bulgular.

Madde No	t	P	Ayırt Edicilik İndeksi	Madde Güçlük İndeksi	Madde No	T	p	Ayırt Edicilik İndeksi	Madde Güçlük İndeksi
M.1	2.37	.01	0.22	0.40	M.22	2.52	.01	0.24	0.50
M.2	2.97	.00	0.25	0.69	M.24	3.51	.00	0.29	0.71
M.7	5.09	.00	0.40	0.71	M.25	3.02	.00	0.27	0.39
M.8	2.02	.04	0.18	0.65	M.27	2.74	.00	0.25	0.47
M.10	3.17	.00	0.29	0.55	M.28	4.15	.00	0.36	0.60
M.11	6.58	.00	0.51	0.65	M.30	4.89	.00	0.40	0.67
M.12	3.86	.00	0.33	0.33	M.31	5.09	.00	0.36	0.78
M.13	2.06	.04	0.16	0.77	M.33	3.60	.00	0.22	0.87
M.14	4.01	.00	0.33	0.71	M.34	4.65	.00	0.33	0.80
M.15	3.65	.00	0.33	0.58	M.35	2.00	.04	0.18	0.64
M.17	3.05	.00	0.24	0.77	M.37	4.01	.00	0.33	0.71
M.18	2.89	.00	0.24	0.74	M.39	2.08	.03	0.18	0.69
M.19	3.44	.00	0.29	0.69	M.40	6.33	.00	0.47	0.71
M.20	3.26	.00	0.29	0.38					

Tablo 6’dan dinleme testinde kalan 27 soruya yönelik madde ayırt edicilik indeksleri incelendiğinde 8., 13., 35. ve 39. maddelerin t değerleri anlamlı olsa dahi madde ayırt edicilik indekslerinin .20’nin altında olduğu görülmüş ve testten çıkarılmışlardır. Tablo 6’dan dinleme testinde yer alan maddelerin madde güçlük indeksleri hesapları yapıldığında 33. maddenin maddelerin güçlük indeksinin 0.87 olduğu görülmüştür. Bu soru çok kolay soru sınıflamasına girdiğinden dolayı ($p \leq 0.80$) bu madde de başarı testinden çıkarılmıştır.

Madde ayırt edicilik indeksine göre 4 maddenin çıkması ve madde güçlük indeksine göre 1 maddenin çıkmasıyla dinleme testinde 22 madde yer almıştır. Kalan 22 maddenin 13 maddesi kolay, 5 maddesi orta ve 4 maddesi ise zor madde olarak sınıflandırılmışlardır. 22 maddenin 100’lük puanlama sisteminde tam puana karşılık gelmemesinin testin kullanılabilirliğini olumsuz etkileyebileceği nedeniyle maddelerin güçlük dağılımları arasındaki denge de gözetilerek kolay sorulardan ayırt ediciliği en düşük iki maddenin ölçme aracından çıkarılmasına karar verilmiştir. Maddeler incelendiğinde kolay soru sınıfında bulunan maddelerden 17. ve 18. madde ölçme aracından çıkarılmışlardır. Çıkan maddeler sonrasında, güçlük dağılımlarının Tablo 7’deki gibi gruplandığı görülmüştür.

Tablo 7.
Madde Güçlük Değerlerine Göre Nihai Testteki Madde Dağılımı.

Güçlük Değeri	Madde Sayısı	Madde No	Yorum
0.61≤p≤0.80	11	2.7.11.14.19.24.30.31. 34.37.40	Kolay madde
0.41≤p≤0.60	5	10.15.22.27.28	Orta madde
0.20≤p≤0.40	4	1.12.20.25	Zor madde

Tablo 7'den görüldüğü üzere nihai testte yer alan maddeler değişik zorluk seviyelerindedir ve bu nedenle farklı türdeki öğrenci seviyelerine hitap edebilecek yeterliliktedir. Testin iç tutarlılığını ölçmek için KR-20 (Kuder-Richardson) değeri hesaplanmıştır. Bilgi testlerinde maddeler doğru ve yanlış olarak iki kategoriden oluşuyorsa, iç tutarlılığı Kuder-Richardson formülleri ile hesaplanmaktadır (KR20; KR21). Bu formüllerden KR20, test maddelerin benzer güçlükte olmadığı varsayımı altında; KR21 ise test maddelerin yaklaşık olarak benzer güçlükte olduğu varsayımı altında kullanılmaktadır (Alpar, 2012; Baykul, 2010). Geliştirilen bu başarı testinde ise her güçlükten maddelere yer verildiği için KR-20 formülünün kullanılarak hesaplanmanın yapılması tercih edilmiştir. Hesaplanan güvenilirlik katsayısının 0.70 ve daha yüksek olması test puanlarının güvenilirliği için genel olarak yeterli görülmektedir (Büyüköztürk, 2012). Yapılan analiz sonucu dinleme testinin iç tutarlılık katsayısı 0.72 olarak hesaplanmıştır. Elde edilen sonuçlar, testin yeterli bir iç tutarlılığa sahip olduğunu göstermiştir.

Konuşma Testine İlişkin Bulgular

Konuşma testinin madde ayırt ediciliğine ilişkin gerçekleştirilen alt ve üst %27'lik gruplar için bağımsız t-testi sonuçları Tablo 8'de gösterilmiştir.

Tablo 8.
T-Testine İlişkin Sonuçlar.

Madde No	t	P	Ayırt Edicilik İndeksi	Madde Güçlük İndeksi	Madde No	T	p	Ayırt Edicilik İndeksi	Madde Güçlük İndeksi
M.1	15.48	.00	9.02	.00	M.14	4.81	.03	8.88	.00
M.2	29.70	.00	7.77	.00	M.15	231.91	.00	8.30	.00
M.3	55.61	.00	5.05	.00	M.16	117.02	.00	3.66	.00
M.4	6.13	.01	10.72	.00	M.17	117.02	.00	7.00	.00
M.5	10.27	.00	9.80	.00	M.18	8.69	.00	13.46	.00
M.6	69.97	.00	3.12	.00	M.19	39.07	.00	7.24	.00
M.7	218.56	.00	5.79	.00	M.20	39.07	.00	7.24	.00
M.8	41.94	.00	4.28	.00	M.21	9504	.00	6.09	.00
M.9	116.01	.00	3.44	.00	M.22	218.56	.00	4.51	.00
M.10	218.56	.00	4.51	.00	M.23	218.56	.00	5.79	.00
M.11	29.70	.00	2.36	.02	M.24	10.27	.00	9.80	.00
M.12	148.31	.00	6.56	.00	M.25	.50	.48	8.82	.00
M.13	1.60	.21	6.91	.00					

Tablo 8'den t değerine karşılık gelen anlamlılık değerleri incelendiğinde konuşma testinde yer alan tüm maddelerin yüksek başarılı ve düşük başarılı öğrencileri yeterince ayırt ettiği anlaşılmış ve madde çıkarılmasına gerek duyulmamıştır.

Madde ayırt edicilik indeksleri her soruya üst grupta doğru yanıt verenlerin sayısından alt grupta doğru cevap verenlerin sayısı çıkarılıp grubun %27'sine (34) bölünerek hesaplanmıştır. Madde ayırt edicilik indekslerinin yanında, madde güçlük indeksleri Tablo 9'da sunulmuştur.

Tablo 9.
Madde Ayırt Edicilik ve Güçlük Değerlerine İlişkin Bulgular.

Madde No	t	P	Ayırt Edicilik İndeksi	Madde Güçlük İndeksi	Madde No	T	p	Ayırt Edicilik İndeksi	Madde Güçlük İndeksi
M.1	9.02	.00	0.74	0.57	M.14	8.88	.00	0.74	0.54
M.2	7.77	.00	0.68	0.40	M.15	8.30	.00	0.68	0.34
M.3	5.05	.00	0.50	0.34	M.16	3.66	.00	0.32	0.19
M.4	10.72	.00	0.79	0.54	M.17	7	.00	0.62	0.34
M.5	9.80	.00	0.76	0.56	M.18	13.46	.00	0.85	0.46
M.6	3.12	.00	0.26	0.84	M.19	7.24	.00	0.65	0.38
M.7	5.79	.00	0.53	0.71	M.20	7.24	.00	0.65	0.62
M.8	4.28	.00	0.44	0.66	M.21	6.09	.00	0.53	0.74
M.9	3.44	.00	0.26	0.87	M.22	4.51	.00	0.41	0.76
M.10	4.51	.00	0.41	0.76	M.23	5.79	.00	0.53	0.71
M.11	2.36	.02	0.21	0.84	M.24	9.80	.00	0.76	0.56
M.12	6.56	.00	0.59	0.68	M.25	8.82	.00	0.74	0.49
M.13	6.91	.00	0.65	0.53					

Tablo 9'dan konuşma testinde yer alan 25 soruya yönelik madde ayırt edicilik indeksleri incelendiğinde tüm maddelerin ayırt edicilik indekslerinin .20'nin üstünde olduğu ve ayırt edici olduğu görülmüştür. Bu nedenle, madde ayırt edicilik indekslerine göre madde çıkarımı yapılmamıştır. Tablo 9'dan konuşma testinde yer alan maddelerin madde güçlük indeksleri hesapları yapıldığında 6., 9. ve 11. maddelerin güçlük indekslerinin sırasıyla 0.84, 0.87 ve 0.84 olduğu tespit edilmiştir. Bu sorular çok kolay soru sınıflamasına girdiklerinden dolayı ($p \leq 0.80$) başarı testinden çıkarılmışlardır. Ayrıca, 16. maddenin güçlük indeksi 0.19 olarak bulunmuştur. Bu soru çok zor soru sınıflamasına girdiğinden dolayı başarı testinden çıkarılmıştır. Yapılan bu işlem sonrasında konuşma testi 21 soruluk bir teste dönüşmüştür.

Kalan 21 maddenin 8 maddesi kolay, 8 maddesi orta ve 5 maddesi ise zor madde olarak sınıflandırılmışlardır. 21 maddenin 100'lük puanlama sisteminde tam puana karşılık gelmemesinin testin kullanılabilirliğini olumsuz etkileyebileceği nedeniyle maddelerin güçlük dağılımları arasındaki denge de göz önüne alınarak kolay veya orta sorulardan ayırt ediciliği en düşük bir maddenin ölçme aracından çıkarılmasına karar verilmiştir. Maddeler incelendiğinde ayırt ediciliği en düşük maddenin (0,41) kolay güçlük sınıfındaki 10. ve 22. madde olduğu görülmüş ve 22. maddenin ölçme aracından çıkarılması tercih edilmiştir. Çıkan maddeler sonrasında, güçlük dağılımlarının Tablo 10'daki gibi gruplandığı görülmüştür.

Tablo 10.
Madde Güçlük Değerlerine Göre Nihai Testteki Madde Dağılımı

Güçlük Değeri	Madde Sayısı	Madde No	Yorum
$0.61 \leq p \leq 0.80$	7	7.8.10.12.20.21.23	Kolay madde
$0.41 \leq p \leq 0.60$	8	1.4.5.13.14.18.24.25	Orta madde
$0.20 \leq p \leq 0.40$	5	2.3.15.17.19	Zor madde

Tablo 10'dan görüldüğü üzere nihai testte yer alan maddeler değişik zorluk seviyelerindedir ve bu nedenle farklı türdeki öğrenci seviyelerine hitap edebilecek yeterliliktedir. Testin iç geçerliliğini ölçmek için KR-20 (Kuder-Richardson) değeri hesaplanmıştır. Yapılan analiz sonucu konuşma testinin iç geçerlilik katsayısı 0.87 olarak hesaplanmıştır. Elde edilen sonuçlar, konuşma testinin yüksek bir güvenilirliğe sahip olduğunu göstermiştir.

Tartışma, Sonuç ve Öneriler

Bu çalışmada ilkokul 2. sınıf İngilizce öğretim programı kapsamında öğrencilerin başarılarını ölçen geçerli ve güvenilir bir ölçme aracı geliştirmek ve bu aracın madde analizlerini ortaya koymak hedeflenmiştir. İlk defa 2013-2014 öğretim yılında uygulanmaya başlayan ilkokul 2. sınıf İngilizce öğretim programına devam etmekte olan öğrencilerde öngörülen amaçların gerçekleşme durumunu ortaya çıkarmak amacıyla konuşma ve dinleme becerilerine odaklanan başarı testi geliştirilmiştir. Yeni uygulanmaya başlanan programların öğrencilerin başarısına ne derece etki ettiğinin belirlenmesi önem taşımaktadır. Ölçme değerlendirme kapsamında sık kullanılan testlerden biri de başarı testidir.

Gerçekleştirilen madde analizleri sonucunda geçerliği ve güvenilirliği yüksek, 20 soruluk dinleme ve 20 soruluk konuşma testi oluşturulmuştur. Bunun yanı sıra, 2. sınıf programının hedeflediği temel beceriler olan dinleme ve konuşma becerilerine odaklanması oluşturulan testlerin güçlü yanısıdır.

Dinleme testi tüm sınıfa aynı anda uygulanabilmekte, konuşma testinde ise öğrenciler tek tek teste tabi tutulmaktadır. Dinleme testinde her sorunun iki defa dinletilmesiyle tüm sınıfa uygulanması yaklaşık 15-20 dakika sürmektedir. Konuşma bölümünün tamamlanması her bir öğrenci için yaklaşık beş dakika sürmektedir. Bu testlerin sınıf ortamında uygulanabilir ve kolaylıkla puanlanabilir olması sayesinde öğretmenler tarafından rahatlıkla kullanılabilen bir ölçme aracıdır. Geliştirilen testler 2. sınıf programında yer alan ölçülebilir tüm kazanımlara yönelik olduğu için dönem sonunda düzey belirleme testleri olarak kullanılabilirler. Geliştirilen testler öğrencilerin eksiklerini ölçme ve değerlendirmesi açısından da önemli materyallerdir. Türkiye’de ilk defa 2. sınıfta dil öğrenimi başladığı için ön teste gerek kalmaksızın bu testler son test olarak kullanılabilirler.

Extended Abstract

Introduction

Early foreign language learning is one of the important topics of discussion in language education. When the approaches of linguists about the age to start foreign language learning are analyzed, it is pointed out that there are differences between approaches. Çayıroğlu-İşgören (2005) gathers these differences into three titles: Approaches about Starting to Learn a Foreign Language at an Early Age, Approaches about Starting to Learn a Foreign Language after a Good Mother Tongue's Education and Approaches about Starting to Learn a Foreign Language at any age.

Early foreign language practices are applied in European Union (EU) countries. When the curriculum studies about foreign language education in Turkey are analyzed, the tendency is to start foreign language teaching early. In 1997, when the elementary education was extended to eight years, 4th and 5th grades started to take compulsory English courses for the first time and foreign language learning started at the age of 10 (Haznedar, 2010). In Turkey, with the decision taken by the Ministry of National Educational from the 2012-2013 education year on, twelve years compulsory education started instead of 8 years non-stop compulsory education and it was decided that foreign language teaching starts at the 2nd grade (MEB, 2012). In the new system, since English language teaching started to take place at the 2nd grade, it was needed that a new curricula locating second and third grades had to be prepared. Thus, a new curriculum considering the needs of young learners was necessary and in this line new curricula were designed. The second grade teaching program is being applied since the education years of 2013-2014. When the second grade teaching program is analyzed, it is seen that the core aim of the program is to develop positive attitudes towards English from an early age. In the 2nd grade teaching program, listening and speaking skills are aimed as the basic skills (MEB, 2013).

In the 2nd grade teaching program, it is essential to revise the previous topics and activities and remind them to students so that students will constantly remember them. Moreover, 2nd grade teaching program proposes an enjoyable and motivating learning context for young learners to make them feel comfortable and supported through the learning process. Besides, in the program it is planned for language learners/users to learn via game-based activities enriched with enjoyable audio-visual tools.

In a teaching program, it is important to find out how students' academic improvement is affected from the program (Parlett & Hamilton, 1977). For this reason, effective implementation of measurement and assessment in all grades is vital. In the measurement and assessment dimension of the 2nd grade English teaching curriculum, project and portfolio assessment, self and peer evaluation and teacher observations and evaluations are suggested. Furthermore, measurement and assessment methods measuring listening and speaking skills are recommended for objective evaluation (MEB, 2013). For that reason, in this study, an achievement test focusing on listening and speaking skills was developed in order to find out whether the 2nd grade aims were realized by 2nd grade students. Since listening and speaking skills are the core aims of 2nd grade English teaching program, in the achievement test items measuring listening and speaking skills took place.

Method

Research Design

"Survey models are research approaches that aim to describe a situation, which existed in the past or still exists, as it existed (Karasar, 2012; p. 77). In this study, survey model was used. An achievement test was developed and the present situation can be identified through this test.

Universe and Sample

The universe of the study is primary school second grade students. The participants for the reliability analysis of the draft listening test are 202 second grade students, who study in eight different classes of three primary schools in the spring semester of 2013-2014 education years. 108 of the participants are female and 94 of them are male. The draft speaking test was administered to 125 second grade students, who study in eight different classes of three primary schools. 69 of these students are female and 56 of them are male.

1st Study: The Design of Listening and Speaking Achievement Test and Validity Study

Within the framework of this study, the units of the program were analyzed and measurable aims were determined in order to find out whether the listening and speaking aims of the program were realized by the students.

In the program, there are nine measurable aims for listening skills. At least three questions were prepared for each listening aim. Students' age and developmental levels were taken into consideration while preparing the questions. Thus, pictures and colorful cards were utilized.

Experts were consulted for the content validity of the listening test. Necessary changes were realized after three English language field experts and three second grade English teachers' views and feedback. Finally, 40 questions took place in the draft listening test.

Due to the nature of the listening skill and the developmental features of students, the listening test did not include application, analysis, synthesis and evaluation level of questions of Bloom's taxonomy of educational aims.

Five second grade students were made listen to the listening texts in order to determine the clarity and relevance of the items to students before the pilot application of the achievement test; and necessary changes were realized after having obtained students' feedback. A native speaker of English read the listening records.

In the program, there are eight measurable aims for speaking skills. At least two questions were prepared for each speaking aim. Students' age and developmental levels were taken into consideration while preparing the questions. Thus, pictures and colorful cards were utilized.

Experts were consulted for the content validity of the speaking test. Necessary changes were realized after three English language field experts and three second grade English teachers' views and feedback. Finally, 25 questions took place in the draft speaking test.

Generally knowledge and comprehension level questions were asked to the students. Due to the age level of students, the speaking test did not include application, analysis, synthesis and evaluation level of questions.

Before the pilot application, the speaking questions were asked to five second grade students in order to determine the clarity and relevance of the items to students and necessary changes were realized after having obtained students' feedback.

2nd Study: The Reliability Study of Listening and Speaking Achievement Test Implementation and Results of Data Analysis

The draft test including listening and speaking parts was administered to second grade students. The listening part was administered to 202 second grade students who were studying in eight different classes of three primary schools, the speaking part was administered to 125 second grade students who were studying in eight different classes of three primary schools when the second grade English teaching

program was completed in May in the spring semester of 2013-2014 education year and item analyses were done. The item discrimination and item difficulty indexes were calculated for each item.

Item discrimination index serves to discriminate between students with higher and lower levels of knowledge (Özçelik, 2013). Item discrimination index may vary from -1 to +1. If the index has minus value, it shows that the item does not discriminate between individuals. Hence, such items must be omitted from the test (Büyüköztürk, Kılıç Çakmak, Akgün, Karadeniz, & Demirel; 2013).

Item discrimination value may be calculated by using the upper and lower 27% technique. Independent samples t-test is used to calculate the difference between item average scores of both groups. Therefore, in this study students' scores were ordered from high to low and according to upper and lower %27 technique, two groups were formed: for the listening test 55 upper and 55 lower groups of students, for the speaking test 34 upper and 34 lower groups of students. Item discrimination index was calculated for each item according to upper and lower group students' responses. Statistical Package for Social Sciences (SPSS) was used to analyze the data. In the process of analysis, items having less than 0.20 index value were accepted as low discriminator and omitted from the test (Özçelik, 2013; Turgut & Baykul, 2012).

Item difficulty index expresses the proportion or percentage of students who answered the item correctly. Item difficulty can range from 0.0 to 1.0. When item difficulty index comes closer to 1, it means the item is easier and when it comes closer to 0, the item gets harder (Turgut & Baykul, 2012). The average level of difficulty must be between 0.20 and 0.80 (Özçelik, 1992). When classifying the item difficulty indexes, items between $0.20 \leq p \leq 0.40$ were commented as difficult, items between $0.41 \leq p \leq 0.60$ were commented as average and items between $0.61 \leq p \leq 0.80$ were commented as easy (Kutlu, 2004).

Results

Results Regarding Listening Test

It is necessary to pay attention to whether there are variance differences between groups to calculate t-test of two independent groups. Therefore, in this study testing whether the variances were equal was needed. Levene test is one of the tests to test the variance equality (Büyüköztürk, 2012). If the significance value for Levene test is below 0.05, it means the variance of two groups is different. When there are different variances, the value in the line of "Equal variances not assumed" is used.

When the significance values are analyzed for t value, 13 items were found not to discriminate or discriminate vice versa. Therefore, items I.3, I.4, I.5, I.6, I.9, I.16, I.21, I.23, I.26, I.29, I.32, I.36, and I.38 were omitted from the draft listening test. According to t test value, 13 items were omitted and 27 items remained.

Item discrimination indexes were calculated by subtracting the number of lower group respondents who gave correct answers to each item from the number of upper group respondents who gave correct answers to each item and then dividing to 55 which is %27 of each group. As a result, I8., I13., I35. and I39. were omitted because their discrimination indexes were below 0.20. Also, I33 was omitted because it was found too easy as a result of the item difficulty test.

As a result of the item discrimination index, four items were removed and one item was removed as a result of the item difficulty index. Thus, 22 items remained. 13 of the items were easy, five of them were average and four items were difficult. In order to provide measurement easiness, two more easy items whose item discrimination indexes were the lowest were removed from the test. I17 and I18 were removed from the test.

In order to measure inner reliability of the test, KR-20 (Kuder-Richardson) value was calculated. If the items are true-false items and have two categories, Kuder Richardson formulas (KR20; KR21) are used.

Reliability value must be above 0.70 and more to be enough (Büyüköztürk, 2012). As a result of the analysis, the inner reliability value was found to be 0.72 which is an enough value.

Results Regarding Speaking Test

When the significance values for t value were analyzed, it was found out that all the items could discriminate between high achieving and low achieving students enough so there was no need to remove any item. Item discrimination indexes were calculated by subtracting the number of lower group respondents who gave correct answers to each item from the number of upper group respondents who gave correct answers to each item and then dividing to 34 which is %27 of each group.

When the item discrimination indexes were analyzed, it was seen that item discrimination indexes were above 0.20 for each item. Hence, no items were omitted from the speaking test. However, after having done the item difficulty analysis, three items I6, I9 and I11 were removed from the test because they were too easy. Also, I16 was removed because it was too difficult with 0.19 item difficulty value. After that, 21 items remained in the test. Eight of the items were easy, eight were average and five of them were easy. In order to provide scoring easiness, it was decided that one more item whose item discrimination index was the lowest from easy or average level items would be omitted from the test. When the items were analyzed, the item having the lowest item discrimination value was I22. Hence, I22 was omitted from the test. In the final speaking test, different level difficulty items took place to address different level of students.

In order to test inner reliability value, KR-20 value was calculated. As a result of the analysis, the inner reliability value of the speaking test was found to be 0.87 which is a highly reliable value.

Discussion, Conclusion & Implementation

In this study, it was aimed to develop a valid and reliable measurement tool within the framework of primary school 2nd grade English teaching program and to identify the item analyses of this tool.

An achievement test, focusing on listening and speaking skills, was developed in order to find out whether students studying the primary school 2nd grade English teaching curriculum, which first started to be implemented in the education year of 2013-2014, attained the aims or not. It is important to determine how a new teaching program affects students' success. An achievement test is one of the most frequently used tests for measurement and assessment.

As a result of the item analyses, 20 items listening and 20 items speaking test were developed which are highly valid and reliable. Furthermore, the test focused on listening and speaking skills that are the core skills of the program, which makes the test strong.

The listening test can be administered to the whole class at the same time; however, speaking test is administered to students one by one. It takes 15 to 20 minutes to administer the listening test with each item listened to twice. It takes five minutes to apply the speaking test to each student. These tests are easily applied and scored so it is a teacher-friendly measurement and assessment tool. The tests can be used as level tests at the end of the year since they address all the measurable aims in the second grade English teaching program. These tests are also good measurement and assessment tools to determine the weak parts of students. Since students start to learn English at the second grade for the first time in Turkey, these tests can be used as post-test without needing pre-tests.

Kaynakça

- Alpar, R. (2012). *Uygulamalı istatistik ve geçerlik-güvenirlik*. Ankara: Detay Yayıncılık
- Baykul, Y. (2010). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Ankara: Pegem Akademi.
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö.E., Karadeniz, Ş., & Demirel, F. (2013). *Bilimsel araştırma yöntemleri*. Ankara: Pegem Yayıncılık.
- Büyüköztürk, Ş. (2012). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem Akademi.
- Çayıroğlu-İşgören, O. (2005). *Yabancı dil öğretimine başlamada en uygun yaşın ne olduğuna ilişkin dilbilimsel yaklaşımlar ve öğretmen görüşleri*. Unpublished master's thesis, Abant Baysal Üniversitesi Sosyal Bilimler Enstitüsü, Bolu.
- Haznedar, B. (2010). *Türkiye’de yabancı dil eğitimi: Reformlar, yönelimler ve öğretmenlerimiz*. Paper presented at International Conference on New Trends in Education and Their Implications, Antalya, Turkey. Retrieved from <http://www.iconte.org/FileUpload/ks59689/File/166.pdf>
- Karasar, N. (2012). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayıncılık.
- Kutlu, Ö. (2004). *Ölçme ve değerlendirme*. Unpublished course notes. Ankara Üniversitesi Eğitim Bilimleri Fakültesi, Ankara.
- MEB (2012). *12 yıl zorunlu eğitim: Sorular-cevaplar*. Retrieved May 26, 2014, from http://www.meb.gov.tr/duyurular/duyurular2012/12Yil_Soru_Cevaplar.pdf
- MEB (2013). İlköğretim kurumları (ilkokullar ve ortaokullar) İngilizce dersi (2,3,4,5,6,7 ve 8. sınıflar) öğretim programı. Ankara.
- Mirici, İ.H. (2001). *Çocuklara yabancı dil eğitimi*. Ankara: Gazi Kitabevi.
- Özçelik, D.A. (1992). *Ölçme ve değerlendirme*. Ankara: ÖSYM.
- Özçelik, D.A. (2013). *Okullarda ölçme ve değerlendirme: Öğretmen el kitabı*. Ankara: Pegem Akademi.
- Parlett, M. & Hamilton, D. (1977). Evaluation as illumination: A new approach to the study of innovatory programs. In M. Parlett & G. Dearden (Eds.), *Introduction to illuminative evaluation: Studies in higher education* (pp. 9-29). California: Pacific Soundings Press.
- Turgut, M.F. & Baykul, Y. (2012). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem Akademi.

Ekler

Ek 1. Dinleme Başarı Testi

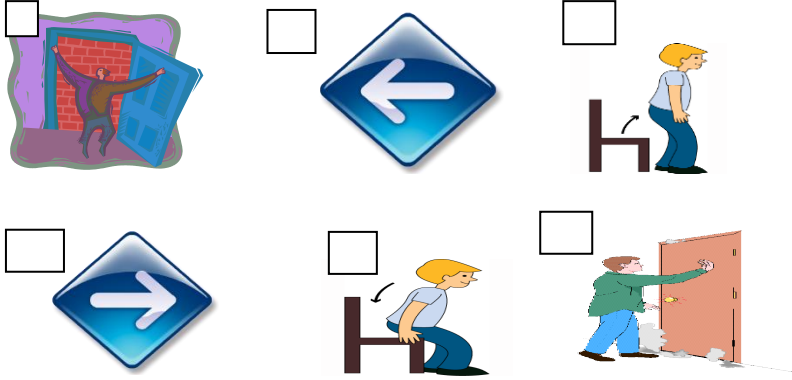
Ad-Soyad:

Dinleme Soruları

1. Bu soruda bazı harflerin okunuşunu dinleyeceksiniz. Dinlediğiniz harflerin kutusuna çarpı (X) işareti koyunuz.

a	i	n	u
d	k	o	y
e	l	s	w

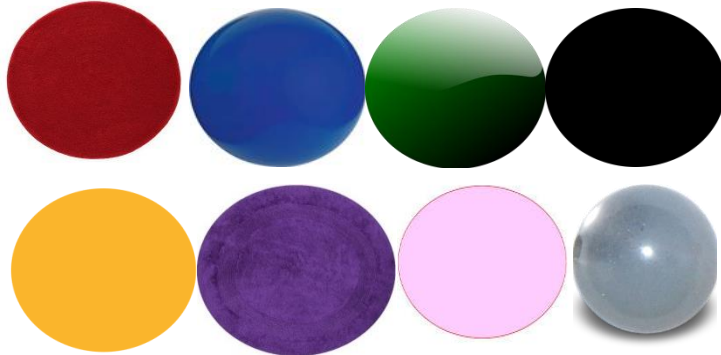
2. Bu soruda bazı emir cümlesi dinleyeceksiniz. Dinlediğiniz cümleler ile ilgili resimlerin kutusuna X (çarpı) işareti koyunuz.



3. Bu soruda bazı sayıları dinleyeceksiniz. Dinlediğiniz sayıların kutusuna çarpı (X) işareti koyunuz.

0 3 6 9 1 4 7 10 2 5 8

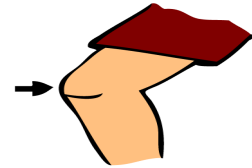
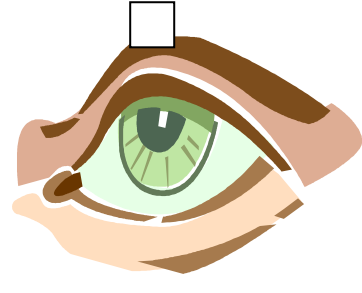
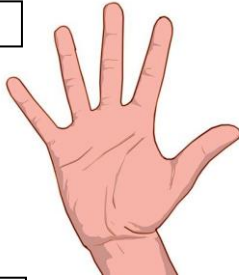
4. Bu soruda bazı renk adlarını dinleyeceksiniz. Dinlediğiniz renklerle ilgili resimlerin kutusuna X (çarpı) işareti koyunuz.



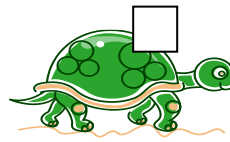
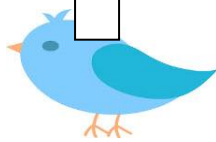
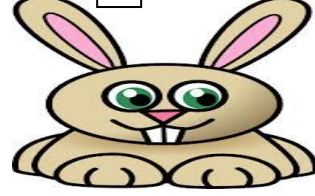
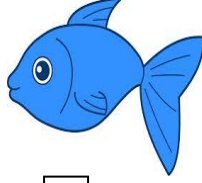
5. Bu soruda bazı öneri cümleleri dinleyeceksiniz. Dinlediğiniz cümleler ile ilgili resimlerin kutusuna X (çarpı) işareti koyunuz.



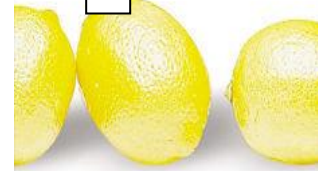
6. Bu soruda vücudun bazı bölümlerini dinleyeceksiniz. Dinlediğiniz vücudun bölümleri ile ilgili resimlerin kutusuna X (çarpı) işareti koyunuz.



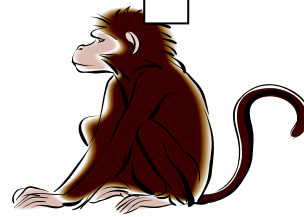
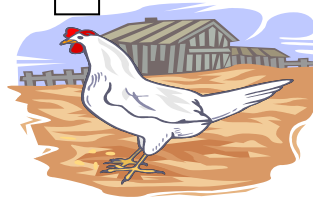
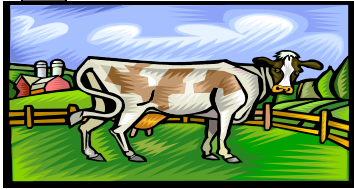
7. Bu soruda bazı evcil hayvanları dinleyeceksiniz. Dinlediğiniz evcil hayvanlar ile ilgili resimlerin kutusuna X (çarpı) işareti koyunuz.

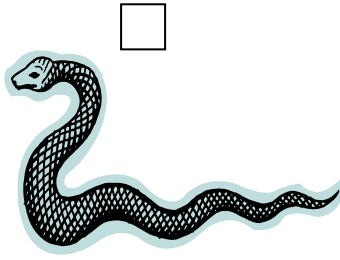
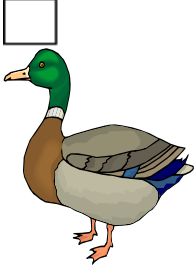
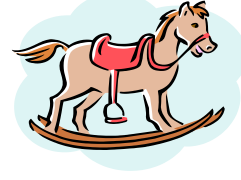


8. Bu soruda bazı meyve isimlerini dinleyeceksiniz. Dinlediğiniz meyveler ile ilgili resimlerin kutusuna X (çarpı) işareti koyunuz.



9. Bu soruda bazı hayvan isimlerini dinleyeceksiniz. Dinlediğiniz hayvanlar ile ilgili resimlerin kutusuna X (çarpı) işareti koyunuz.





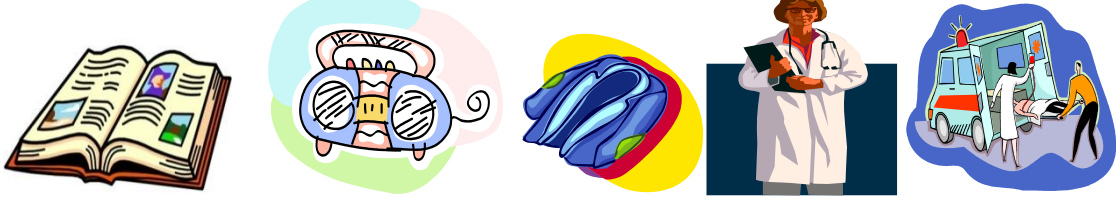
İlkokul 2. Sınıf İngilizce Dersi Dinleme Sınavı Kayıtları

1. 1. A 2. K
2. 1. Sit down! 2. Stand up!
3. 1. Nine 2. Zero 3. Eight 4. Two
4. 1. Black 2. Yellow
5. 1. Let's run 2. Let's play
6. 1. Eye 2. Knee 3. Shoulder
7. 1. Turtle 2. Bird
8. 1. Banana
9. 1. Chicken 2. Duck

Ek 2. Konuşma Başarı Testi

Konuşma Testi Soruları

1. Öğretmen öğrencilerden resimlerde gördükleri nesnelerin ne olduklarını sorar.



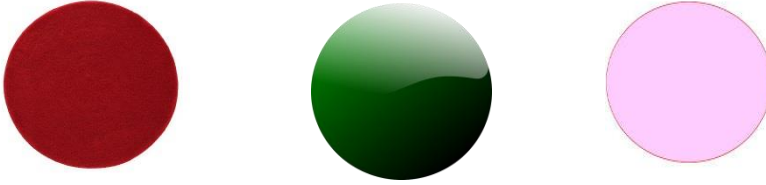
2. A: Hello. B: _____.

A: My name is _____. What is your name? B: _____

3. Öğretmen öğrencilerden gösterilen resimdeki nesnenin sayısını söylemelerini ister.



4. Öğretmen öğrencilere bazı kartlar gösterir ve bunların rengini sorar.



5. Öğretmen bir çocuk resmi gösterir ve işaretli vücudun bazı bölümlerinin ne olduklarını sorar.



6. Öğretmen evcil hayvan resimlerini gösterir ve ne olduklarını sorar.



7. Öğretmen meyve resmi gösterir ve ne olduğunu sorar.



8. Öğretmen bazı hayvan resimlerini gösterir ve ne olduklarını sorar.



İlkokul 2. Sınıf İngilizce Dersi Konuşma Sınavı Cevapları

1. 1. Book 2. Radio 3. Jacket 4. Doctor 5. Ambulance
2. 1. Hello. 2. My name is
3. 1. Four
4. 1. Red 2. Green 3. Pink
5. 1. Nose 2. Finger 3. Head
6. 1. Mouse 2. Cat 3. Fish
7. 1. Orange
8. 1. Lion 2. Elephant

