

ARAPÇA İÇİN ULUSLARARASI STANDART BİR SINAV GELİŞTİRME MODELİ

İsmail EKİNCİ*

Makale Bilgisi

Makale Türü: Araştırma Makalesi, **Geliş Tarihi:** 02 Mart 2026, **Kabul Tarihi:** 08 Haziran 2026, **Yayın Tarihi:** 26 Haziran 2026, **Atıf:** Ekinci, İsmail. "Arapça İçin Uluslararası Standart Bir Sınav Geliştirme Modeli". *Dinbilimleri Akademik Araştırma Dergisi* 26/2 (Haziran 2026): 715-743.

DOI: 10.33415/daad.1900856



Öz

Arapçanın küresel stratejik önemine rağmen, dil yeterliliğini uluslararası standartlarda, şeffaf ve karşılaştırılabilir şekilde belgeleyen kapsamlı bir ölçme sisteminin eksikliği akademik ve profesyonel hareketliliği kısıtlamaktadır. Bu çalışma, mevcut sınav uygulamalarının Avrupa Ortak Dil Çerçevesi-CEFR ve Amerikan Yabancı Dil Öğretimi Konseyi-ACTFL gibi küresel referans çerçeveleriyle uyum sorunlarını gidermek amacıyla; Arapçanın diglossia olgusu, kök-kalıp morfolojisi ve harekesiz yazı sistemi gibi tipolojik özelliklerini merkeze alan özgün bir sınav modeli önermektedir. Modern Standart Arapça temelinde dayalı bu model, ileri düzeylerde kontrollü lehçe maruziyeti ile otantikliği sağlarken; Madde Tepki Kuramı (IRT) ve Bilgisayar Ortamında Bireyselleşmiş Test (CAT) gibi çağdaş psikometrik yaklaşımları sürece entegre etmektedir. Üretken becerilerin değerlendirilmesinde insan ve yapay zekâ iş birliğine dayalı hibrit puanlama mekanizmalarını öngören araştırma, Arapçayı modern profesyonel bağlamlarda etkin bir iletişim aracı olarak konumlandırarak bilimsel bir ölçme-değerlendirme ekosistemi sunmaktadır.

Anahtar Kelimeler: Arap Dili ve Belagati, Dört Beceri Ölçme, Dil Yeterlik Testi, Sınav Geliştirme, Dil Değerlendirme.

* Doç. Dr., Ankara Hacı Bayram Veli Üniversitesi, İlahiyat Fakültesi, Temel İslam Bilimleri Bölümü, Arap Dili ve Belagati Anabilim Dalı, Ankara/Türkiye, ismailekinci1980@gmail.com, ORCID: <https://orcid.org/0000-0003-2728-9097>.

An International Standardized Examination Development Model for Arabic

İsmail EKİNCİ*

Article Information

Article Types: Research Article, **Received:** 02 March 2026, **Accepted:** 08 June 2026, **Published:** 26 June 2026, **Cite as:** Ekinci, İsmail. "An International Standardized Examination Development Model for Arabic". *Journal of Academic Research in Religious Sciences* 26/2 (June 2026): 715-743.

DOI: 10.33415/daad.1900856



Extended Abstract

This study aims to present the theoretical and practical framework of a standardized proficiency examination model for Arabic that possesses international validity and psychometric reliability. Arabic occupies a strategic position in global communication, diplomacy, academia, and commerce due to its historical depth, wide geographical distribution, and its status as one of the six official languages of the United Nations. Despite this prominence, there remains a notable absence of an internationally comparable and transparent proficiency examination system capable of certifying Arabic language competence based on comprehensive validity and reliability analyses. A significant proportion of existing Arabic examinations are designed to meet specific institutional needs; however, they often lack full alignment with global reference frameworks, balanced skill equivalence, and transparent reporting of item statistics. This fragmentation hinders the use of Arabic proficiency as an objective and widely recognized criterion in international labour markets and academic admissions processes. The present study grounds its theoretical framework in internationally recognized reference systems such as the Common European Framework of Reference for Languages (CEFR) and the American Council on the Teaching of Foreign Languages (ACTFL). Nevertheless, these frameworks are recalibrated in accordance with the typological and sociolinguistic characteristics of Arabic. Core structural variables—such as the phenomenon of diglossia, the root-and-pattern morphological system, and the tradition of an unvowelled orthography—are treated as fundamental considerations in test design. Within this context, the proposed model establishes Modern Standard Arabic as the ontological foundation of the examination while recommending limited and controlled

* Assoc. Prof. Dr., Ankara Hacı Bayram Veli University, Faculty of Theology, Department of Basic Islamic Sciences, Department of Arabic Language and Rhetoric, Ankara/Turkey, ismailekinci1980@gmail.com, ORCID: <https://orcid.org/0000-0003-2728-9097>.

exposure to dialectal varieties at advanced proficiency levels to enhance authenticity. Methodologically, the research adopts a qualitatively oriented theoretical analysis. A comprehensive review of the international language assessment literature informs the development of a feasible and context-sensitive testing architecture for Arabic. The four core language skills—listening, reading, speaking, and writing—are operationalized through level-based task design and analytic rating scales. To ensure psychometric robustness, the integration of contemporary measurement approaches is proposed, including Item Response Theory (IRT), Differential Item Functioning (DIF) analyses, and Computerized Adaptive Testing (CAT) models. In the assessment of productive skills, a hybrid scoring mechanism combining human raters and artificial intelligence is envisioned, supported by multiple-rater systems and statistical consistency coefficients to enhance inter-rater reliability. In conclusion, this study emphasizes that Arabic should be regarded not only as a language of historical and cultural heritage but also as an effective instrument of contemporary global communication in academic and professional contexts. By proposing a scientifically grounded assessment ecosystem based on principles of validity, fairness, and digital sustainability, the model demonstrates the potential to transform Arabic language education into a goal-oriented system and to establish a credible and internationally recognized proficiency standard. The study is considered not only as a contribution to the standardization of Arabic language instruction, but also as a structural step that will strengthen the academic and professional status of Arabic within global circulation.

Keywords: Arabic Language and Rhetoric, Four-Skill Assessment, Language Proficiency, Test Development, Language Evaluation.



Giriş

Günümüzde İngilizce, Fransızca, Almanca ve İspanyolca gibi dünya dilleri için Common European Framework of Reference for Languages (CEFR) ile uyumlu, standartlaştırılmış ve uluslararası geçerliliğe sahip sınav sistemleri bulunmaktadır. Bu sistemler, bireylerin dil yeterliliklerini eğitim kurumları, işverenler ve devlet otoriteleri nezdinde karşılaştırılabilir ve güvenilir ölçekler aracılığıyla belgelemelerine imkân tanımaktadır. Arapça ise 25'in üzerinde ülkede resmî dil statüsüne sahip, 400 milyondan fazla kişi tarafından konuşulan, aynı zamanda İslam medeniyetinin ve klasik bilginin temel kaynak dili olan, ayrıca Birleşmiş Milletler'in altı resmî dilinden biri konumunda bulunan küresel bir dil olmasına rağmen, dünya çapında kabul görmüş, dört temel dil becerisini bütüncül biçimde ve CEFR çerçevesine uygun olarak ölçebilen standart bir sınav sistemine sahip değildir. Ölçemediğiniz dil sizin değildir düsturunca, Arapça için standart bir sınav sisteminin olmayışı ciddi bir sorun ve eksiklik olarak karşımızda durmaktadır.

Mevcut Arapça sınavlarının önemli bir kısmı bölgesel kullanımla sınırlı kalmakta; bazıları yalnızca okuma veya gramer gibi tekil becerileri

değerlendirmekte; bir kısmı ise modern sınav tasarım ilkeleri, psikometrik geçerlik ve standardizasyon ölçütleri bakımından yeterli düzeyde yapılandırılmamaktadır. Arapça alanındaki örnekler (ör. AL-ARABIYYA TEST, ALPT, CIMA, Hamza Testi) çoğu zaman kurumsal ölçekli, tek beceri odaklı veya yalnızca Modern Standart Arapça (MSA) üzerine inşa edilmiş yapılarla sınırlıdır. Psikometrik geçerlik analizleri, madde kalibrasyonu, standard-setting süreçleri ve fairness (adalet) ilkeleri çoğu uygulamada sistematik ve şeffaf biçimde raporlanmadığı için, farklı ülkelerde uygulanan sınav sonuçlarının karşılaştırılabilirliği zayıflamaktadır. Bu durum, özellikle göç süreçleri, uluslararası eğitim programları, Arap ülkelerinde çalışma izinleri, akademik başvurular ve bilinçli dil öğrenimi planlamaları açısından ciddi bir boşluk doğurmaktadır. Üniversiteler, göçmenlik ofisleri, mesleki kurumlar ve araştırma enstitüleri, adayların Arapça yeterlilik düzeylerini ölçmek için ortak ve güvenilir bir referans sisteminden yoksun kalmaktadır.

718 | db

Bu çalışma, söz konusu boşluğu gidermeye yönelik olarak, Arapça için uluslararası düzeyde geçerliliğe, güvenilirliğe ve karşılaştırılabilirliğe sahip standart bir yeterlilik sınavı modelinin kuramsal ve uygulamalı çerçevesini ortaya koymayı amaçlamaktadır. Temel hedef, Arapçanın tipolojik ve sosyo-dilbilimsel özelliklerini dikkate alan; dinleme, okuma, konuşma ve yazma becerilerini dengeli, bütüncül ve bilimsel ölçütlerle değerlendiren; uluslararası standartlarla uyumlu bir sınav modelinin kuramsal altyapısını geliştirmektir. Bu bağlamda Arapçanın diglossia olgusu, kök-kalıp morfolojisi ve harekesiz yazı sistemi gibi özgün nitelikleri, sınav tasarımında merkezi değişkenler olarak ele alınmaktadır. Amaç, bu dilsel gerçeklikleri göz ardı etmeyen; ancak ölçme adaleti ve standardizasyon ilkelerinden taviz vermeyen bir model geliştirmektir. Ayrıca sınav sonuçlarının yalnızca tek bir toplam puan şeklinde değil, her bir beceriye ilişkin ayrıntılı yeterlilik profilleri biçiminde raporlanması ve böylece hem bireysel gelişim planlamasına hem de kurumsal karar süreçlerine katkı sunması hedeflenmektedir.

Yöntem bakımından çalışma, nitel ağırlıklı kuramsal bir tasarım modeline dayanmaktadır. Uluslararası dil ölçme-değerlendirme literatürü incelenmiş; özellikle American Council on the Teaching of Foreign Languages (ACTFL) ve CEFR'in ölçme felsefesi, seviye tanımları ve performans betimleyicileri analiz edilmiştir. Bu çerçeveler, Arapçanın dilbilimsel yapısı ve sosyokültürel bağlamı ışığında yeniden yorumlanarak uyarılama önerileri geliştirilmiştir. Psikometrik boyutta ise Madde Tepki Kuramı (IRT), madde yanlılığı (DIF) analizleri ve Bilgisayar Ortamında Bireysel-

leşmiş Test (CAT) modelleri gibi çağdaş ölçme teknikleri kavramsal düzeyde ele alınmış; geçerlik, güvenilirlik ve adalet ilkelerini teminat altına alacak bütüncül bir değerlendirme mimarisi tasarlanmıştır.

Sonuç olarak bu çalışma, Arapça için CEFR uyumlu, dört temel beceriyi kapsayan, geçerli, güvenilir ve adil bir uluslararası sınav sisteminin oluşturulmasını hem akademik hem de uygulamalı dil bilimi açısından acil bir gereklilik olarak temellendirmekte ve bu doğrultuda kapsamlı bir yol haritası sunmaktadır. Bu girişim, yalnızca Arapça öğretiminin standardizasyonuna katkı sağlamakla kalmayacak; aynı zamanda Arapçanın küresel dolaşımdaki akademik ve profesyonel statüsünü güçlendirecek yapısal bir adım olarak değerlendirilmektedir.

1. Arapça İçin Uluslararası Bir Dil Sınavı Geliştirme İhtiyacı

Arapça, tarihsel derinliği, jeopolitik önemi ve Birleşmiş Milletler'in altı resmi dilinden biri olması hasebiyle küresel iletişim ekosisteminde önemli bir konuma sahiptir. Ancak bu dilin yetkinliğini ölçen uluslararası standartlara sahip, standardize edilmiş bir sınav sisteminin eksikliği akademik ve profesyonel çevrelerde giderek daha fazla hissedilmektedir. Dilin morfolojik yapısındaki kök-kalıp sistemi üzerine kurulu sentaktik zenginlik, ölçme ve değerlendirme süreçlerini diğer Hint-Avrupa dillerinden ayıran temel bir zorluk teşkil etmektedir.¹ Özellikle Modern Standart Arapça (MSA) ile bölgesel lehçeler ('ammîce) arasındaki keskin diglossia (bir topluluğun içinde aynı dilin iki farklı lehçesinin bir arada bulunması) olgusunun varlığı,² adayların dilsel edinimlerini değerlendirirken hangi varyantın temel alınacağı sorunsalını doğurmaktadır.³ Literatürde diglossia, bireyin akademik ortamda MSA kullanırken günlük hayatta lehçeye yönelmesi sonucu oluşan bir işlevsel ayrışma olarak tanımlanmakta,⁴ bu durum ise sınav ortamındaki otantiklik kavramını karmaşıktırmaktadır.⁵ Standart bir sınavın başarısı, MSA'nın yüksek prestijli işlevlerini korurken, adayın gerçek hayat durumlarındaki iletişimsel yetisini ne derece yansıtılabildiğine bağlıdır. Dolayısıyla, tasarlanacak sınavın sadece gramer

¹ Karin C. Ryding, *A Reference Grammar of Modern Standard Arabic* (New York: Cambridge University Press, 2005), 18.

² Charles A. Ferguson, "Diglossia", *Word* 15/2 (1959), 325-340.

³ Mahmoud al-Batal (ed.), *Arabic as One Language: Integrating Dialect in the Arabic Language Curriculum* (Washington: Georgetown University Press, 2018), 45.

⁴ el-Sa'îd. M. Bedâvî, *Mustevyât el-'Arabiyye el-Mu'âsira fî Mısır* (Kahire: Dâru'l-Me'arif, 1973), 89; Clive Holes, *Modern Arabic: Structures Functions and Varieties* (Washington: Georgetown University Press, 2004), 5.

⁵ E. Saiegh-Haddad, "Linguistic Distance and Initial Reading Acquisition: The Case of Arabic Diglossia", *Applied Psycholinguistics* 24 (2003), 431-451.

bilgisini değil, dilin bağlamsal ve sosyo-linguistik katmanlarını da kapsamı bir zorunluluktur.⁶

Mevcut Arapça yeterlilik sınavlarının (ALPT-Arabıc Language Proficiency Test, CIMA-Certificat International de Maîtrise en Arabe, DLPT-Defense Language Proficiency Test vb.) temel sorunu, Avrupa Ortak Dil Referans Çerçevesi (CEFR) veya American Council on the Teaching of Foreign Languages (ACTFL) gibi küresel standartlarla tam bir korelasyon içerisinde olmamaları ve psikometrik geçerlilik analizlerinin yeterince şeffaf bir şekilde kamuoyuyla paylaşılmamasıdır.⁷ Bir sınavın uluslararası geçerliliği için sadece içerik geçerliliği değil, aynı zamanda yapı geçerliliği ve ölçüt geçerliliğinin de bilimsel yöntemlerle kanıtlanması gerekmektedir.⁸ Arapça için geliştirilecek uluslararası bir sınavın, okuma, yazma, dinleme ve konuşma becerileri arasındaki denklik analizlerini titizlikle yürütmesi⁹ ve her seviye için dilsel üretim beklentilerini netleştirilmesi şarttır.¹⁰ Bölgesel düzeyde yapılan Arapça sınavlarının pek çoğu, dilin zengin morfolojik ve sentaks yapısını test ederken adayları taktiksel bir mantıkla hareket etmeye itmekte ve dil eğitimindeki yapısal sorunları derinleştirmektedir.¹¹ Bu durum da dilin işlevsel kullanımının ihmal edilmesine sebep olmaktadır. Oysa modern ölçme kuramları, dilin sadece bir kurallar bütünü değil, bir toplumsal etkileşim aracı olduğunu vurgular. Bu bağlamda, geliştirilecek yeni modelin, Arapça ölçme sistemlerindeki geleneksel açıklar göz önünde bulundurularak, test maddesi analizi (IRT-Item Response Theory) gibi ileri istatistiksel yöntemlerle desteklenmesi ve farklı kültürlerden gelen adaylar için madde yanlılığı (DIF-Differential Item Functioning) içermemesi akademik bir gerekliliktir.¹²

Sonuç olarak, Arapçanın uluslararası statüsünün tahkim edilmesi ve küresel iş gücü piyasasında, akademik dünyada ve diplomatik ilişkilerde

⁶ Abdulkafi Albirini, "The Sociolinguistic Functions of Codeswitching between Standard Arabic and Dialectal Arabic", *Language in Society* 40/5 (2011), 713-735.

⁷ Aziz Khalil, "Assessing Arabic Language Proficiency. Al-'Arabiyya", *Journal of the American Association of Teachers of Arabic* 43 (2010), 100-115.

⁸ Ciyiril. J. Weir, *Language Testing and Validation: An Evidence-Based Approach* (England: Palgrave Macmillan, 2005), 12.

⁹ Lyle Bachman - Adrian Palmer, *Language Assessment in Practice* (İngiltere: Oxford University Press, 2010), 45.

¹⁰ J. Charles Alderson, *Assessing Reading* (New York: Cambridge University Press, 2000), 116.

¹¹ Hanada Taha-Thomure, "Arabic Language Education in the Arab World", *International Journal of West Asian Studies* 1 (2008), 175-192.

¹² A. S. Palmer, *Testing Arabic as a Foreign Language* (Washington: Georgetown University Press, 2007), 65.

nesnel bir ölçüt haline gelmesi, bilimsel temellere dayalı bir ölçme-değerlendirme ekosistemin inşasına bağlıdır. Bu sistem, dilin dijitalleşme sürecindeki yerini (NLP- Doğal Dil İşleme kapasitesi gibi) de hesaba katarak, adayların bilişsel süreçlerini ve dilsel performanslarını evrensel bir dille raporlamalıdır. Eğitim sistemlerinin bu standartlar doğrultusunda yenilenmesi, Arapça öğretiminde amaç odaklı bir dönüşümü tetikleyecektir.¹³ Pragmatik açıdan bakıldığında, standardize bir sınavın varlığı hem işverenler hem de yükseköğretim kurumları için güvenilir bir filtreleme mekanizması sunacak, akademik açıdan ise Arapça dilbilimi çalışmalarını modern dil pedagojisiyle entegre edecektir.¹⁴ Bu vizyoner adım, Arapçanın sadece bir miras dili veya klasik dil olarak algılanmasının önüne geçerek,¹⁵ onun dinamik ve küresel bir iletişim aracı olarak prestijini pekiştirecek ve dil üzerindeki ideolojik mitleri kıracaktır.¹⁶ Ayrıca Arapça yeterliliğini ölçen mevcut sınavlar genellikle CEFR doğrulaması, çok uluslu güvenilirlik çalışmaları ve beceriler arası denklik analizleri açısından eksik olarak görülmekte, bu durum eğitici kadronun yetkinlik ölçümlerine de yansımaktadır.¹⁷ Bu nedenle, Arapçanın uluslararası statüsünü destekleyecek, bilimsel temellere dayalı bir sınav sisteminin oluşturulması hem akademik hem pragmatik açıdan zorunludur.

2. Kuramsal Temeller ve Ölçme Felsefesi

Uluslararası standartlarda bir Arapça yeterlilik sınavının inşası, ölçme ve değerlendirme biliminin evrensel ilkeleri ile Arapçanın kendine özgü dilbilimsel paradokslarının sentezlenmesini zorunlu kılan, yüksek düzeyde bir kuramsal mimari gerektirmektedir. Bu mimarinin ilk sacayağını oluşturan CEFR, dili statik bir bilgi kümesi olarak değil, bireyin toplumsal, akademik ve mesleki bağlamlarda sergilediği işlevsel bir eylem kapasitesi olarak tanımlar.¹⁸ CEFR'in 'yapabilirim' (can-do) ifadeleri üzerine kurulu taksonomisi, Arapça özelinde MSA'nın yüksek prestijli işlevle-

¹³ Wahhab Kassem vd. (ed.), *Handbook for Arabic Language Teaching Professionals in the 21st Century* (New York City: Lawrence Erlbaum Associates, 2006), 155.

¹⁴ Munther Younes, *The Integrated Approach to Arabic Instruction* (London: Routledge, 2015), 42.

¹⁵ Emma Trentman, "Arabic Heritage Learners in Study Abroad", *Heritage Language Journal* 10/3 (2013), 432-451.

¹⁶ John Eisele, "Myth Ideology and Arabic Dialects", *Journal of Arabic Linguistics* 42 (2003), 198-215.

¹⁷ Hezi Brosh, "Arabic Language Proficiency of Elementary School Teachers" 2/1 (2013), 154-165.

¹⁸ Council of Europe, *Common European Framework of Reference for Languages: Learning Teaching Assessment* (New York: Cambridge University Press, 2001), 43.

rini kapsayacak şekilde genişletilmelidir.¹⁹ Ancak Arapçanın sosyolengüistik/toplum dil bilimi gerçeği olan diglossia, CEFR'in Avrupa dilleri için öngördüğü tek dilli varsayımları sarsmaktadır;²⁰ zira akademik bir metni anlama yetisi MSA ile bu metin üzerine sözlü bir tartışma yürütme pratiği (genellikle eğitilmiş lehçe kullanımı) arasındaki geçişkenlik,²¹ sınavın geçerlik katsayısını doğrudan etkilemekte ve konuşma becerisinin ölçümünde özgün metodolojiler gerektirmektedir.²² Bu bağlamda, eyleme dayalı yaklaşımların Arapça sınavları tasarlanırken, adayın sadece dil bilgisi kurallarını bilip bilmediğine bakılmamalı; bu kuralları farklı toplumsal ve kültürel ortamlarda ne kadar etkili kullanabildiği de ölçülmelidir.²³

Ölçme sisteminin teknik güvenilirliği ve etik meşruiyeti, Amerikan Eğitim Araştırmaları Derneği (AERA), Amerikan Psikoloji Derneği (APA) ve Ulusal Eğitimde Ölçme Konseyi (NCME) tarafından 2014 yılında güncellenen standartlar bütününe dayanmalıdır. Bu standartlar, bir testin sadece istatistiksel tutarlılığını değil, aynı zamanda adalet ve yanlılık denetimini de zorunlu kılar.²⁴ Arapça gibi geniş bir coğrafyada farklı eğitim ekollerinden gelen adayların girdiği sınavlarda, madde karakteristiği analizleri ve madde yanlılığı çalışmaları hayati önem taşımaktadır.²⁵ Özellikle metinlerin harekelemesi meselesi, bu standartlar ışığında ele alınması gereken kritik bir yapısal sorundur; başlangıç seviyelerindeki metinlerin tam harekeli olması gerekirken, ileri düzeylerde (C1-C2) harekesiz metinlerin kullanılması, adayın morfo-sentaktik çıkarım yeteneğini ölçen bir yapı geçerliliği unsuru haline gelmektedir.²⁶ Harekesiz bir metni doğru seslendirmek, sadece bir okuma eylemi değil, aynı zamanda derin bir dil bilgisi ve bağlam çözümü kapasitesinin kanıtıdır.²⁷ Bu nedenle, test tasarımında kullanılan her bir metnin ve sorunun, adayın arka plan bilgilerinin-

¹⁹ David Little, "The Common European Framework of Reference for Languages: Content Purpose Origin Reception and Impact", *Language Teaching* 39/3 (2006), 167.

²⁰ Kees Versteegh, *The Arabic Language* (Edinburgh: Edinburgh University Press, 2014), 189.

²¹ Bedâvî, *Musteveyât el-'Arabiyye el-Mu'âsira fî Mısır*, 105.

²² Martin Bygate, *Speaking* (London: Oxford University Press, 1987), 12.

²³ Mahmoud al-Batal, "Diglossia Proficiency: The Need for an Intermediate Level of Arabic", *The Modern Language Journal* 75/2 (1992), 174.

²⁴ Robert L. Ebel - David A. Frisbie, *Essentials of Educational Measurement* (New Delhi: Prentice Hall, 1991), 78.

²⁵ Lyle Bachman, *Fundamental Considerations in Language Testing* (London: Oxford University Press, 1990), 238.

²⁶ Hanan Khalifa - Ciyril. J. Weir, *Examining Reading* (Cambridge: Cambridge University Press, 2009), 45.

²⁷ Alderson, *Assessing Reading*, 122.

den bağımsız olarak sadece dilsel yetkinliğini ölçecek şekilde standardize edilmesi, uluslararası güvenilirlik protokollerinin temel bir gereğidir.²⁸

Kuramsal çerçevenin üçüncü bileşeni olan ACTFL- American Council on the Teaching of Foreign Languages (Amerika Yabancı Dil Öğretimi Konseyi)/OPI- Oral Proficiency Interview (Sözlü Yeterlilik Mülakatı) modeli, özellikle üretim odaklı becerilerin (konuşma ve yazma) değerlendirilmesinde nesnel bir derecelendirme sistemi (rubric) sunar.²⁹ Arapçanın morfolojik karmaşıklığı, fiil çekimlerindeki zenginlik ve türetme sisteminin yoğunluğu, değerlendiriciler arası tutarlılığın sağlanmasını zorlaştırmaktadır.³⁰ ACTFL kriterleri, adayın performansını başlangıçtan seçkine kadar uzanan bir süreklilikte, dilin biçimsel doğruluğu ile içeriksel derinliğini dengeleyerek puanlar. Dinleme ve okuma materyallerinde Modern Standart Arapça omurgayı oluşturmakla birlikte, ileri düzey yeterlilik tanımlarında sınırlı ve kontrollü bir lehçe ('ammice) maruziyeti, sınavın otantiklik niteliğini güçlendirmektedir.³¹ Nitekim gerçek bir dil kullanım ortamında, bir Arapça konuşanı sadece haber bültenlerini değil, entelektüel tartışmalardaki lehçe geçişlerini de anlamak durumundadır.³² Bu durum, adalet ilkesiyle çalışmayacak şekilde, lehçe çeşitliliğinin bir engel değil, üst düzey dilsel bir yetkinlik göstergesi olarak sınava entegre edilmesini gerektirir.³³ Sonuç olarak, bu üçlü kuramsal yapı (CEFR, AERA/APA/NCME Standartları ve ACTFL/OPI Modeli) üzerine inşa edilen bir sınav, Arapçanın hem klasik mirasını koruyan hem de modern dünyanın iletişimsel ihtiyaçlarına cevap veren bilimsel bir ölçüt haline gelecektir.

3. Arapça Sınavı İçin İçerik Geliştirme ve Madde Yazımı

Uluslararası standartlarda bir Arapça sınavının içerik geliştirme ve madde yazımı süreci, dilin yapısal hususiyetlerini ölçme-değerlendirme biliminin teknik gereklilikleriyle harmanlayan sofistike bir tasarım aşamasıdır. Bu süreçte soru hazırlayan uzmanların sadece dili çok iyi bilmesi yetmez; aynı zamanda ölçme yöntemlerinden, soru istatistiklerinden ve öğrenme seviyelerinden (bilişsel taksonomilerden) de anlamaları gere-

²⁸ Samuel Messick, "Validity", *Creative Education* 3/6A (2012), 13-104.

²⁹ Judith e. Liskin-Gasparro, "The ACTFL Proficiency Guidelines and the Oral Proficiency Interview: A Retrospect", *Foreign Language Annals* 36/4 (2003), 483-490.

³⁰ Elvira Swender, "Oral Proficiency Testing in the Real World", *Foreign Language Annals* 36/4 (2003), 520-526.

³¹ Renate Schulz, "The Challenge of Assessing Cultural Understanding in the Context of Foreign Language Instruction", *Foreign Language Annals* 40/1 (2006), 9-26.

³² Mahmoud al-Batal - R. Kirk Belnap, *The Teaching and Learning of Arabic in the United States: Realities Needs and Future Directions* (New York: Routledge, 2006), 392.

³³ Younes, *The Integrated Approach to Arabic Instruction*, 55.

kir.³⁴ Arapça özelinde madde yazımı; morfolojik farkındalık, sentaktik çözümleme ve bağlamsal çıkarım becerilerini eşzamanlı olarak hedeflemelidir.³⁵ Özellikle başlangıç ve orta seviyelerde (A1-B1), adayların kelime köklerini ve vezinleri tanıma yetileri, izole gramer soruları yerine bağlam içi tamamlama (cümle testleri) veya metin bütünlüğü üzerinden sorgulanmalı³⁶ ve adayın girdileri doğru işleme süreçleri titizlikle tasarlanmalıdır.³⁷ Bu yaklaşım, adayın dili bir formüller bütünü olarak değil, yaşayan bir mekanizma olarak algılayıp algılamadığını test eder. İleri düzeylerde ise (C1-C2), madde yazımı stratejik bir dönüşüm geçirerek, doğrudan ifade edilmeyen anlamların, metinler arası ilişkilerin ve retorik sanatların (belâgat) kavranmasına odaklanmalıdır. Bu aşamada, seçeneklerin çeldirici gücü, adayın nüansları ayırt etme kapasitesini ölçecek şekilde kalibre edilmelidir.³⁸

Soru tiplerinin yapılandırılmasında otantiklik ve etkileşim anahtar kavramlardır.³⁹ Dinleme modülünde, başlangıç düzeyinde net telaffuzlu MSA haber metinleri kullanılırken; ileri seviyelerde, akademik seminer kayıtları, entelektüel paneller ve sınırlı ölçüde eğitimli lehçe unsurları içeren tartışma programları kullanılmalıdır.⁴⁰ Okuma modülünde ise, edebi metinlerden teknik raporlara, hukuki metinlerden dijital medya içeriklerine kadar geniş bir yelpaze sunulmalıdır.⁴¹ Özellikle harekesiz metin okuma becerisi, ileri düzey test maddelerinde adayın sözcük türünü (isim/fiil ayrımı) ve i'rab bilgisini bağlamdan çıkarıp çıkarmadığını ölçen bir yapı geçerliliği unsuru olarak kurgulanmalıdır. Yazma ve konuşma becerilerinde ise görev temelli maddeler tercih edilmelidir. Örneğin, adaya bir grafik yorumlatma, resmi bir e-posta yazdırma veya belirli bir argümanı savunma gibi gerçek hayat senaryoları sunulmalıdır. Bu maddelerin puanlanmasında kullanılacak dereceli puanlama anahtarları,

³⁴ Thomas M. Haladyna - Michael C. Rodriguez, *Developing and Validating Test Items* (New York: Routledge, 2013), 112.

³⁵ Lyle Bachman - Adrian S. Palmer, *Language Testing in Practice* (London: Oxford University Press, 1996), 67.

³⁶ Saleh al-Shuwairekh, *Vocabulary Selection for Arabic as a Foreign Language* (Riyadh: King Saud University Press, 2001), 45.

³⁷ Gary Buck, *Assessing Listening* (New York: Cambridge University Press, 2001), 94.

³⁸ Karin Ryding, *Teaching and Learning Arabic as a Foreign Language: A Guide for Teachers* (Washington: Georgetown University Press, 2013), 138.

³⁹ Glenn Fulcher, *Practical Language Testing* (London: Routledge, 2010), 156.

⁴⁰ Khalil, "Assessing Arabic Language Proficiency. Al-'Arabiyya", 100-115.

⁴¹ Taha-Thomure, "Arabic Language Education in the Arab World", 180-192.

dilin doğruluğu, akıcılığı, kelime dağarcığının genişliği ve sosyo-linguistik uygunluk kriterlerini dengeli bir biçimde içermelidir.⁴²

Teknik altyapıda ise, Bilgisayar Ortamında Bireyselleşmiş Testler (Computer Adaptive Testing-CAT) modelinin benimsenmesi, sınavın verimliliğini ve güvenliğini maksimize edecektir.⁴³ Bu sistemde, Madde Tepki Kuramı (Item Response Theory-IRT) algoritmaları kullanılarak, adayın her bir doğru cevabından sonra karşısına daha zor, yanlış cevabından sonra ise daha kolay bir sorunun gelmesi sağlanır.⁴⁴ Bu model, Arapçanın geniş seviye yelpazesinde (A1'den C2'ye) çok daha kısa sürede, çok daha kesin bir yetenek kestirimi yapılabilmesine olanak tanır.⁴⁵ Ayrıca, açık uçlu soruların (yazma ve konuşma) değerlendirilmesinde Yapay Zekâ Destekli Puanlama sistemlerinin insan puanlayıcılar ile entegre edilmesi, değerlendirme sürecindeki yanlılığı minimize ederken nesneliği artıracaktır. Bu dijital dönüşüm, Arapçanın küresel bir dil olarak dijital okuryazarlık gereksinimleriyle uyumlanmasını, ölçme biliminin sosyo-bilişsel temelleriyle desteklenmesini⁴⁶ ve sınav sonuçlarının uluslararası kurumlar nezdinde tartışmasız bir güvenilirlik kazanmasını sağlayacaktır.

4. Dört Becerinin Ölçülmesi: İlkeler ve Uygulama

db | 725

4.1. Dinleme

Uluslararası bir Arapça yeterlilik sınavında dinleme becerisinin ölçülmesi, dilin fonolojik, morfolojik ve semantik katmanlarının yanı sıra Modern Standart Arapçanın kendine özgü retorik yapısını kapsayan bütüncül bir tasarımı gerektirir. Bu tasarım sürecinde, Arapçanın iki farklı dilsel varyantı arasındaki sesbilimsel ve sözcüksel mesafenin adayların işitsel algı süreçlerine olan etkisi göz önünde bulundurulmalıdır.⁴⁷ Dinleme modülü, sadece işitsel verinin kodunu çözmeyi değil, aynı zamanda veriler arasındaki hiyerarşik anlam ilişkilerini saptamayı, konuşmacının

⁴² Andrew D. Cohen, *Assessing Language Ability in the Classroom* (Boston: Heinle & Heinle, 1994), 204.

⁴³ Howard Wainer, *Computerized Adaptive Testing: A Primer* (New York: Routledge, 2000), 13.

⁴⁴ Frederic M. Lord, *Applications of Item Response Theory to Practical Testing Problems* (New York: Routledge, 1980), 42.

⁴⁵ Cynthia G. Parshall vd., *Practical Considerations in Computer-Based Testing* (Berlin: Springer International Publishing, 2002), 65.

⁴⁶ Robert J. Mislevy, *Sociocognitive Foundations of Educational Measurement* (New York: Routledge, 2018), 210.

⁴⁷ Saiegh-Haddad, "Linguistic Distance and Initial Reading Acquisition: The Case of Arabic Diglossia", 431-451.

niyetini çözümlmeyi ve örtük bilgiyi açığa çıkarmayı hedeflemelidir.⁴⁸ Bu süreç, alt düzey bilişsel süreçlerden üst düzey meta-bilişsel süreçlere doğru kademeli bir yapı sergilemeli ve adayın girdileri anlık olarak işleme kapasitesini ölçmelidir.⁴⁹ Başlangıç düzeylerinde (A1–A2), adayların sınırlı kelime dağarcığı ve temel dil bilgisi seviyeleri göz önünde bulundurularak; somut ihtiyaçlara yönelik kısa, doğrudan ve yüksek frekanslı kelimelerden oluşan günlük diyaloglar temel alınmalı ve bu seviyelendirme uluslararası ölçeklerin dilsel üretim beklentileriyle tam uyumlu yürütülmelidir (ACTFL, 2025). Bu aşamada ölçme odağı, adayın belirli anahtar kelimeleri yakalaması ve temel düzeydeki bildirimsel ifadeleri anlamasıdır.⁵⁰

Orta ve ileri düzey yeterlilik basamaklarında (B2–C2) ise dinleme materyalleri, adayın bilişsel yükünü artıracak şekilde akademik konferanslar, resmi diplomatik bildirimler, entelektüel tartışmalar ve teknik raporlar gibi karmaşık içeriklere evrilmelidir.⁵¹ Bu seviyelerde ölçme felsefi, adayın metnin sadece lafzi anlamını değil, aynı zamanda metnin bütünsel mantığını ve konuşmacının bakış açısını ne ölçüde kavradığına odaklanır.⁵² Özellikle MSA'nın esas alınması, sınavın standardizasyonu ve adalet ilkesi açısından hayati bir zorunluluktur.⁵³ MSA kullanımı, adayın farklı bölgesel aksanların veya yerel lehçe müdahalelerinin yarattığı ek yükten arındırılmasını, böylece sadece standart dil yetkinliğinin ölçülmesini sağlar. Ses kayıtlarının doğal konuşma hızında gerçekleştirilmesi, sınavın otantiklik katsayısını artırırken; stüdyo kalitesindeki kayıtlar, aksan dengesizliği veya gürültü gibi dışsal değişkenlerin adayın performansını olumsuz etkilemesini engeller. İleri düzeylerde, konuşmacıların retorik vurguları ve vurgulanan i'rab nüansları, adayın morfo-sentaktik çıkarım yeteneğini test eden birer ayıraç vazifesi görür.⁵⁴

Dinleme becerisinin değerlendirilmesinde, madde yazımı stratejileri, sesli uyarının süresi ve karmaşıklığı ile doğru orantılı olmalıdır.⁵⁵ A1 seviyesinde bir adaya fiyat, saat, yön gibi somut veriler sorulurken; C1 ve C2

⁴⁸ Michael Rost, *Teaching and Researching Listening* (New York: Routledge, 2011), 124.

⁴⁹ Buck, *Assessing Listening*, 96.

⁵⁰ "ACTFL Assigning CEFR Ratings to ACTFL Assessments" (Erişim 10 Kasım 2025).

⁵¹ Larry Vandergrift - Christine Goh, *Teaching and Testing L2 Listening* (New York: Routledge, 2012), 235.

⁵² Weir, *Language Testing and Validation: An Evidence-Based Approach*, 89.

⁵³ Palmer, *Testing Arabic as a Foreign Language*, 74.

⁵⁴ Alderson, *Assessing Reading*, 132.

⁵⁵ Penny Ur, *Teaching Listening Comprehension* (New York: Cambridge University Press, 1984), 34.

seviyelerindeki adaylara, sunulan argümanlardaki mantıksal boşlukları tespit etme veya konuşmacının ironi ve metafor kullanımını analiz etme görevleri verilmelidir. Bu süreçte kullanılacak çeldiricilerin, adayın sadece metinde geçen ses benzerliklerine dayalı bir tahmin yürütmesini engelleyecek, ancak metnin genel bağlamıyla uyumlu olacak şekilde tasarlanması akademik bir gerekliliktir.⁵⁶ Ayrıca, dinleme materyallerinde kullanılan MSA'nın, klasik metinlerin ağıdalı yapısından ziyade, modern medya ve akademi dilinin işlevsel formunda olması, sınavın güncellik ve mesleki geçerlilik hedefleriyle örtüşmektedir. Sonuç olarak, yüksek kaliteli bir dinleme testi, dilin fonetik zenginliğini standart bir normda sunarak, adayların işitsel anlama kapasitelerini evrensel ölçme kriterleri dahilinde raporlamalıdır.⁵⁷

4. 2. Okuma

Uluslararası bir Arapça yeterlilik sınavında okuma becerisinin ölçülmesi, dilin yazılı formlarındaki yapısal çeşitliliği ve Modern Standart Arapça'nın farklı türlerdeki kullanımını kapsayan sistematik bir metin seçimi stratejisini zorunlu kılar.⁵⁸ Bu strateji belirlenirken, Arapçadaki diglossia olgusunun ve yazılı standart dil ile sözlü varyantlar arasındaki mesafenin okuma edinişi süreçlerindeki belirleyici rolü göz önünde bulundurulmalıdır.⁵⁹ Okuma modülü, sadece sözcük düzeyinde anlamlandırmayı değil, metnin türsel özelliklerini, yazarın amacını ve dilin retorik araçlarını kavrama yetisini değerlendirmelidir.⁶⁰ Bu bağlamda metin havuzu; günlük yaşamın pratik ihtiyaçlarını karşılayan broşür ve ilanlardan, güncel olayları aktaran haber metinlerine,⁶¹ derinlemesine analiz içeren makalelerden yüksek düzeyde soyutlama gerektiren akademik ve edebi yazılara kadar geniş bir yelpazeden seçilmelidir.⁶² Bu çeşitlilik, sınavın temsil gücünü artırırken, adayın dili farklı sosyal ve bilimsel bağlamlarda kullanma kapasitesini doğrular. Bu doğrulanma süreci, modern dil öğre-

⁵⁶ Janeette Hansen, *Second Language Listening: Theory and Practice* (New York: Cambridge University Press, 2010), 210.

⁵⁷ Kassem vd., *Handbook for Arabic Language Teaching Professionals in the 21st Century*, 162.

⁵⁸ William Grabe, *Reading in a Second Language: Moving from Theory to Practice* (New York: Cambridge University Press, 2009), 214.

⁵⁹ Saiegh-Haddad, "Linguistic Distance and Initial Reading Acquisition: The Case of Arabic Diglossia", 431-451.

⁶⁰ Alderson, *Assessing Reading*, 167.

⁶¹ Shuwaikeh, *Vocabulary Selection for Arabic as a Foreign Language*, 58.

⁶² Ryding, *Teaching and Learning Arabic as a Foreign Language: A Guide for Teachers*, 142.

tim politikalarının evrensel ve kültürel boyutlarıyla uyumlu yürütülmeli,⁶³ dilin dijitalleşme ve doğal dil işleme süreçlerindeki morfo-sentaktik standartları da göz ardı edilmemelidir.⁶⁴

Düşük düzeylerde (A1–A2), okuma görevleri adayın temel bilgi bulma ve genel hatları belirleme yetilerini hedeflemelidir. Bu aşamada, morfolojik açıdan yalın, kelime sıklığı bakımından yüksek ve somut içerikli metinler tercih edilmeli ve test tasarım döngüsü bu doğrultuda yapılandırılmalıdır.⁶⁵ Adaydan, bir broşürdeki saati bulması veya kısa bir haber girişinden ana fikri saptaması beklenir. Orta düzeylere (B1–B2) geçildiğinde, metinlerin dilsel karmaşıklığı artırılarak adaydan çıkarım yapma ve metni özetleme gibi daha ileri bilişsel görevleri yerine getirmesi istenir.⁶⁶ Bu seviyedeki metinler, sadece açıkça ifade edilen bilgileri değil, metnin satır aralarında gizli olan anlam ilişkilerini de içermelidir.⁶⁷ Özellikle morfolojik yoğunluğun artmasıyla birlikte, adayın kelime köklerinden hareketle türetilmiş sözcüklerin anlamını bağlamdan çıkarması bir ölçüt haline gelir.⁶⁸

İleri düzey yeterlilik basamaklarında (C1–C2), okuma becerisi eleştirel bir boyut kazanarak metinler arası ilişki kurma ve yazarın ideolojik veya retorik tutumunu çözümleme görevlerini kapsamalıdır.⁶⁹ Bu seviyedeki akademik ve edebi metinler, harekesizlik unsurunu bir zorluk değişkeni olarak kullanarak, adayın sentaktik çözümlene kabiliyetini en üst düzeyde test eder.⁷⁰ C1 ve C2 düzeyindeki aday, farklı kaynaklardan gelen bilgileri sentezlemeli, metnin üslup özelliklerini fark etmeli ve karmaşık argüman yapılarını deşifre edebilmelidir.⁷¹ Metinlerin zorluk derecesi, sadece kelime sayısı ile değil, aynı zamanda kavramsal yoğunluk ve dilin figüratif kullanımı (mecaz, kinâye, istiâre) ile belirlenmeli, test maddele-

⁶³ Tulus Musthofa, “CEFR-Based Policy in Arabic Language Teaching and Cultural Dimension in Indonesian Islamic Higher Education”, *Eurasian Journal of Applied Linguistics* 8/2 (2022), 96-107.

⁶⁴ Nizar Y. Habash, “Introduction to Arabic Natural Language Processing”, (*Synthesis Lectures on Human Language Technologies*, ed. Graeme Hirst (Colombia: Columbia University, 2010), 5.

⁶⁵ Bachman - Palmer, *Language Assessment in Practice*, 256.

⁶⁶ Khalifa - Weir, *Examining Reading*, 62.

⁶⁷ Kassem vd., *Handbook for Arabic Language Teaching Professionals in the 21st Century*, 175.

⁶⁸ Reima al-Jarf, *Processing of Arabic Syntax* (Riyadh: King Saud University Press, 2001), 45.

⁶⁹ Thom Hudson, *Teaching Second Language Reading* (London: Oxford University Press, 2007), 240.

⁷⁰ Taha-Thomure, “Arabic Language Education in the Arab World”, 175-192.

⁷¹ Fulcher, *Practical Language Testing*, 178.

rinin sınıf içi ve geniş ölçekli pratik geçerlilik süreçleri bu karmaşıklığı karşılamalıdır.⁷² Sonuç olarak, bilimsel bir ölçme-değerlendirme süreci, CEFR düzeylerine göre kalibre edilmiş bu metinler ve görevler aracılığıyla, Arapça okuma yetkinliğini tesadüfi sonuçlardan arındırılmış, nesnel ve geçerli bir zemine oturtmalıdır.

4. 3. Konuşma

Uluslararası bir Arapça yeterlilik sınavında konuşma becerisinin ölçülmesi, dilin en dinamik ve diglossia olgusundan en çok etkilenen boyutu olması hasebiyle, titizlikle yapılandırılmış bir değerlendirme mimarisi gerektirir.⁷³ Bu mimari kurulurken, Modern Standart Arapça ile yerel diyalektler arasındaki sesbilimsel ve sözcüksel mesafenin adayın sözlü üretim hızı üzerindeki etkileri dikkate alınmalıdır.⁷⁴ Konuşma modülü, adayın sadece mekanik ses üretim kapasitesini değil,⁷⁵ kelime dağarcığından doğru yapıları seçme hızını, bu yapıları bağlama uygun sosyolinguistik normlarla harmanlama yetisini ve iletişimsel bir hedefi gerçekleştirme başarısını ölçmelidir. Bu doğrultuda sınav, yapılandırılmış mülakatlar ve görev temelli diyaloglar olmak üzere iki ana eksenle kurgulanmalı, test tasarım döngüsü ve ölçme kriterleri bu evrensel standartlarla desteklenmelidir.⁷⁶ Yapılandırılmış mülakatlar, adayın kendisiyle ilgili somut bilgilerden başlayarak soyut kavramlar ve toplumsal meselelere uzanan bir çizgide dilsel performansını örneklemesini sağlarken; sözlü dilin test edilmesindeki pratik formatlar⁷⁷ ve görev temelli diyaloglar adayı gerçek hayat senaryoları içine yerleştirerek dilin işlevsel kullanımını test eder. Bu süreç, uluslararası ölçeklerin sunduğu seviye karşılıklarına ve kültürel boyut analizlerine tam uyumlu olarak yürütülmelidir.⁷⁸

Değerlendirme süreci, sübjektiviteden arındırılmış, nesnel ve şeffaf ölçütlere dayanmalıdır.⁷⁹ Bu bağlamda beş temel kriter ön plana çıkmaktadır: Akıcılık, adayın konuşma sırasında sergilediği doğal tempo ve du-

⁷² Cohen, *Assessing Language Ability in the Classroom*, 215.

⁷³ Batal, "Diglossia Proficiency: The Need for an Intermediate Level of Arabic", 182.

⁷⁴ Saiegh-Haddad, "Linguistic Distance and Initial Reading Acquisition: The Case of Arabic Diglossia", 431-451.

⁷⁵ Bygate, *Speaking*, 42.

⁷⁶ Bachman - Palmer, *Language Assessment in Practice*, 284.

⁷⁷ Nic Underhill, *Testing Spoken Language* (Cambridge University Press, 1987), 12.

⁷⁸ "ACTFL Assigning CEFR Ratings to ACTFL Assessments"; Musthofa, "CEFR-Based Policy in Arabic Language Teaching and Cultural Dimension in Indonesian Islamic Higher Education", 96-107.

⁷⁹ Arthur Hughes, *Testing for Language Teachers* (New York: Cambridge University Press, 2003), 110.

raksama yönetimidir; doğruluk, Modern Standart Arapçanın morfo-sentaktik kurallarının (i'rab, fiil çekimleri, cinsiyet uyumu) doğru uygulanması ve değerlendiriciler arası tutarlılığın sağlanmasıdır;⁸⁰ telaffuz, harflerin mahreçlerine uygunluğu ve entonasyonun anlamı desteklemesidir; söylem organizasyonu, cümlelerin mantıksal bir sıra ve bağlaçlarla (edatlar/bağlaçlar) birbirine eklenerek bütüncül bir anlatı oluşturmalarıdır;⁸¹ etkileşim becerisi ise, adayın bir diyalog içerisinde sırasını yönetme, karşı tarafın ifadelerine uygun tepki verme ve iletişimsel stratejiler geliştirme kapasitesidir.⁸² Bu kriterlerin her biri, CEFR ve ACTFL seviyelerine göre özelleştirilmiş analitik rubriklerle tanımlanmalıdır.⁸³ Örneğin, B2 düzeyindeki bir adaydan beklenen performans, sadece dil bilgilere doğruluk değil, aynı zamanda belirli bir kültürel etkinliği veya fikri ikna edici bir dille tanıtabilmesi, karşı argümanlara cevap verebilmesi ve konuşmasını giriş-gelişme-sonuç dengesinde yürütebilmesidir.

İleri düzey (C1-C2) konuşma değerlendirmelerinde ise, adayın kod değiştirme veya eğitilmiş konuşma dili arasındaki geçişkenliği yönetme kabiliyeti bir ayıraç olarak ele alınabilir.⁸⁴ Bu seviyelerde adaydan, karmaşık konuları akademik bir üslupla tartışması, hipotezler kurması ve kültürel nüansları (deyimler, atasözleri, kültürel göndermeler) yerinde kullanması beklenir.⁸⁵ Rubriklerdeki doğruluk kriteri, bu seviyede yerini dilsel zenginlik ve uygunluk kavramına bırakmalıdır. Değerlendiriciler, adayların lehçe etkisindeki konuşmalarını yanlış olarak değil, MSA normlarından sapma derecesine ve iletişimin sürdürülebilirliğine göre puanlamalı,⁸⁶ üst düzey dilsel yetkinlik tanımlarında bütünleşik dil öğretim modelinin sunduğu esnekliği korumalıdır.⁸⁷ Bu yaklaşım, sınavın hem bilimsel ciddiyetini hem de dilin yaşayan gerçekliğini yansıtmasını sağlar. Sonuç olarak, yapılandırılmış görevler ve net tanımlanmış rubrikler aracılığıyla gerçekleştirilen bir konuşma sınavı, adayın Arapça dünyasındaki sesini küresel standartlarda tescilleyecektir.

⁸⁰ Swender, "Oral Proficiency Testing in the Real World", 520-526.

⁸¹ Weir, *Language Testing and Validation: An Evidence-Based Approach*, 145.

⁸² Kevin G. Knight, "Assessing Speaking Proficiency: A Test a Measure or a Concept?.", *The Modern Language Journal* 76/3 (1992), 288-298.

⁸³ Mahdi Alish, *Strategies for Proficiency* (New Haven: Yale University Press, 2005), 38.

⁸⁴ Willem J. M. Levelt, *Speaking: From Intention to Articulation* (Massachusetts: MIT Press, 1989), 158.

⁸⁵ Jack C. Richards, *Teaching Listening and Speaking* (New York: Cambridge University Press, 2008), 19.

⁸⁶ Kassem vd., *Handbook for Arabic Language Teaching Professionals in the 21st Century*, 204.

⁸⁷ Younes, *The Integrated Approach to Arabic Instruction*, 67.

4. 4. Yazma

Uluslararası bir Arapça yeterlilik sınavında yazma becerisinin ölçülmesi, Modern Standart Arapçanın sentaktik hassasiyeti ve yazım kurallarının karmaşıklığı nedeniyle, yapısal bir kademelendirme ve yüksek düzeyde nesnel bir değerlendirme protokolü gerektirir.⁸⁸ Yazma modülü, adayın sadece grafik sembolleri kâğıda dökme yetisini değil, düşüncelerini belirli bir türün normlarına göre organize etme, dil bilimsel yapıları bağlamsal doğruluğa dönüştürme ve hedef kitleye uygun üslup seçme becerisini ölçmelidir.⁸⁹ Bu süreçte, Arapçanın iki dilli (diglosik) yapısı ve konuşma dili ile yazma dili arasındaki dilsel mesafe, adayın başlangıç aşamasındaki üretim süreçlerinde doğrudan belirleyici bir faktör olarak dikkate alınmalıdır.⁹⁰ Seviye bazlı görev tasarımı, adayın bilişsel gelişimini ve dilsel üretim kapasitesini CEFR standartlarıyla uyumlu hale getirmeyi amaçlar. Bu bağlamda, başlangıç düzeylerinde (A1–A2) somut ve kişisel odaklı üretimler (kısa e-postalar, tebrik kartları, kişisel notlar) istenirken; orta düzeylerde (B1–B2) daha işlevsel ve mesleki odaklı (betimleyici raporlar, kısa bilgilendirici makaleler) görevler;⁹¹ ileri düzeylerde ise (C1–C2) yüksek bilişsel süreçler gerektiren argümantatif metinler, eleştirel analizler ve farklı kaynaklardan gelen bilgilerin harmanlandığı sentezleme görevleri sunulmalıdır.⁹² Kurumsallaşan bu ölçeklendirme politikalarında, CEFR standartlarının yerel ve küresel müfredat politikalarındaki sosyo-kültürel boyutları da test geçerliliği açısından göz önünde bulundurulmalıdır.⁹³

db | 731

Yazılı üretimin niteliğini tayin etmek amacıyla kullanılacak analitik dereceli puanlama anahtarları, çok boyutlu bir ölçme felsefesine dayanmalıdır.⁹⁴ Bu kapsamda belirlenen kriterlerden uygunluk, adayın verilen görevin tüm gerekliliklerini ve tonunu ne ölçüde karşıladığını; tutarlılık ve metin organizasyonu, yazının giriş, gelişme, sonuç dengesini, paragraflar

⁸⁸ Sara Cushing Weigle, *Assessing Writing* (New York: Cambridge University Press, 2002), 116.

⁸⁹ Barbara Kroll, *Second Language Writing: Research Insights for the Classroom* (New York: Cambridge University Press, 1990), 82.

⁹⁰ Saiegh-Haddad, "Linguistic Distance and Initial Reading Acquisition: The Case of Arabic Diglossia", 431-451.

⁹¹ Ken Hyland, *Second Language Writing* (New York: Cambridge University Press, 2003), 45.

⁹² Bachman - Palmer, *Language Assessment in Practice*, 295.

⁹³ Musthofa, "CEFR-Based Policy in Arabic Language Teaching and Cultural Dimension in Indonesian Islamic Higher Education", 96-107.

⁹⁴ Deborah Crusan, *Assessment in the Second Language Writing Classroom* (Michigan: University of Michigan Press, 2010), 92.

arasındaki mantıksal akışı ve tipolojik formuna uygunluğunu;⁹⁵ dil doğruluğu, MSA'nın morfolojik ve sentaktik kurallarının (i'rab, fiil-fail uyumu vb.) hatasız uygulanmasını; sözcük çeşitliliği, kelime hazinesinin genişliğini ve eş anlamlı/yan anlamlı sözcüklerin yerinde kullanımını; metin organizasyonu ise, yazının giriş, gelişme, sonuç dengesini ve tipolojik formuna (mektup, makale vb.) uygunluğunu değerlendirir. Bu kriterlerin yazılı üründeki sosyo-bilişsel geçerlilik katmanları ve puanlayıcılar arası tutarlılık mekanizmaları ölçme biliminin evrensel ilkeleriyle desteklenmelidir.⁹⁶ Özellikle Arapça yazımında noktalama işaretlerinin işlevsel kullanımı ve geleneksel metin yapılarından modern yazı normlarına geçişin yönetilmesi, ileri düzey puanlamalarında kritik birer ayrıraftır.⁹⁷

Ölçme sonuçlarının güvenilirliğini garanti altına almak için, her bir aday metni en az iki bağımsız ve eğitilmiş değerlendirici tarafından puanlanmalıdır. Değerlendiriciler arası tutarlılığın (inter-rater reliability) sağlanması ve subjektif yargıların puan üzerindeki etkisinin minimize edilmesi, performans tabanlı bu görevlerin yapı geçerliliği açısından zorunludur.⁹⁸ Eğer iki puanlayıcı arasındaki fark belirli bir eşik değeri (örneğin toplam puanın %10'u) aşarsa, üçüncü bir kıdemli değerlendiricinin hakemliğine başvurulmalıdır.⁹⁹ Ayrıca, puanlayıcı eğitimlerinde standardizasyon oturumları düzenlenerek, her seviyeye ait referans kağıtlar üzerinden ortak bir kanaat oluşturulması şarttır. Bu titiz süreç, adayın Arapça yazma yetkinliğinin tesadüfi değişkenlerden arındırılmış, bilimsel olarak doğrulanmış bir kanıtı olarak raporlanmasını sağlar. Sonuç olarak, bu yöntemle geliştirilen bir yazma testi, adayın sadece gramer bilgisini değil, düşüncelerini MSA normlarında birer entelektüel ürüne dönüştürme kapasitesini de tescil eder.¹⁰⁰

732 | db

5. Metin Seçimi, Kültürel Denge ve Adalet

Uluslararası bir Arapça yeterlilik sınavının başarısı, kullanılan metinlerin dilsel niteliği ile bu metinlerin temsil ettiği kültürel çeşitlilik arasındaki hassas dengenin kurulmasına bağlıdır. Arapça, kelimelerin belirli

⁹⁵ Christopher Tribble, *Writing* (New York: Cambridge University Press, 1996), 130.

⁹⁶ Stuart D. Shaw - Cyril J. Weir, *Examining Writing: Research and Practice in Assessing Writing* (New York: Cambridge University Press, 2007), 64.

⁹⁷ Batal, "Diglossia Proficiency: The Need for an Intermediate Level of Arabic", 185.

⁹⁸ Robert L. Johnson vd., *Assessing Performance: Designing Scoring and Validating Performance Tasks* (New York: Guilford Press, 2009), 154.

⁹⁹ Dana R. Ferris - John S. Hedgcock, *Teaching L2 Composition: Purpose Process and Practice* (New York: Routledge, 2014), 340.

¹⁰⁰ Kassem vd., *Handbook for Arabic Language Teaching Professionals in the 21st Century*, 215.

köklerden türetilerek kalıplara döküldüğü ve az sözcükle çok fazla şey anlatılabilen, anlam yoğunluğu yüksek bir dildir. Bu durum, bir metnin zorluk derecesini belirlerken sadece kelime sayısını değil, aynı zamanda morfolojik türetim sıklığını ve sentaktik karmaşıklığı da temel bir değişken haline getirir.¹⁰¹ Metin seçimi sürecinde benimsenen kuramsal çerçeve, CEFR ile tam uyumlu olmalı; ancak bu uyum, Arapçanın yapısal hususiyetleri göz önünde bulundurularak yeniden kalibre edilmelidir.¹⁰² A1-A2 seviyelerinde metinler, somut referans noktaları içeren, kısa ve yüksek bağlamsal ipuçlarına sahip yapıda kurgulanırken;¹⁰³ okuma testlerindeki görev tasarımı adayın temel bilgi bulma yetisini hedeflemelidir.¹⁰⁴ C1-C2 seviyelerinde disiplinler arası bir nitelik kazanan, felsefi, hukuki veya teknik derinliği olan, soyutlama düzeyi yüksek metinler seçilmelidir. Bu kademelendirme, adayın sadece metni okumasını değil, dilin derin yapısındaki anlamsal katmanları deşifre etmesini hedefler.¹⁰⁵

Dilsel bütünlüğün sağlanması noktasında, Arapçadaki bağdaşıklık ve tutarlılık unsurları, metin seçiminde birincil kriterdir. Arapça metinlerin karakteristik özelliği olan uzun isim tamlamaları (izâfet), karmaşık atıf sistemleri ve vav, fe, sümme gibi bağlaçlarla kurulan retorik akış, sınavın yapı geçerliliğini belirleyen unsurlardır.¹⁰⁶ Metinlerin tarz çeşitliliği ilkesi gereği; dijital medya içerikleri, resmi raporlar, akademik makaleler ve edebi pasajlar dengeli bir biçimde sunulmalı,¹⁰⁷ okuma testlerindeki metin mimarisi adayın satır arası anlam ilişkilerini saptayabileceği bir derinlikte kurgulanmalıdır.¹⁰⁸ Bu süreçte en kritik teknik eşik, lehçe müdahalesinden kaçınma prensibidir.¹⁰⁹ Sınavın uluslararası alanda güvenilir olması için tüm soruların sadece standart (resmi) Arapça ile hazırlanması gerekir.¹¹⁰ Dil bilimsel bir olgu olarak diglossianın getirdiği bu çok katmanlı yapı sınav adaletini zedelememelidir.¹¹¹ Böylece hiçbir aday, bölgesel bir

¹⁰¹ Grabe, *Reading in a Second Language: Moving from Theory to Practice*, 214.

¹⁰² Saiegh-Haddad, "Linguistic Distance and Initial Reading Acquisition: The Case of Arabic Diglossia", 431-451.

¹⁰³ Shuwairekh, *Vocabulary Selection for Arabic as a Foreign Language*, 58.

¹⁰⁴ Alderson, *Assessing Reading*, 167.

¹⁰⁵ "ACTFL Assigning CEFR Ratings to ACTFL Assessments"; Musthofa, "CEFR-Based Policy in Arabic Language Teaching and Cultural Dimension in Indonesian Islamic Higher Education", 96-107.

¹⁰⁶ Jarf, *Processing of Arabic Syntax*, 45.

¹⁰⁷ Kassem vd., *Handbook for Arabic Language Teaching Professionals in the 21st Century*, 175.

¹⁰⁸ Khalifa - Weir, *Examining Reading*, 62.

¹⁰⁹ Holes, *Modern Arabic: Structures Functions and Varieties*, 5.

¹¹⁰ Bachman - Palmer, *Language Assessment in Practice*, 256.

¹¹¹ Ferguson, "Diglossia", 325-340.

şivenin getirdiği avantajı veya dezavantajı yaşamaz ve herkes eşit şartlarda yarışmış olur. MSA'nın bu standartlaştırıcı rolü, farklı coğrafyalardan gelen adaylar için ortak bir dilsel zemin teşkil eder.

Kültürel tarafsızlık ve erişilebilirlik, sınavın etik ve pragmatik boyutunu oluşturur. Sınav materyalleri, belirli bir Arap ülkesinin yerel kültürüne, siyasi görüşüne veya dini yorumuna münhasır içeriklerden arındırılmalı; evrensel, bilimsel ve tarafsız bir içerik havuzu oluşturulmalıdır. Bu, kültürel duyarlılık analizi süreciyle desteklenmeli;¹¹² sınavda kullanılan bir metnin, Fas'tan Umman'a kadar geniş bir coğrafyadaki hiçbir aday için yabancılaştırıcı etkisi olmamalıdır.¹¹³ Ayrıca adalet ilkesi çerçevesinde, farklı sosyo-ekonomik ve eğitim geçmişlerine sahip adayların metinlerdeki arka plan bilgisine değil, sadece dilsel yetkinliğine dayalı olarak değerlendirilmesi esastır.¹¹⁴ Test maddelerinin etik standartları, adayların farklı kültürel ve coğrafi havzalardan gelmesiyle oluşabilecek sapmaları (construct-irrelevant variance) önleyecek psikometrik denetimlerle garanti altına alınmalıdır.¹¹⁵ Bu bilimsel yaklaşım, Arapça sınavını sadece bir dil testi olmaktan çıkarıp, küresel eğitim ve istihdam piyasasında güvenle kabul edilen, etik standartları yüksek bir ölçme-değerlendirme enstrümanına dönüştürecektir.

6. Arapçaya Özgü Zorluklar ve Çözüm Önerileri

Arapçanın uluslararası ölçekte standardize edilmiş bir ölçme-değerlendirme sistemine dahil edilmesi sürecinde karşılaşılan yapısal zorluklar, bu dilin tipolojik özelliklerinden kaynaklanan nev-i şahsına münhasır bir metodolojik yaklaşımı zorunlu kılmaktadır.¹¹⁶ Bu zorlukların başında dil ikiliği (diglossia) gelmektedir.¹¹⁷ Yani kitaplarda, resmi yazışmalarda kullanılan standart Arapça ile halkın günlük hayatta konuştuğu yerel lehçeler arasındaki büyük fark, sınavın adayın gerçek dil yeteneğini doğru ölçmesini zorlaştırmaktadır. Bir konuşma topluluğunda aynı dilin iki farklı varyantının (yüksek ve düşük kod) bir arada bulunması,¹¹⁸ adayların hangi dilsel bağlamda değerlendirileceği konusunda belirsizlik yaratmaktadır. Bu sorunun akademik çözümü, sınavın temel ontolojisini

¹¹² Antony John Kunnan, *Fairness and Validation in Language Assessment* (New York: Cambridge University Press, 2000), 35.

¹¹³ Taha-Thomure, "Arabic Language Education in the Arab World", 180-198.

¹¹⁴ Hudson, *Teaching Second Language Reading*, 240.

¹¹⁵ Fulcher, *Practical Language Testing*, 178.

¹¹⁶ Habash, "Introduction to Arabic Natural Language Processing".

¹¹⁷ Ferguson, "Diglossia", 325-340.

¹¹⁸ Holes, *Modern Arabic: Structures Functions and Varieties*, 12.

tamamen MSA üzerine inşa etmek ve aday el kitabında test edilen dil varyantını kesin çizgilerle tanımlamaktır.¹¹⁹ Ancak bu tanımlama, dilin yaşayan gerçekliğinden kopuk bir mekanik yapıya dönüşmemeli; sosyolinguistik katmanlar arasındaki geçişkenliği,¹²⁰ MSA'nın prestijli fonksiyonları ile uluslararası iletişimdeki rolü arasındaki dengeyi korumalıdır.¹²¹

İkinci büyük zorluk, Arapçanın kelime türetme gücüdür. Dil, sabit kökleri alıp belirli kalıplara yerleştirerek binlerce yeni sözcük ürettiği için çok geniş ve zengin bir yapıya sahiptir. Arapça, sınırlı sayıda üçlü veya dörtlü kökten binlerce kelime türetilebilen bir yapıya sahip olması nedeniyle,¹²² Hint-Avrupa dilleriyle karşılaştırıldığında metin başına düşen anlamsal yoğunluk açısından çok daha yüklü bir dildir. Bu durum, okuma ve dinleme metinlerinin zorluk derecesini belirlerken sadece kelime sayısının yeterli bir metrik olmadığını, morfolojik yoğunluğun da ölçülmesi gerektiğini göstermektedir.¹²³ Bir metindeki nadir vezinlerin kullanımı veya karmaşık türetimlerin sıklığı, adayın işleme yükünü artırmaktadır.¹²⁴ Çözüm olarak, test geliştirme sürecinde morfolojik analiz yazılımları kullanılmalı¹²⁵ ve CEFR düzeylerine göre her seviyede izin verilen morfolojik karmaşıklık eşikleri matematiksel olarak belirlenmelidir.¹²⁶

Üçüncü ve belki de en teknik zorluk, Arap alfabesinin doğasından kaynaklanan harekeleme/sesli harf eksikliği meselesidir. Arapça metinlerde ünlülerin (harekelerin) yazılmaması, okuma eylemini sadece bir kod çözme sürecinden çıkarıp, yüksek düzeyli bir sentaktik çıkarım işlemine dönüştürmektedir.¹²⁷ A1-B1 düzeyindeki adaylar, henüz dilin gramer sistemini içselleştirmedikleri için harekesiz metinlerde ciddi bir belirsizlik ile karşılaşır.¹²⁸ Bu durumun ölçme adaletini bozmaması için, başlangıç seviyelerinde (A1-A2) tam harekeli; orta seviyelerde (B1) kısmi harekeli (özellikle i'rabı belirleyen son harf harekeleri); ileri seviyelerde ise (B2-C2) tamamen harekesiz metin kullanımı bir standart haline getirilmelidir.

¹¹⁹ Batal, "Diglossia Proficiency: The Need for an Intermediate Level of Arabic", 178.

¹²⁰ Bedâvi, *Musteveyât el-'Arabiyye el-Mu'âsira fî Msr*, 102.

¹²¹ Reem Bassiouney, *Arabic Sociolinguistics* (Edinburg: Edinburgh University Press, 2009), 28.

¹²² Versteegh, *The Arabic Language*, 192.

¹²³ Ryding, *A Reference Grammar of Modern Standard Arabic*, 45.

¹²⁴ Alish, *Strategies for Proficiency*, 54.

¹²⁵ Kassem vd., *Handbook for Arabic Language Teaching Professionals in the 21st Century*, 185.

¹²⁶ Shuwairekh, *Vocabulary Selection for Arabic as a Foreign Language*, 62.

¹²⁷ Jarf, *Processing of Arabic Syntax*, 52.

¹²⁸ Elinor Saiegh-Haddad, "Correlates of Reading Fluency in Arabic", *Reading and Writing* 18/6 (2005), 559-582.

Bu kademeli yapılandırma, adayın test maddelerini çözerken karşılaşacağı zihinsel değişkenleri dengeler.¹²⁹ B2 ve üzeri seviyelerde harekesiz metin kullanılması, adayın kelimenin cümledeki konumuna göre doğru seslendirmeyi yapabilme yeteneğini ölçen doğal bir yapı geçerliliği unsuru olarak işlev görecektir.¹³⁰

7. Değerlendirme ve Puanlama Sistematiği

Uluslararası standartlarda bir Arapça sınavının psikometrik başarısı, öznel yargılardan arındırılmış ve ileri istatistiksel modellerle desteklenmiş bütüncül bir değerlendirme mimarisine dayanmalıdır. Üretken beceriler olan konuşma ve yazma modüllerinde, puanlama güvenilirliğini maksimize etmek amacıyla çoklu değerlendirici yaklaşımı benimsenmeli; her aday performansı, birbirinin notundan habersiz en az iki kıdemli puanlayıcı tarafından analitik rubrikler ışığında derecelendirilmelidir.¹³¹ Puanlayıcılar arasındaki tutarlılık, Cohen'in Kappa katsayısı gibi istatistiksel yöntemlerle denetlenmeli, anlamlı bir sapma durumunda üçüncü bir hakem değerlendiricinin görüşüne başvurulmalıdır.¹³² Bu nesnel süreç, Madde Tepki Kuramı (Item Response Theory - IRT) tabanlı madde analizleriyle tahkim edilerek, her bir sorunun zorluk ve ayırt edicilik parametreleri sürekli olarak kalibre edilmelidir.¹³³ IRT modellemesi, adayın yetenek düzeyini sadece doğru cevap sayısına göre değil, cevapladığı maddelerin karakteristik özelliklerine göre de hesaplayarak, farklı test formları arasında matematiksel bir denklik ve ölçme hassasiyeti sağlar.¹³⁴ Sınavın hedef dil kullanım alanı (TLU) ile olan bu psikometrik ve kuramsal uyumu, ölçme sonuçlarının küresel ölçekteki geçerliliğini garanti altına alırken,¹³⁵ testlerin uluslararası akreditasyon süreçleri, standardizasyon politikaları ve sosyo-kültürel dil gereksinimleri arasındaki dengenin de korunmasını sağlar.¹³⁶

Sınav sonuçlarının raporlanması aşamasında, adaya sunulan belge sadece kaba bir puan tablosundan ibaret kalmamalı, CEFR tanımlarıyla

¹²⁹ Alderson, *Assessing Reading*, 128.

¹³⁰ Khalifa - Weir, *Examining Reading*, 74.

¹³¹ Weir, *Language Testing and Validation: An Evidence-Based Approach*, 162.

¹³² Johnson vd., *Assessing Performance: Designing Scoring and Validating Performance Tasks*, 158.

¹³³ Lord, *Applications of Item Response Theory to Practical Testing Problems*, 45.

¹³⁴ Fulcher, *Practical Language Testing*, 215.

¹³⁵ Bachman - Palmer, *Language Assessment in Practice*, 312.

¹³⁶ "ACTFL Assigning CEFR Ratings to ACTFL Assessments"; Musthofa, "CEFR-Based Policy in Arabic Language Teaching and Cultural Dimension in Indonesian Islamic Higher Education", 96-107.

uyumlu, derinlemesine bir yeterlilik profili içermelidir.¹³⁷ Bu raporlama sistematığı, adayın genel dil seviyesinin yanı sıra okuma, dinleme, yazma ve konuşma becerilerinin her biri için ayrı düzey belirlemeleri yapmalı ve her bir alt beceride sergilenen dilsel performansla dair betimsel, yapıcı geri bildirimler sunmalıdır.¹³⁸ Örneğin, bir adayın akademik okuma becerisinde C1 düzeyinde olup sözlü üretimde B2 düzeyinde kalması gibi asimmetrik yetkinlikler, kariyer ve eğitim planlaması için somut verilerle belgelenmelidir.¹³⁹ Bu detaylı sonuç raporu, adayın güçlü yönlerini tescil ederken, gelişim alanlarını da CEFR'in eylem odaklı 'yapabilirim' ifadeleriyle ilişkilendirerek uluslararası kurumlar nezdinde şeffaf ve güvenilir bir dil pasaportu işlevi görmeli,¹⁴⁰ Arapça dil eğitimindeki kurumsal akrediteasyon ve standartlaştırma politikalarını desteklemelidir.¹⁴¹

8. Uygulama Önerileri

Uluslararası standartlarda bir Arapça yeterlilik sınavının hayata geçirilmesi, kuramsal tasarımın operasyonel bir ekosisteme dönüştürülmesini gerektiren çok katmanlı bir süreçtir. Bu sürecin başarıyla yönetilmesi için önerilen uygulama stratejileri, akademik geçerlilik ile teknolojik sürdürülebilirliği bir araya getirmelidir.

db | 737

8.1. Akademik ve Operasyonel Yapılanma Önerileri

Sınavın uluslararası meşruiyet kazanması için ilk adım, çok uluslu bir Arapça Ölçme ve Değerlendirme Konsorsiyumu kurulmasıdır. Bu yapı, sınavın içerik havuzunu denetlemeli ve farklı coğrafyalardaki dil kullanım varyantlarını izlemelidir. Uygulama aşamasında, pilot uygulamalardan elde edilen verilerin Madde Tepki Kuramı (IRT) algoritmaları ile analiz edilmesi ve madde güçlük derecelerinin CEFR seviyeleriyle korelasyonunun istatistiksel olarak doğrulanması şarttır.¹⁴² Ayrıca, sınavın güvenliğini sağlamak amacıyla biyometrik veri doğrulama sistemleri ve uzaktan gözetmenlik teknolojileri entegre edilmelidir.¹⁴³

¹³⁷ Council of Europe, *Common European Framework of Reference for Languages: Learning Teaching Assessment*, 174.

¹³⁸ Alderson, *Assessing Reading*, 232.

¹³⁹ Khalifa - Weir, *Examining Reading*, 210.

¹⁴⁰ Kunnan, *Fairness and Validation in Language Assessment*, 82.

¹⁴¹ Taha-Thomure, "Arabic Language Education in the Arab World", 180-198.

¹⁴² Fulcher, *Practical Language Testing*, 192.

¹⁴³ Bachman - Palmer, *Language Assessment in Practice*, 422.

8. 2. Teknolojik Entegrasyon ve Dijital Sertifikasyon

Geleneksel kâğıt-kalem sınavlarından ziyade, Bilgisayar Ortamında Bireyselleşmiş Test (Computer Adaptive Testing - CAT) modelinin uygulanması, hem adayın seviyesini daha kısa sürede tespit etmeyi sağlar hem de sınavın yapı geçerliliğini artırır.¹⁴⁴ Yazma ve konuşma modüllerinde, yapay zekâ tabanlı puanlama motorlarının insan puanlayıcılarla hibrit bir şekilde kullanılması, değerlendirme sürecindeki nesnelliği pekiştirecektir.¹⁴⁵ Sınav sonuçları, sadece bir puan olarak değil, adayın her bir alt becerideki güçlü ve zayıf yönlerini gösteren ayrıntılı bir yeterlilik profili olarak raporlanmalı¹⁴⁶ ve dijital sertifikalar blok zincir (blockchain) alt yapısı ile güvence altına alınarak, ölçme çıktılarının sosyo-bilişsel temelleri modern eğitimsel ölçme felsefesiyle tahkim edilmelidir.¹⁴⁷

8. 3. Sürekli İyileştirme ve Geribildirim Döngüsü

Uygulamanın son aşaması, sınavın etkisini ve paydaş memnuniyetini ölçen periyodik değerlendirme döngüleridir. Sınavın, Arapça öğretim müfredatları üzerindeki etkisi yakından izlenmeli;¹⁴⁸ adaylardan, eğitimcilerden ve işverenlerden gelen geri bildirimler doğrultusunda madde havuzu sürekli güncellenmelidir.¹⁴⁹ Ayrıca, sınavın adalet kriterlerini korumak için farklı etnik ve kültürel kökenlerden gelen aday grupları üzerinde düzenli olarak Madde Yanlılığı (DIF) analizleri yapılmalıdır.¹⁵⁰ Bu sürekli iyileştirme mekanizması, Arapça sınavının küresel eğitim standartlarının zirvesinde kalmasını sağlayacaktır.¹⁵¹

Sonuç

Arapçanın küresel statüsü, Birleşmiş Milletler'in altı resmî dilinden biri olması ve geniş coğrafi yayılımı dikkate alındığında, bu dilde yetkinliği ölçen sistemlerin dağınık, metodolojik olarak heterojen ve uluslararası çerçevelerle sınırlı düzeyde uyumlu olması önemli bir yapısal boşluğa işaret etmektedir. Mevcut sınavların bir kısmı belirli kurumsal veya ulusal

¹⁴⁴ Wainer, *Computerized Adaptive Testing: A Primer*, 34.

¹⁴⁵ Chapelle Carol - Dan Douglas, *Assessing Language Through Computer Technology* (New York: Cambridge University Press, 2006), 102.

¹⁴⁶ Alesh, *Strategies for Proficiency*, 52.

¹⁴⁷ Mislevy, *Sociocognitive Foundations of Educational Measurement*, 245.

¹⁴⁸ J. Charles Alderson - Dianne Wall, "Does Washback Exist?", *Applied Linguistics* 14/2 (1993), 115-129.

¹⁴⁹ Kassem vd., *Handbook for Arabic Language Teaching Professionals in the 21st Century*, 230.

¹⁵⁰ Kunnan, *Fairness and Validation in Language Assessment*, 65.

¹⁵¹ Taha-Thomure, "Arabic Language Education in the Arab World", 180-198.

ihtiyalara cevap vermekle birlikte, ierik geerlilięi, yapı geerlilięi ve lut geerlilięi bakımından kapsamlı ve kamusal olarak doęrulanmıř kanıtlar sunma noktasında yetersiz kalmaktadır. Bu durum, Arapa yeterlilięinin kresel iř gc piyasasında, akademik kabul srelerinde ve diplomatik baęlantılarda karřılařtırılabilir bir lut olarak kullanılmasını gleřtirmektedir.

alıřmada nerilen model, kuramsal temelini CEFR ve ACTFL gibi uluslararası referans sistemlerinden almakta; ancak Arapanın diglossia, kk-kalıp morfolojisi ve harekesiz yazı sistemi gibi tipolojik zelliklerini dikkate alarak bu ereveleri yeniden kalibre etmektedir. Bu baęlamda, sınavın ontolojik temelini Modern Standart Arapa zerine inřa edilmesi, standardizasyon ve adalet ilkesi aısından zorunlu grnmektedir. Bununla birlikte, ileri dzeylerde kontroll ve sınırlı lehe maruziyetinin otantiklik ilkesini glendirecek biimde tasarlanması, dilin sosyolinguistik gereklięi ile lme nesnelilięi arasında dengeli bir uzlařı sunabilir.

Drt temel becerinin lmnde, seviye bazlı grev tasarımı, biliřsel yk kademelendirmesi ve aık tanımlanmıř analitik rubrikler, sınavın i tutarlılıęını ve deęerlendiriciler arası gvenilirlięini artıracaktır. zellikle retken becerilerde oklu puanlayıcı modeli, istatistiksel tutarlılık katsayıları ile desteklenmeli; madde bankası ise Madde Tepki Kuramı temelli analizlerle srekli kalibre edilmelidir. Arapaya zg olarak, bařlangı seviyelerinde tam harekeli, orta seviyelerde kısmi harekeli ve ileri seviyelerde harekesiz metin kullanımı hem pedagojik hem de psikometrik aıdan yapı geerlilięini glendiren rasyonel bir zm olarak deęerlendirilebilir. Bu yaklařım, okuma ve dinleme srelerinde morfo-sentaktik ıkarım yeteneęini llebilir bir deęiřken haline getirecektir.

Teknolojik entegrasyon boyutunda, Bilgisayar Ortamında Bireyselleřmiř Test (CAT) modelinin benimsenmesi ve yapay zekâ destekli hibrit puanlama sistemlerinin kullanımı hem lme hassasiyetini artıracak hem de uygulama leęini geniřletecektir. Bununla birlikte, teknolojik araların kullanımı, etik ilke ve veri gvenlięi standartlarıyla uyumlu biimde dzenlenmeli; zellikle madde yanlılıęı (DIF) analizleri ve kltrel duyarlılık incelemeleri dzenli olarak gerekleřtirilmelidir. Sınavın sadece bir puan retmekle kalmayıp, CEFR temelli ayrıntılı yeterlilik profilleri sunması, onu bireysel geliřim planlaması ve kurumsal karar alma sreleri iin iřlevsel bir ara haline getirecektir.

Bu teknolojik ve kuramsal çerçevenin operasyonel bir düzleme aktarılması ve modelin uygulanabilir kılınması, belirli şartlar altında, somut araçlar ve net süreç yönetimleriyle mümkündür. Modelin iddia ettiği özgün değerlendirme ekosistemi, üç temel aşamadan oluşan bir uygulama matrisi (Action Plan) üzerinden yapılandırılmaktadır:

Şartlar ve Altyapı Araçları (Aşama I): Sınavın hayata geçirilmesi için öncelikli şart, açık kaynak kodlu doğal dil işleme (NLP) kütüphaneleriyle bütünleşmiş bir Sınav Yönetim Sistemi (LMS) altyapısının kurulmasıdır. Bu bağlamda, Arapçanın kök-kalıp morfolojisini analiz edebilen ve metinlerin “anlamsal-morfolojik yoğunluk skorunu” otomatik hesaplayan Arapça Kelime İşleme ve Frekans Analiz Yazılımları (Morphological Parsers), madde yazım sürecinin ana aracı olarak konumlandırılmalıdır. Soru bankasının kalibrasyonunda ise, maddelerin zorluk ve ayırt edicilik indekslerini tek, iki ve üç parametrelili lojistik modellere göre hesaplayan R tabanlı mirt (Multidimensional Item Response Theory) paketi veya eşdeğer psikometrik analiz araçları kullanılmalıdır.

740 | db

Süreç ve Uygulama Koşulları (Aşama II): Diglossia ve harekeleme dengesi, soru havuzunda matematiksel bir kota sistemiyle işletilecektir. A1-B1 düzeyindeki test maddelerinin veri tabanında %100 tam harekeli fontlarla saklanması; B2 düzeyinde i’rab harflerine odaklanan %25’lik kısmi harekeleme parametresinin devreye sokulması; C1-C2 seviyelerinde ise sistemin tamamen harekesiz metin akışını zorunlu kılması mimari bir şarttır. Üretken becerilerin (yazma ve konuşma) değerlendirilmesinde, iki bağımsız puanlayıcının sisteme girdiği analitik rubrik skorları arasındaki tutarlılık, g-katsayısı (Generalizability Theory) yazılımları ile anlık olarak denetlenmelidir. Puanlayıcılar arasındaki sapma payının %10 eşliğini aşması durumunda, sistem otomatik olarak üçüncü bir kıdemli hakem puanlayıcıya görevi devredecektir.

Somut Çıktılar ve Doğuracağı Sonuçlar (Aşama III): Geliştirilen bu somut araçlar ve süreçler neticesinde, adaya sınav sonunda sunulacak olan “Yeterlilik Profili”, geleneksel kaba puanların (raw scores) ötesinde, adayın morfo-sentaktik çıkarım becerisini, bağlamsal kelime hızını ve işlevsel iletişim kapasitesini radar grafikleri ve CEFR “yapabilirim” (can-do) ifadeleriyle dökümanate eden dijital bir karne (Blockchain-verified Badge) üretecektir. Bu durum, ölçme sonuçlarının manipülasyondan uzak, küresel istihdam ve akademi pazarlarında doğrudan kabul gören yüksek güvenilirlikli standart bir çıktıya dönüşmesini sağlayacaktır.

Bu bağlamda önerilen sınav modeli, Arapçayı yalnızca tarihsel ve kültürel bir miras dili olarak değil, çağdaş akademik, diplomatik ve ekonomik iletişim alanlarında aktif bir küresel dil olarak konumlandırmayı hedeflemektedir. Bu şekilde bilimsel geçerlik, ölçme adaleti ve dijital sürdürülebilirlik ilkeleri üzerine inşa edilecek bir Arapça yeterlilik sistemi, hem öğretim programlarını amaç odaklı biçimde dönüştürecek hem de uluslararası kurumlar nezdinde güvenilir bir referans standardı olacaktır. Dolayısıyla, disiplinler arası iş birliği, çok uluslu akademik konsorsiyumlar ve sürekli psikometrik izleme mekanizmalarıyla desteklenen bu girişim, Arapça dil öğretimi ve değerlendirmesi alanında paradigmatik bir dönüşüm potansiyeli taşımaktadır.

Kaynakça

- Albirini, Abdulkafi. "The Sociolinguistic Functions of Codeswitching between Standard Arabic and Dialectal Arabic". *Language in Society* 40/5 (2011), 713-735.
- Alderson, J. Charles. *Assessing Reading*. New York: Cambridge University Press, 2000.
- Alderson, J. Charles - Wall, Dianne. "Does Washback Exist?" *Applied Linguistics* 14/2 (1993), 115-129. <https://doi.org/https://doi.org/10.1093/applin/14.2.115>
- Alosh, Mahdi. *Strategies for Proficiency*. New Haven: Yale University Press, 2005.
- Bachman, Lyle. *Fundamental Considerations in Language Testing*. London: Oxford University Press, 1990.
- Bachman, Lyle - Palmer, Adrian. *Language Assessment in Practice*. İngiltere: Oxford University Press, 2010.
- Bachman, Lyle - Palmer, Adrian S. *Language Testing in Practice*. London: Oxford University Press, 1996.
- Bassiouney, Reem. *Arabic Sociolinguistics*. Edinburgh: Edinburgh University Press, 2009.
- Batal, Mahmoud al- (ed.). *Arabic as One Language: Integrating Dialect in the Arabic Language Curriculum*. Washington: Georgetown University Press, 2018.
- Batal, Mahmoud al-. "Diglossia Proficiency: The Need for an Intermediate Level of Arabic". *The Modern Language Journal* 75/2 (1992), 174.
- Batal, Mahmoud al- - Belnap, R. Kirk. *The Teaching and Learning of Arabic in the United States: Realities Needs and Future Directions*. New York: Routledge, 2006.
- Bedâvî, el-Sa'îd. M. *Mustevyât el-'Arabiyye el-Mu'âsira fî Mısır*. Kahire: Dâru'l-Me'arif, 1973.
- Brosh, Hezi. "Arabic Language Proficiency of Elementary School Teachers" 2/1 (2013), 154-165.
- Buck, Gary. *Assessing Listening*. New York: Cambridge University Press, 2001.
- Bygate, Martin. *Speaking*. London: Oxford University Press, 1987.
- Carol, Chapelle - Douglas, Dan. *Assessing Language Through Computer Technology*. New York: Cambridge University Press, 2006.
- Cohen, Andrew D. *Assessing Language Ability in the Classroom*. Boston: Heinle & Heinle, 1994.
- Council of Europe. *Common European Framework of Reference for Languages: Learning Teaching Assessment*. New York: Cambridge University Press, 2001.
- Crusan, Deborah. *Assessment in the Second Language Writing Classroom*. Michigan: University of Michigan Press, 2010.
- Ebel, Robert L. - Frisbie, David A. *Essentials of Educational Measurement*. New Delhi: Prentice Hall, 1991.

- Eisele, John. "Myth Ideology and Arabic Dialects". *Journal of Arabic Linguistics* 42 (2003), 198-215.
- Ferguson, Charles A. "Diglossia". *Word* 15/2 (1959), 325-340. <https://doi.org/https://doi.org/10.1080/00437956.1959.11659702>
- Ferris, Dana R. - Hedgcock, John S. *Teaching L2 Composition: Purpose Process and Practice*. New York: Routledge, 2014.
- Fulcher, Glenn. *Practical Language Testing*. London: Routledge, 2010.
- Grabe, William. *Reading in a Second Language: Moving from Theory to Practice*. New York: Cambridge University Press, 2009.
- Habash, Nizar Y. "Introduction to Arabic Natural Language Processing". (*Synthesis Lectures on Human Language Technologies*. ed. Graeme Hirst. Colombia: Columbia University, 2010.
- Haladyna, Thomas M. - Rodriguez, Michael C. *Developing and Validating Test Items*. New York: Routledge, 2013.
- Hansen, Janeette. *Second Language Listening: Theory and Practice*. New York: Cambridge University Press, 2010.
- Holes, Clive. *Modern Arabic: Structures Functions and Varieties*. Washington: Georgetown University Press, 2004.
- Hudson, Thom. *Teaching Second Language Reading*. London: Oxford University Press, 2007.
- Hughes, Arthur. *Testing for Language Teachers*. New York: Cambridge University Press, 2003.
- Hyland, Ken. *Second Language Writing*. New York: Cambridge University Press, 2003.
- Jarf, Reima al-. *Processing of Arabic Syntax*. Riyadh: King Saud University Press, 2001.
- Johnson, Robert L. vd. *Assessing Performance: Designing Scoring and Validating Performance Tasks*. New York: Guilford Press, 2009.
- Kassem, Wahhab vd. (ed.). *Handbook for Arabic Language Teaching Professionals in the 21st Century*. New York City: Lawrence Erlbaum Associates, 2006.
- Khalifa, Hanan - Weir, Ciyril. J. *Examining Reading*. Cambridge: Cambridge University Press, 2009.
- Khalil, Aziz. "Assessing Arabic Language Proficiency. Al-'Arabiyya". *Journal of the American Association of Teachers of Arabic* 43 (2010), 100-115.
- Knight, Kevin G. "Assessing Speaking Proficiency: A Test a Measure or a Concept?." *The Modern Language Journal* 76/3 (1992), 288-298.
- Kroll, Barbara. *Second Language Writing: Research Insights for the Classroom*. New York: Cambridge University Press, 1990.
- Kunnan, Antony John. *Fairness and Validation in Language Assessment*. New York: Cambridge University Press, 2000.
- Levelt, Willem J. M. *Speaking: From Intention to Articulation*. Massachusetts: Mit Press, 1989.
- Liskin-Gasparro, Judith e. "The ACTFL Proficiency Guidelines and the Oral Proficiency Interview: A Retrospect". *Foreign Language Annals* 36/4 (2003), 483-490.
- Little, David. "The Common European Framework of Reference for Languages: Content Purpose Origin Reception and Impact". *Language Teaching* 39/3 (2006), 167.
- Lord, Frederic M. *Applications of Item Response Theory to Practical Testing Problems*. New York: Routledge, 1980.
- Messick, Samuel. "Validity". *Creative Education* 3/6A (2012), 13-104.
- Mislevy, Robert J. *Sociocognitive Foundations of Educational Measurement*. New York: Routledge, 2018.
- Musthofa, Tulus. "CEFR-Based Policy in Arabic Language Teaching and Cultural Dimension in Indonesian Islamic Higher Education". *Eurasian Journal of Applied Linguistics* 8/2 (2022), 96-107.

- Palmer, A. S. *Testing Arabic as a Foreign Language*. Washington: Georgetown University Press, 2007.
- Parshall, Cynthia G. vd. *Practical Considerations in Computer-Based Testing*. Berlin: Springer International Publishing, 2002.
- Richards, Jack C. *Teaching Listening and Speaking*. New York: Cambridge University Press, 2008.
- Rost, Michael. *Teaching and Researching Listening*. New York: Routledge, 2011.
- Ryding, Karin. *Teaching and Learning Arabic as a Foreign Language: A Guide for Teachers*. Washington: Georgetown University Press, 2013.
- Ryding, Karin C. *A Reference Grammar of Modern Standard Arabic*. New York: Cambridge University Press, 2005.
- Saiegh-Haddad, E. "Linguistic Distance and Initial Reading Acquisition: The Case of Arabic Diglossia". *Applied Psycholinguistics* 24 (2003), 431-451. <https://doi.org/http://dx.doi.org/10.1017/S0142716403000225>
- Saiegh-Haddad, Elinor. "Correlates of Reading Fluency in Arabic". *Reading and Writing* 18/6 (2005), 559-582.
- Schulz, Renate. "The Challenge of Assessing Cultural Understanding in the Context of Foreign Language Instruction". *Foreign Language Annals* 40/1 (2006), 9-26.
- Shaw, Stuart D. - Weir, Cyril J. *Examining Writing: Research and Practice in Assessing Writing*. New York: Cambridge University Press, 2007.
- Shuwairekh, Saleh al-. *Vocabulary Selection for Arabic as a Foreign Language*. Riyadh: King Saud University Press, 2001.
- Swender, Elvira. "Oral Proficiency Testing in the Real World". *Foreign Language Annals* 36/4 (2003).
- Taha-Thomure, Hanada. "Arabic Language Education in the Arab World". *International Journal of West Asian Studies* 1 (2008), 180-198.
- Trentman, Emma. "Arabic Heritage Learners in Study Abroad". *Heritage Language Journal* 10/3 (2013), 432-451.
- Tribble, Christopher. *Writing*. New York: Cambridge University Press, 1996.
- Underhill, Nic. *Testing Spoken Language*. Cambridge University Press, 1987.
- Ur, Penny. *Teaching Listening Comprehension*. New York: Cambridge University Press, 1984.
- Vandergrift, Larry - Goh, Christine. *Teaching and Testing L2 Listening*. New York: Routledge, 2012.
- Versteegh, Kees. *The Arabic Language*. Edinburgh: Edinburgh University Press, 2014.
- Wainer, Howard. *Computerized Adaptive Testing: A Primer*. New York: Routledge, 2000.
- Weigle, Sara Cushing. *Assessing Writing*. New York: Cambridge University Press, 2002.
- Weir, Cyril. J. *Language Testing and Validation: An Evidence-Based Approach*. England: Palgrave Macmillan, 2005.
- Younes, Munther. *The Integrated Approach to Arabic Instruction*. London: Routledge, 2015.
- "ACTFL Assigning CEFR Ratings to ACTFL Assessments". Eriřim 10 Kasım 2025. <https://www.actfl.org/assessments/assigning-cefr-ratings-to-actfl-assessments>

