Türk Müziği, Eylül 2026
6(3), 251-261, e-ISSN: 2822-3195
turkmuzigidergisi.com

Genç Bilge yayıncılık
gencbilgeyayincilik.com

Araştırma Makalesi

Türk müziği araştırmalarında hesaplamalı etnomüzikoloji çalışmaları için güvenilir bir araç: MakamKit

Seyhan Canyakan[1]

Afyon Kocatepe Üniversitesi Devlet Konservatuvarı, Afyonkarahisar, Türkiye

Makale Bilgisi

Geliş: 21 Haziran 2026
Kabul: 31 Ağustos 2026
Online Yayın: 30 Eylül 2026

Anahtar Kelimeler
Hesaplamalı etnomüzikoloji
Karar perdesi
Koma
Makam kuramı
MakamKit
Nota aktarımı
Türk müziği

Öz

Bu makale, Türk makam müziği için geliştirilmiş MakamKit adlı analiz motorunun yeteneklerini tanıtır, ölçülmüş başarımını bildirir ve bu ölçümler sırasında ortaya çıkan bir aktarım sorununu tartışır. Motorun en temel özelliği, makam müziğinin kendi perde dünyasını bozmadan taşımasıdır: dijital notadan koma bilgisinin yeniden kurulması, 511 eser ve 129.151 nota üzerinde denetlenmiş ve binde iki hata payıyla neredeyse eksiksiz bulunmuştur; yani eser, motorun içinden geçerken Batı'nın on iki sesli düzenine yuvarlanmamaktadır. Bunun üzerine kurulan çözümleme katmanlarında motor, daha önce hiç görmediği 180 eserlik bir sınav takımında makamı eserlerin dörtte üçünden fazlasında ilk adayında, tanıdığı makam adlarıyla sınırlı bakıldığında ise onda dokuza yakınında doğru bilmektedir. Usul ve biçim tanıma, bu çalışmayla birlikte ilk kez sayısal olarak ölçülmüştür: on yedi usulde 2.586 eser üzerinde usulün ilk üç aday içinde doğru bulunma oranı yaklaşık yüzde doksan üç, yedi biçimde 1.820 eser üzerinde biçimin ilk adayda doğru bilinme oranı yaklaşık yüzde seksen yedidir. Ses kaydından makam tanıma, seyrin en çıplak duyulduğu solo taksimde güvenilir, geçkilerle örülü topluluk icralarında ise beklendiği gibi daha kırılgandır. Basılı nota tanımada ise motor, sayfadaki bir işaretin hangi perdeye karşılık geldiğini on notadan dokuzunda bir koma yanılma payı içinde doğru hesaplarken, o sayfadaki eserin makamını güvenilir biçimde adlandıramamakta; motor bunu gizlemek yerine, uzman onayı tamamlanmadan temiz çıktı vermeyi reddetmektedir. Motorun tasarım ilkesi budur: bilmediğinde susar, emin olmadığında aday sıralar, kaynağı yetersizse hiç cevap vermez. Bu titizlik sayesinde ortaya bir de beklenmedik bulgu çıkmıştır: aynı eserin iki farklı dijital nüshası, eserin karar perdesi konusunda her dört eserden birinden fazlasında birbiriyle anlaşamamaktadır. Bu, motorun değil külliyatın sorunudur ve bugüne dek bu ölçekte ölçülmemiştir; meşkten notaya, notadan baskıya uzanan aktarım zincirine dijital nüshanın da bir halka olarak eklendiğini gösterir. Makalenin katkısı iki yönlüdür: makam müziği için ölçülmüş başarımı ve kaynak-güven düzeni açıkça belgelenmiş bir araştırma aracı tanıtmak, ve bu aracın imkân verdiği bir aktarım bulgusunu alanın gündemine taşımak. Yazılımın iç tasarımı, lisans durumu gereği kapsam dışıdır.

Makaleye atıf için

Canyakan, S. (2026). Türk müziği araştırmalarında hesaplamalı etnomüzikoloji çalışmaları için güvenilir bir araç: MakamKit. *Türk Müziği, 6*(2), 251-261. DOI: https://doi.org/10.5281/zenodo.2

Giriş

Türk makam müziğinin bilgisayarla incelenmesi son yirmi yılda kendi geleneğini kurdu. Bozkurt'un (2008) kayıtlardan perde çıkarma yöntemi ile Gedik ve Bozkurt'un (2010) perde histogramı yaklaşımı bu incelemenin ölçüm zeminini kurdu; Bozkurt, Yarman, Karaosmanoğlu ve Akkoç'un (2009) çalışması kuram kitaplarındaki dizilerle icradaki fiilî perdelerin her noktada örtüşmediğini gösterdi. Karaosmanoğlu'nun (2012) hazırladığı SymbTr nota külliyatı ile Uyar,

[1] Doç.Dr., Afyon Kocatepe Üniversitesi Devlet Konservatuvarı, Afyonkarahisar, Türkiye: Email: scanyakan@aku.edu.tr ORCID: 0000-0001-6373-4245

Atlı, Şentürk, Bozkurt ve Serra'nın (2014) araştırma korpusu, bu tür incelemelerin binlerce eser üzerinde tekrarlanabilmesini sağladı. Şentürk, Holzapfel ve Serra (2014) notayla kaydı hizalamayı, Holzapfel (2014) yüzeydeki ritim ile usul arasındaki ilişkiyi, Ünal, Bozkurt ve Karaosmanoğlu (2014) notadan makam tanımayı, Bozkurt, Karaosmanoğlu, Karaçalı ve Ünal (2014) usul ve makam bilgisiyle yönlendirilen cümle bölmeyi ele aldı. Bozkurt, Ayangil ve Holzapfel'in (2014) derlemesi bu birikimi toparladı.

Bu birikimin bir eksiği vardı: araştırmacının eline geçen, uçtan uca çalışan, ne yaptığı ve ne kadar doğru yaptığı açıkça belgelenmiş bir araç. Yayımlanan yöntemlerin çoğu tek bir soruna odaklanır; bir eseri nota dosyasından alıp perdesini ölçmek, makamını önermek, usulünü ve biçimini tahmin etmek, kaydıyla karşılaştırmak ve bunların hepsini aynı kayıt düzeni içinde tutmak, ayrı ayrı çözülmüş parçaların bir araya getirilmesini gerektirir. MakamKit bu boşluk için geliştirilmiştir.

Aracın kurulmasında belirleyici olan, Türk müziğinin kendi aktarım tarihidir. Behar'ın (1998) ayrıntılı biçimde anlattığı gibi, Osmanlı/Türk müziğinde aktarımın esas yolu yüzyıllarca yazılı nota değil meşkti: eser, usta ile çırak arasında tekrar tekrar icra edilerek bedenden bedene geçerdi. Batı notasının bu geleneğe girişi, Ayangil'in (2008) izlediği üzere geç, tartışmalı ve hiçbir zaman tam olmayan bir süreçtir. Nota, meşkin taşıdığı bilginin ancak bir bölümünü kâğıda geçirebilmiştir. Signell'in (1977) seyir üzerine yaptığı betimleme, Popescu-Judetz'in (1996) terim tarihi ve Wright'ın (1992) erken kaynak incelemeleri, bu zincirin her halkasında bir yorum kararı bulunduğunu gösterir. Perde sistemi de bunun dışında değildir: bugün dijital külliyatlarda ortak dil olan Arel–Ezgi–Uzdilek çerçevesi (Arel, 1968) yirminci yüzyılda kurulmuş bir düzendir ve Bozkurt ve diğerlerinin (2009) ölçümleri onun icrayla her noktada örtüşmediğini ortaya koymuştur.

Bu tarih, bir aracın karşısına iki ödev koyar. Birincisi, makamın perde dünyasını bozmadan taşımak: bir eser motorun içinden geçerken komalarını yitirip Batı'nın on iki sesli düzenine yuvarlanıyorsa, çıkan sonuç artık o eser hakkında değildir. İkincisi, bilmediğini bilmek: notanın taşıyamadığı bilgi hesapla geri getirilemez, ve bir araç bunu itiraf etmiyorsa araştırmacıyı yanıltır. MakamKit'in tasarımı bu iki ödev üzerine kuruludur.

Aracın ölçülmesinde ise müzik bilgisayarlığı literatürünün uyarıları esas alınmıştır. Sturm (2013, 2014) yüksek başarı gösteren bir programın müzikal bir şey öğrendiğinin kanıtlanmış olmadığını, verinin tesadüfi düzenliliklerine tutunmuş olabileceğini uzun süredir hatırlatır. Kapoor ve Narayanan (2023), bildirilen makine başarımlarının önemli bir bölümünün, sınav sorularının çalışma kâğıdında bulunmasına benzer bir kusur yüzünden şişkin çıktığını göstermiştir. Panteli, Benetos ve Dixon (2018) aynı uyarıyı dünya müziği külliyatları için tekrarlar. Bu nedenle aşağıdaki bütün sayılar, motorun daha önce hiç görmediği eserler üzerinde ya da eser bütünlüğü korunarak kurulmuş sınavlarda ölçülmüş; her sayının hangi sınavdan çıktığı ayrıca belirtilmiştir.

MakamKit nedir, bu makalede ne kadarı anlatılmaktadır?

MakamKit'in beş işi vardır: nota dosyasından perde, süre ve bölüm bilgisini çıkarmak; ses kaydından perde seyrini çıkarıp karar perdesi ve makam adayları önermek; usul ve aksak yapıyı incelemek; güfteyi çözümlemek; eserler arası benzerlik, kümelenme ve ağ haritaları üretmek. Bunların yanında, basılı ya da taranmış notayı uzman denetimine açık bir inceleme paketine çeviren bir tanıma hattı ve uzman her kararı onaylamadan temiz çıktı vermeyen bir nota editörü bulunur. Editör, koma doğruluğunda seslendirme yapabildiği için, motorun önerisi kulakla da denetlenebilir.

Bu makale motorun *ne yaptığını ve ne kadar doğru yaptığını* anlatır. Programın iç tasarımı — hangi özelliklerden yararlandığı, hangi model ailesini kullandığı, eşiklerini nasıl türettiği, dosyalarının nasıl düzenlendiği — yazılımın lisans durumu gereği anlatılmamaktadır. Buna karşılık denemelerin nasıl kurulduğu eksiksiz verilmektedir; çünkü bu bilgi olmadan sonuç okunamaz.

Araştırma Problemi

Makale, motorun yeteneklerini yedi soruda ele alır:

- Makamın perde dünyası, motorun içinden geçerken bozulmadan korunuyor mu?
- Motor bir eserin makamını notadan ne kadar doğru tanıyor?
- Usulü nota yüzeyinden tanıyabiliyor mu; devri tanımakla devrin ağırlığını tanımak aynı şey mi?
- Bir eserin biçimini tanıyabiliyor mu; ve tanıyamadığında bunu söyleyecek kadar tedbirli mi?

- Bir ses kaydından makamı tanırken hangi icra biçiminde güvenilir?
- Basılı bir notayı okurken hangi bilgisi kullanılabilir, hangisi kullanılamaz?
- Bütün bunları ölçerken, külliyatın kendisi hakkında ne öğrendik?

Ortak Zemin

Veri. Nota üzerinden yapılan bütün denemeler SymbTr külliyatı (Karaosmanoğlu, 2012) ve ondan türetilen yerel çalışma kopyaları üzerinde yapılmıştır. Ses denemeleri solo taksim kayıtları ile CompMusic/OTMM çevresinde derlenmiş topluluk icralarına dayanır (Uyar vd., 2014). Basılı nota denemeleri, aynı eserlerin nota görüntüleri ile dijital karşılıklarının eşleştirilmesinden oluşan sabit bir listeyle yapılmıştır. Külliyatın kullanımı ticari olmayan araştırma amaçlıdır ve atıf zorunludur.

Örneklem uyarısı. Bu külliyatlar repertuvarın rastgele bir kesiti değildir. İçlerindeki eserler günümüze ulaşmış, meşk zincirinde yaşatılmış, notaya alınmış, basılmış, dijitalleştirilmiş ve bir besteciye bağlanmış eserlerdir; her aşama bir seçimdir (Behar, 1998; Ayangil, 2008). Aşağıdaki hiçbir sonuç repertuvarın tamamı hakkında bir yargı değildir.

Ölçmek ile tahmin etmek. Motorun verdiği sonuçlar iki ayrı cinstendir ve bu ayrım aracın tasarımına gömülüdür. Birincisi **ölçüm**dür: bir perdenin frekansı, iki perde arasındaki aralık, bir sesin komadaki sapması, notadaki bir sesin süresi, bir eserin son notası. Bunlar hesaplanır, tahmin edilmez; aynı dosyaya her zaman aynı cevabı verirler. Cetvelin doğruluk oranı olmaz.

İkincisi yargıdır: makam adı, usul adı, biçim adı, taranmış bir sayfadaki işaretin ne olduğu, bir kayıttaki karar perdesinin yeri. Bunlar tahmindir ve yanılma payı taşır.

Motorun sözleşmesi bu ayrım üzerine kuruludur: ölçümü verir, yargıyı adaylarıyla ve payıyla birlikte verir**,** ikisini karıştırmaz. Her cevap dört durumdan birine düşer — kesin, öneri (sıralı adaylar), çekimser ("bilmiyorum"), ret (hiç cevap yok). Sessizce yanlış cevap vermek bu düzende tanımlı bir durum değildir.

Notun hangi sınavdan geldiği. Bir yargının başarı oranı üç ayrı biçimde kaydedilir ve bunlar birbirinin yerine kullanılmaz: motorun kullanıcıya fiilen verdiği yoldan her tür eserle; aynı yoldan, yalnızca motorun tanıdığı adlar arasında kalan eserlerle; ve modelin kendi iç denemesinden aldığı not. Üçüncüsü gerçek bir sayıdır ama kullanıcının aldığı sonucu ölçmez, bu yüzden ayrı tutulur.

Bölüm 1. Makamın perde dünyası bozulmadan taşınıyor mu?

Soru

Makam müziğini bilgisayara aktarmanın ilk ve en temel engeli perdedir. Yaygın nota yazılımlarının çoğu Batı'nın on iki sesli düzenine göre çalışır; bir bakiye ile bir küçük mücennebi ayırt etmez, eseri en yakın yarım sese yuvarlar. Böyle bir yuvarlama olduğunda, sonrasında yapılan bütün çözümleme başka bir eser hakkındadır. Bu yüzden bir makam aracının ilk sınavı, komayı koruyup korumadığıdır.

Deneme ve bulgu

Dijital notada koma bilgisi her zaman doğrudan yazılı değildir; sayısal bir alandan yeniden kurulması gerekir. Bu yeniden kurulumun gerçek karşılığıyla örtüşmesi 511 eser ve 129.151 nota üzerinde tek tek denetlenmiş, binde iki hata payıyla neredeyse eksiksiz bulunmuştur. Yani eser motorun içinden geçerken perde kimliğini yitirmemektedir. Motor yine de bunun bir *okuma* değil bir *yeniden kurulum* olduğunu kaydında saklar; ayrım korunur.

İkinci bir sağlamlık ölçüsü, külliyatın açılabilirliğidir. SymbTr'nin XML külliyatındaki üç bin dosyanın tamamı tek tek açılmaya çalışılmış, yaklaşık yüzde doksan sekizi (%98) sorunsuz açılmıştır. Açılamayan altmış yedi dosyanın hepsi kaynakta yarıda kesilmiş, bozuk dosyalardır; motor bunlar için hiçbir sonuç üretmez, sessizce eksik bir çıktı vermez.

Bu iki ölçüm birlikte, motorun üzerine bina kurulabilecek bir zemin sunduğunu gösterir: eser bozulmadan giriyor, bozuk dosya sessizce geçmiyor. Sonraki bütün bölümler bu zeminin üzerinde durur.

Bölüm 2. Notadan makam tanıma

Soru

Makam müziğinde bir eserin kimliği büyük ölçüde nerede karar kıldığına bağlıdır. Karar perdesi, seyrin sonunda varılan yerdir ve makamı tanımanın en güçlü ipucudur. Nota bu bilgiyi açıkça taşır: eserin son notası çoğu zaman doğrudan

okunabilir. Ünal, Bozkurt ve Karaosmanoğlu (2014) notadan makam tanımanın kurulabileceğini göstermiştir. Buradaki soru, MakamKit'in bu işi ne kadar doğru yaptığıdır.

Deneme

Motorun daha önce hiç görmediği 180 eserlik sabit bir sınav takımı kullanılmıştır (deneme tarihi: 19 Haziran 2026). Model, kırk makam adı tanıyan sabit bir modeldir ve deneme sırasında hiç değiştirilmemiştir.

Bulgu

Tablo 1. Motorun hiç görmediği 180 eserde makam tanıma.

Eser kümesi	n	İlk aday doğru	İlk üç aday içinde doğru
Tümü	180	%77,8	%86,1
Motorun tanıdığı makam adları	161	%87,0	%96,3

Motor, kendi tanıdığı makam adları arasında kalan eserlerde on eserden yaklaşık dokuzunu ilk adayında doğru bilmekte; ilk üç adayına bakıldığında ise yirmi beş eserden yirmi dördünde doğru makam listede bulunmaktadır. Bu, makam müziği için bildirilen sonuçlar arasında güçlü bir düzeydir ve doğru kullanım biçimini de gösterir: motor üç aday üretir, uzman aralarından seçer.

Motorun tanımadığı makamlar ise ayrı bir başlıktır. Model kırk makam bilmektedir; rastgele seçilmiş bir SymbTr örnekleminde eserlerin yaklaşık beşte biri bu kırkın dışındadır ve bu eserlerde doğru cevap baştan imkânsızdır. Motor bunu gizlemez: her cevabında kaç makam tanıdığını ve hangilerini tanıdığını bildirir, böylece araştırmacı sonucu kullanmadan önce aradığı makamın listede olup olmadığını görebilir. Kapsam genişletme, aracın açık geliştirme başlığıdır.

Bu ne anlama geliyor

Bu bölümün asıl mesajı, sayının kendisi kadar sayının nasıl sunulduğudur. Motor tek bir ad dayatmaz; üç aday, tanıdığı makam listesi ve o listeye göre ölçülmüş başarı payı ile birlikte cevap verir. Sturm'un (2013) uyarısı bir sayının hangi sınavdan çıktığı söylenmedikçe kullanıcı hakkında hiçbir şey söylemeyeceği burada tasarım ilkesine çevrilmiştir. Motorun kendi iç denemesinden aldığı not, kullanıcının aldığı sonuçtan yüksektir; bu iki sayı motorun kayıt düzeninde ayrı ayrı tutulur ve iç deneme notu hiçbir zaman kullanıcı başarısı gibi gösterilmez.

Bölüm 3. Usul: yeni bir yetenek ve beklenmedik bir sınır

Soru

Usul, Türk müziğinde ritmin adı değil düzenidir. Bir usulü tanımak, kaç zamanlı olduğunu bilmek değil; düm ile tekin nereye düştüğünü, hangi vuruşun ağır hangisinin hafif olduğunu bilmektir. Holzapfel (2014) nota yüzeyindeki ritim ile usulün kendisi arasındaki mesafeyi ölçülebilir biçimde ele almış, Bozkurt ve diğerleri (2014) usul bilgisinin cümle bölmeyi yönlendirdiğini göstermiştir. Soru şudur: usulün adı notanın süre yazısından ne kadar çıkarılabilir?

Bu soru MakamKit için ayrıca önemlidir, çünkü motorda notadan usul tanıma bu çalışmayla birlikte ilk kez kurulmuş ve ölçülmüştür; daha önce motor bir nota dosyası için usul adı üretmiyordu.

Deneme

SymbTr metin külliyatından, en az otuz eseri bulunan usuller alınmıştır: on yedi usulde toplam 2.586 eser. Usul alanı bozuk olan 414 dosya elenmiştir. Eserler bütün hâlinde beş takıma bölünmüş, her seferinde bir takım sınav olarak ayrılmıştır (deneme tarihi: 20 Haziran 2026).

Bir noktanın altı çizilmelidir: külliyattaki dosyaların yalnızca yüzde biri kadarında ölçü rakamı yazılıdır. Motor usulü ölçü rakamından okuyamaz; süre ve vuruş dokusundan çıkarmak zorundadır. Bu, işi kolaylaştıran değil zorlaştıran bir koşuldur.

Bulgu

Motor eserlerin yaklaşık yüzde yetmişinde usulü ilk adayında doğru bilmekte, ilk üç adayı içinde doğru usul eserlerin yaklaşık yüzde doksan üçünde (%93) bulunmaktadır. Ölçü rakamı olmadan, yalnızca süre dokusundan elde edilen bu sonuç, aracın kullanılabilir bir usul önerisi üretebildiğini gösterir.

Usul usul bakıldığında aksak, aksak semâî ve semâî gibi devirler yüksek oranda yakalanmaktadır. Buna karşılık kapalı curcuna, nim sofyan ve evfer gibi usullerde oran belirgin biçimde düşmektedir.

Tablo 2. Usullere göre yakalama oranı (seçilmiş)

Usul	Oran	Usul	Oran
aksak	0,90	düyek	0,70
aksak semâî	0,87	müsemmen	0,67
semâî	0,86	sofyan	0,55
devrihindî	0,80	evfer	0,39
ağır aksak	0,78	nim sofyan	0,29
curcuna	0,74	kapalı curcuna	0,21

Bu düşüşün nedenini bulmak için ikinci bir deneme yapılmıştır: aynı tahminler, bu kez usul aileleri düzeyinde değerlendirilmiştir — ağır aksak ile aksak, kapalı curcuna ile curcuna, yürük semâî II ile yürük semâî aynı sayılmıştır.

Tablo 3. Aynı tahminlerin usul ailesi düzeyinde değerlendirilmesi.

Aile	Oran	Alt biçimin kendi oranı
curcuna	**0,89**	kapalı curcuna: 0,21
aksak	**0,92**	ağır aksak: 0,78
yürük semâî	0,67	yürük semâî II: 0,51

Bu ne anlama geliyor

İki tablo yan yana konduğunda örüntü açıktır: motor devri doğru tanımakta, devrin niteleyicisini tanıyamamaktadır. Kapalı curcuna adıyla beş eserden birinde bilinirken, aynı tahminler curcuna ailesi olarak sayıldığında on eserden dokuzunda doğrudur. Motor aslında "curcuna" demektedir; yanıldığı yer hangi curcuna olduğudur.

Bu, aracın bir kusuru değil, notanın bir sınırıdır — ve bulgunun kendisi etnomüzikolojik bir sonuçtur. Ağır, kapalı, yürük gibi niteleyiciler notanın süre yazısında değil, düm-tek ağırlık düzeninde, tempoda ve icra tavrında yaşar. Bir usulün "ağır" olması, sürelerin uzun yazılmasından ibaret değildir; vuruşların taşıdığı yükün başka dağılmasıdır. Kâğıt bu yükü kaydetmez; meşk kaydeder. Usta çırağa devrin kaç zamanlı olduğunu değil, nasıl basılacağını öğretir. Holzapfel'in (2014) kavramsal ayrımı burada, iki tablonun farkı olarak sayısallaşmaktadır.

Pratikte bu, aracın nasıl kullanılacağını da söyler: motorun ilk üç adayı doğru usulü çok yüksek oranda içerdiğinden, doğru kullanım tek ad dayatmak değil, üç adayı uzmanın önüne koymaktır. Motor zaten böyle davranmaktadır.

Kapsam beyanı

Bu deneme motorun kendi iç sınavıdır; kullanıcının kullandığı yoldan geçen bir ölçüm değildir. Deneme tarihinde bu yetenek henüz kullanıcıya açılmamıştı; sayılar tavan değeri sayılmalıdır. Usul adları külliyatın kendi dosya adlarından gelmekte, bozuk alanlar tahmin edilmek yerine reddedilmektedir.

Bölüm 4. Biçim tanıma ve tedbir

Soru

Şarkı, türkü, peşrev, saz semâîsi gibi adlar hem eserin yapısını hem repertuvardaki yerini anlatır. Stokes'un (1992) tür sınırlarının müzikal olduğu kadar toplumsal da olduğunu gösteren tartışması, bu kategorilerin masum olmadığını hatırlatır. Soru iki katmanlıdır: motor biçimi tanıyabilir mi, ve tanıyamadığında bunu söyleyecek kadar tedbirli midir? Usulde olduğu gibi, biçim tanıma da motorda ilk kez bu çalışmayla ölçülmüştür.

Deneme ve bulgu

SymbTr temelli bir çalışma külliyatından, en az elli eseri bulunan yedi biçim alınmıştır: 1.820 eser. Eser bütünlüğü korunarak beş takıma bölünmüştür (deneme tarihi: 21 Haziran 2026).

Motor eserlerin yaklaşık yüzde seksen yedisinde (%87) biçimi ilk adayında doğru bilmekte, ilk üç aday içinde doğru biçim neredeyse her eserde bulunmaktadır.

Tablo 4. Biçimlere göre doğruluk payı ve yakalama oranı.

Biçim	Eser	Bu adı verdiğinde haklı çıkma	Bu biçimdeki eserleri yakalama
şarkı	994	0,87	0,92
türkü	285	0,74	0,61
seyir	169	0,97	0,97
kupe	120	0,94	0,98
peşrev	93	0,87	0,82
saz semâîsi	86	0,93	0,95
aranağme	73	0,92	0,78

Peşrev ve saz semâîsi gibi klasik çalgısal biçimlerin yüksek oranda tanınması, aracın repertuvar çalışmalarında kullanılabilirliğini gösterir. Motorun en çok yanıldığı yer türküdür ve yanılgı büyük ölçüde şarkı yönündedir; bu şaşırtıcı değildir, çünkü şarkı ile türkü arasındaki sınır notanın yüzeyinde her zaman görünür bir sınır değildir.

Motorun tedbir düzeyi ayrıca ölçülmüştür. "Kesin" diyebilmesi için gereken eşik, yüzde doksan beş doğruluk hedefine göre ayarlanmıştır; bu eşikte motor eserlerin dörtte üçü hakkında kesin konuşmakta, kalanı sıralı aday listesi olarak sunmakta, en belirsiz durumlarda hiç konuşmamaktadır.

Tablo 5. Motorun ne zaman "kesin" dediği

	Değer
Hedeflenen doğruluk payı	0,95
Eşikte gerçekleşen doğruluk payı	0,95
Eşikte "kesin" denebilen eser oranı	0,73

Bu ne anlama geliyor

Tablo 5, aracın en ayırt edici özelliğini sayıya döker: motor yüksek doğrulukla geniş kapsamı aynı anda vaat etmez, ikisi arasındaki pazarlığı açıkça yapar ve sonucunu bildirir. Bir araştırmacı için bu, tek bir yüksek orandan daha kullanışlıdır; hangi eserde motora güvenip hangisinde kendi kararını vereceğini bilir.

Bir uyarı da eklenmelidir. Motorun en başarılı olduğu üç kategori — seyir, kupe, aranağme — külliyatın kendi düzenleme birimleridir ve geleneksel anlamda birer biçim sayılıp sayılmayacakları tartışmalıdır. Bu, makinenin müzikal biçimden çok külliyatın düzenleme alışkanlığını öğrenmiş olabileceği ihtimalini akılda tutmayı gerektirir; Sturm'un (2014) meşhur "at" benzetmesiyle anlattığı durumdur. Şüpheyi dağıtmanın yolu, külliyatların kendi kategorilerinin nasıl kurulduğunu belgelemekten geçer (Gebru vd., 2021; Bender & Friedman, 2018).

Bölüm 5. Ses kaydından makam

Soru

Bir makamı en açık gösteren icra biçimi taksimdir: usul yoktur, güfte yoktur, icracı makamın seyrini serbestçe kurar; makamın çıkıcı mı inici mi olduğu, asma duraklarının nerede olduğu, karara nasıl inildiği burada en çıplak hâliyle duyulur. Bir fasıl icrası ya da topluluk kaydı ise geçkileriyle ve toplu tavrıyla çok daha karışık bir dokudur. Soru, motorun hangi icra biçiminde güvenilir olduğudur.

Deneme ve bulgu

Tablo 6. Ses kaydından makam tanıma.

Kaynak	Örneklem	Tanınması istenen makam	İlk aday doğru	İlk üç aday içinde doğru
Solo taksim	280 kesit	6	%75,5	%94,5
Solo taksim (ortak karar)	277 kesit	6	%75,0	%95,0
Topluluk icrası	1.000 kayıt	20	%59,8	%85,1

Dar kapsamdaki altı makam şunlardır: hicaz, nihavend, rast, saba, segâh, uşşak. Motorun birden çok yöntemin ortak kararına dayanan yolu, kayıtların yaklaşık yüzde kırkında "kesin" diyebilmekte ve bu durumlarda on kayıttan dokuzunda doğru çıkmaktadır — yani motor her kayıtta konuşmamakta, konuştuğunda büyük ölçüde haklı çıkmaktadır.

Yirmi makamlık geniş kapsamda ilk üç aday içinde doğru makamın bulunma oranı yüzde seksen beşin üzerindedir; bin kayıtlık bir örneklemde bu, aracın geniş repertuvarlı arşiv taramalarında ön eleme için kullanılabileceğini gösterir.

Bu ne anlama geliyor

Tablodaki sayılar birbiriyle kıyaslanamaz: altı makam arasından doğru olanı bulmakla yirmi makam arasından bulmak aynı zorlukta değildir. Müzikal olarak okunduğunda tablo şunu söyler: motor, makamın seyrinin en açık göründüğü yerde en güvenilirdir. Bu, Signell'in (1977) seyrin makam kimliğindeki merkezî yerine dair betimlemesiyle uyumludur. Seyir ne kadar açık duyuluyorsa hem insan hem makine için makam o kadar kolay tanınmaktadır.

Bir sınır belirtilmelidir: ses tarafında motorun hiç görmediği eserlerden oluşan sabit bir sınav takımı henüz kurulmamıştır; yukarıdaki sayılar iç deneme değerleridir. Nota tarafında uygulanan disiplinin ses tarafına taşınması aracın açık geliştirme başlığıdır.

Bölüm 6. Basılı nota: ölçebildiği yer ile okuyamadığı yer

Soru

Türk müziği kaynaklarının büyük bölümü basılı ya da elyazması notadır. Bunların bilgisayara aktarılması, sayfadaki işaretlerin otomatik tanınması sorununu doğurur; bu iş Batı müziği için bile çözülmüş sayılmaz (Rebelo vd., 2012; Calvo-Zaragoza, Hajič ve Pacha, 2020). Türk müziğinde daha zordur, çünkü tanınması gereken yalnızca notanın yeri değil, komayı gösteren arıza işaretleridir.

Deneme ve bulgu

Motorun hiç görmediği 150 nota görüntüsü ile dijital karşılıklarından oluşan sabit bir liste kullanılmış, 147'si ölçülebilmiştir (deneme tarihi: 19 Haziran 2026). Bu listedeki hiçbir eser sistemin görsel örnek havuzunda bulunmamaktadır.

Sonuç iki uçludur. Motor, sayfada gördüğü bir işaretin hangi perdeye denk düştüğünü *on notadan dokuzunda bir koma yanılma payı içinde* doğru hesaplamaktadır. Buna karşılık o sayfadaki eserin makamını güvenilir biçimde adlandıramamaktadır. Nota nota bakıldığında hatanın nerede toplandığı da görülür: oktav hiç şaşırılmamış, nota adı neredeyse hiç yanlış okunmamış; kayıpların büyük bölümü hiç görülemeyen notalardan, kalanı ise koma okumasından gelmektedir.

Tablo 7. Basılı notadan tanıma (147 eser).

	Oran
Perde, bir koma yanılma payı içinde doğru	**%88,8**
Karar perdesi doğru	%30,6
Makam, ilk üç aday içinde doğru	%14,3
Makam, ilk aday doğru	%3,4

Bu ne anlama geliyor

Bu iki uç aynı cinsten sonuçlar değildir ve tek bir "tanıma başarısı" rakamına toplanamaz. Motor *ölçmeyi becerdiği yerde ölçmekte, okuyamadığı yerde ise okuduğunu iddia etmemektedir.* Bu, aracın tasarım ilkesinin en açık göründüğü yerdir: taramadan gelen hiçbir sonuç doğrudan temiz nota sayılmaz; her işaret uzman onaylayana kadar "incelenmemiş" kalır; temiz çıktı ancak bütün işaretler onaylandığında verilir, eksik incelemede sistem sessizce eksik bir sonuç üretmek yerine işlemi reddeder. Calvo-Zaragoza ve diğerlerinin (2020) tanıma ile kodlamayı ayırma çağrısının güven tarafındaki karşılığı budur.

Araştırmacı için pratik sonuç açıktır ve olumludur: tarama çıktısı makam kaynağı olarak kullanılamaz, ama nota nota denetlemek koşuluyla koma ölçümü kaynağı olarak kullanılabilir. Elyazması ve eski basma notaların perde içeriğini ölçmek isteyen bir çalışma için bu, hâlihazırda kullanılabilir bir imkândır.

Bölüm 7. Ölçerken ortaya çıkan: külliyatın kendi sorunu

Bulgu

Bölüm 2'nin denemeleri yapılırken beklenmedik bir durum fark edilmiştir. SymbTr külliyatında pek çok eser iki ayrı dijital nüshada bulunur: külliyatın kendi metin dosyası ve yaygın nota değişim biçimi olan MusicXML. Aynı sabit model, aynı eserleri bu iki nüshadan okuduğunda farklı sonuçlar vermektedir. Nedeni aranınca kaynak bulunmuştur: *iki nüsha eserin karar perdesi konusunda dört eserden birinden fazlasında birbiriyle anlaşamamaktadır.*

Tablo 8. İki nüshanın birbiriyle anlaşması (80 eserlik örneklem, aynı model)

Karşılaştırılan	n	anlaşan	oran
Karar (durak) perdesi	76	55	**%72,4**
İlk aday (makam adı)	76	45	%59,2

Kimi eserde fark bir beşli, kimi eserde tam bir oktavdır. Bunun sonucu doğrudan makam tanımaya yansımaktadır: metin nüshasından okunduğunda eserlerin dörtte üçünden fazlasında doğru bilinen makam, MusicXML nüshasından okunduğunda yarısına düşmektedir.

Bu ne anlama geliyor

Bu bulgu MakamKit'in bir kusuru değildir; motor iki nüshayı da aynı biçimde, aynı modelle işlemiştir. Bulgu *külliyatın kendisine* aittir ve bilinen literatürde bu ölçekte ölçülmemiştir. Ortaya çıkması, aracın karar perdesini bağımsız olarak ölçüp kaydetmesi ve iki kaynağı karşılaştırabilmesi sayesinde mümkün olmuştur; yani burada araç, kendi hakkında değil incelediği alan hakkında bilgi üretmiştir.

Etnomüzikolojik anlamı ise şudur. Behar'ın (1998) anlattığı meşk zinciri, eser notaya geçtiğinde bitmez; notadan basılı edisyona, edisyondan dijital dosyaya uzanan halkalar da birer aktarım halkasıdır ve her biri kendi kaybını üretir. Ayangil'in (2008) Batı notasının Türk müziğine girişinde saptadığı uyum sorunları, bugün dosya biçimleri düzeyinde sürmektedir. Bir eserin karar perdesinin bir nüshada rast, ötekinde gerdaniye okunması, o eseri okuyanın onu bambaşka bir perde bölgesinde işitmesi demektir.

Buradan alanı ilgilendiren somut bir öneri çıkar: bir çalışmada "SymbTr'den alınan şu eser" demek yetmez, hangi nüshadan alındığının belirtilmesi gerekir. Bu, künye titizliği değil yöntem zorunluluğudur. MakamKit bu nedenle her cevabında kaynağın türünü ve güven basamağını taşır; aynı eser için iki nüsha varsa hangisinin tercih edildiğini de bildirir.

Sonuç ve Tartışma

MakamKit, Türk makam müziği için üç şeyi bir arada sağlayan bir araştırma aracıdır: makamın perde dünyasını bozmadan taşımak, çözümleme katmanlarını uçtan uca çalıştırmak, ve her sonucun ne kadar güvenilir olduğunu sonucun kendisiyle birlikte bildirmek.

Ölçümler bunu doğrulamaktadır. Koma bilgisi 511 eser ve 129.151 nota üzerinde binde iki hata payıyla korunmakta; motor, daha önce hiç görmediği eserlerde makamı — kendi tanıdığı adlar arasında — on eserden dokuza yakınında ilk adayında bilmekte, ilk üç adayında ise neredeyse her eserde doğru makamı listelemektedir. Usul ve biçim tanıma bu çalışmayla ilk kez ölçülmüş, ilk üç aday düzeyinde sırasıyla yüzde doksan üç ve neredeyse tam doğruluk vermiştir. Ses tarafında motor, seyrin en açık duyulduğu solo taksimde güvenilir, geniş repertuvarlı arşivlerde ise ön eleme için kullanılabilir düzeydedir. Basılı notada perde ölçümü bugün kullanılabilir durumdadır.

Aracın ikinci katkısı tasarımındadır. Motor uydurmaz: bilmediğinde susar, emin olmadığında aday sıralar, kaynak yetersizse hiç cevap vermez, uzman onayı tamamlanmadan temiz nota çıkarmaz. Her cevabı hangi kaynaktan geldiğini, motorun kaç kategori tanıdığını ve o yolun ölçülmüş başarı payını taşır. Sturm'un (2013, 2014) ve Kapoor ile Narayanan'ın (2023) uyarıları burada bir uyarı olarak değil, bir davranış kuralı olarak yaşamaktadır. Bu, açık ve yeniden üretilebilir bilim çağrılarının (Wilkinson vd., 2016; Munafò vd., 2017; Nosek vd., 2018) makam müziği araçlarına uyarlanmasının somut bir denemesidir.

Üçüncü katkı, aracın imkân verdiği iki bulgudur; ikisi de motor hakkında değil, alan hakkındadır. *Birincisi* usuldedir. Motor devri doğru tanımakta, ama ağır/kapalı/yürük gibi niteleyicileri tanıyamamaktadır — ve bunun nedeni açıktır: bu niteleyiciler notanın süre yazısında değil, düm-tek ağırlık düzeninde ve icra tavrında yaşar. Kâğıt bu yükü kaydetmez, meşk kaydeder. Holzapfel'in (2014) kavramsal ayrımı burada iki tablonun farkı olarak sayısallaşmaktadır. Bu bulgunun etnomüzikolojik anlamı şudur: meşkin taşıdığı ama kâğıdın taşımadığı bilgi, hesapla da geri getirilemez; bir usulün ağırlığını bilmek için onu basmış birinden öğrenmek gerekir. Bu, hesaplamalı yöntemin acizliğini değil, iş bölümünü tarif eder.

İkincisi külliyattadır. Aynı eserin iki dijital nüshası karar perdesinde dört eserden birinden fazlasında anlaşamamaktadır. Bu, bilinen literatürde bu ölçekte ölçülmemiş bir bulgudur ve "veri kümesi" dediğimiz şeyin sabit bir zemin olmadığını gösterir. Panteli ve diğerlerinin (2018) dünya müziği külliyatları için işaret ettiği tasarım sorunları, Türk müziğinde eserin kimliği düzeyine kadar inmektedir. Bulgudan çıkan öneri somuttur: dijital nüshanın hangisi olduğu, artık bir künye ayrıntısı değil bir yöntem bilgisidir.

Motorun etnomüzikolojideki yeri bu tablodan açıkça çıkar: MakamKit *yorumcu değil ölçücüdür***.** Bir eserin makamına karar vermez; karar için gereken ölçümleri üretir, adaylarını sıralar, bilmediğini söyler ve gerektiğinde susar. Kararı, kaynağın tarihini, meşk silsilesini ve icra geleneğini bilen araştırmacı verir. Bu iş bölümü bir eksiklik değil, sınırını bilen bir aracın tanımıdır. Meşk geleneğinde de bilgi, kimden alındığı bilinmeden kabul edilmezdi; MakamKit'in kaynak-güven düzeni, o eski sorunun dijital ortamdaki karşılığıdır.

Sınırlılıklar ve Geliştirme Başlıkları

Sayıların kaynağı ve çıkar durumu. Bildirilen bütün sayılar yazarın geliştirdiği motorun kendi ölçüm kayıtlarından alınmıştır; hiçbiri bağımsız bir üçüncü tarafça yeniden koşulmamıştır. Denemelerin kuruluşu eksiksiz verilmiştir, ama bağımsız doğrulama yapılmamıştır.

Tarih. Sayılar 19–21 Haziran 2026 tarihli denemelere aittir; araç gelişmeye devam etmektedir.

Aktarım sınırı. Veri, özgün tarihsel kaynak değil, meşkten notaya, notadan baskıya, baskıdan dijital dosyaya uzanan zincirin son halkasıdır. Perde ölçümlerinin dayandığı Arel–Ezgi–Uzdilek çerçevesi yirminci yüzyılda kurulmuş bir düzendir (Arel, 1968); erken repertuvara geriye dönük uygulanması bir çağ karıştırma riski taşır (Bozkurt vd., 2009).

Kapsam. Nota tarafındaki makam modeli kırk makam, dar ses yolu altı makam, geniş ses yolu yirmi makam, usul modeli on yedi usul, biçim modeli yedi biçim tanımaktadır. Bu kapsamların dışındaki hiçbir kategori için iddia geçerli değildir. Kapsam genişletme açık bir geliştirme başlığıdır.

Sınav takımının bağımsızlığı. Bölüm 2'nin sınav listesi, gelecekte yapılacak bir yeniden eğitimin havuzunun içindedir. Bugünkü sabit modelin sayıları geçerlidir; model yeniden eğitilirse liste "görülmemiş" olma niteliğini yitirir. Bu, kapatılması planlanan bir açıktır.

Ses tarafında sabit sınav takımı. Bölüm 5'in sayıları iç deneme değerleridir; nota tarafındaki disiplinin ses tarafına taşınması gerekmektedir.

Kullanıcıya açılma. Bölüm 3 (usul) ve Bölüm 4 (biçim) yetenekleri, deneme tarihinde kullanıcının kullandığı yoldan geçmemişti; sayılar tavan değeri sayılmalıdır.

Etiketlerin niteliği. Usul ve biçim adları külliyatın kendi dosya adlarından gelmektedir; uzman elinden çıkmış olmakla birlikte bağımsız olarak yeniden denetlenmemiştir. Biçim alanı, geleneksel biçimlerle külliyat-içi kesit adlarını birbirine karıştırmaktadır.

Uzman denemesinin yokluğu. Motorun sonuçlarının uzman yargısıyla ne ölçüde örtüştüğünü sınayan, önceden planlanmış bir uzman denemesi henüz koşulmamıştır. Denemeler makinenin tutarlılığını ölçmekte, müzikal doğruluğu ölçmemektedir; bu, aracın bir sonraki adımıdır.

Beyanlar

Çıkar çatışması. Yazar, bu makalede tanıtılan yazılımın geliştiricisidir. Bildirilen sayıların tamamı, düşük olanlar dâhil, motorun kendi ölçüm kayıtlarından değiştirilmeden aktarılmıştır.

Veri ve materyal erişilebilirliği. Denemelerin dayandığı SymbTr külliyatı açık erişimlidir (Karaosmanoğlu, 2012) ve ticari olmayan araştırma amaçlı kullanıma, atıf koşuluyla açıktır. Denemelerin kuruluşu makalede eksiksiz verilmiştir. Yazılımın kaynak kodu ve iç tasarımı lisans durumu gereği paylaşılmamaktadır; bu, bildirilen davranış iddialarının başka araçlarla ve açık külliyatlar üzerinde sınanmasını engellemez.

Etik kurul onayı. Çalışma insan ya da hayvan katılımcı içermediğinden etik kurul onayı gerektirmemektedir.

Finansman. Çalışma herhangi bir dış finansman almamıştır.

Bilgilendirme

MakamKit'in ortaya çıkmasında Dr. Hasan Said Tortop'un fikrî katkısı etkili olmuştur. Bu değerli katkısından dolayı kendisine teşekkürlerimi sunarım. Dr. Seyhan Canyakan tarafından geliştirilen yazılım, Genç Bilge Yayıncılık bünyesindeki akademik dergilerde münhasır kullanım ve yayın izniyle tahsis edilmiştir ve yayınevinin yazılı onayı olmadan başka kişi veya kurumlarca bilimsel araştırma, analiz ve akademik yayın amacıyla kullanılamaz.

Kaynakça

Arel, H. S. (1968). *Türk mûsıkîsi nazariyatı dersleri*. İstanbul: Hüsnütabiat Matbaası.

Ayangil, R. (2008). Western notation in Turkish music. *Journal of the Royal Asiatic Society*, *18*(4), 401–447. https://doi.org/10.1017/S1356186308008651

Behar, C. (1998). *Aşk olmayınca meşk olmaz: Geleneksel Osmanlı/Türk müziğinde öğretim ve intikal*. İstanbul: Yapı Kredi Yayınları.

Bender, E. M., & Friedman, B. (2018). Data statements for natural language processing: Toward mitigating system bias and enabling better science. *Transactions of the Association for Computational Linguistics*, *6*, 587–604. https://doi.org/10.1162/tacl_a_00041

Bozkurt, B. (2008). An automatic pitch analysis method for Turkish maqam music. *Journal of New Music Research*, *37*(1), 1–13. https://doi.org/10.1080/09298210802259520

Bozkurt, B., Ayangil, R., & Holzapfel, A. (2014). Computational analysis of Turkish makam music: Review of state-of-the-art and challenges. *Journal of New Music Research*, *43*(1), 3–23. https://doi.org/10.1080/09298215.2013.865760

Bozkurt, B., Karaosmanoğlu, M. K., Karaçalı, B., & Ünal, E. (2014). Usul and makam driven automatic melodic segmentation for Turkish music. *Journal of New Music Research*, *43*(4), 375–389. https://doi.org/10.1080/09298215.2014.924535

Bozkurt, B., Yarman, O., Karaosmanoğlu, M. K., & Akkoç, C. (2009). Weighing diverse theoretical models on Turkish maqam music against pitch measurements: A comparison of peaks automatically derived from frequency histograms with proposed scale tones. *Journal of New Music Research*, *38*(1), 45–70. https://doi.org/10.1080/09298210903147673

Calvo-Zaragoza, J., Hajič, J., Jr., & Pacha, A. (2020). Understanding optical music recognition. *ACM Computing Surveys*, *53*(4), 1–35. https://doi.org/10.1145/3397499

Cornelis, O., Lesaffre, M., Moelants, D., & Leman, M. (2010). Access to ethnic music: Advances and perspectives in content-based music information retrieval. *Signal Processing*, *90*(4), 1008–1031. https://doi.org/10.1016/j.sigpro.2009.06.020

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé, H., III, & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM*, *64*(12), 86–92. https://doi.org/10.1145/3458723

Gedik, A. C., & Bozkurt, B. (2010). Pitch-frequency histogram-based music information retrieval for Turkish music. *Signal Processing*, *90*(4), 1049–1063. https://doi.org/10.1016/j.sigpro.2009.06.017

Gómez, E., Herrera, P., & Gómez-Martin, F. (2013). Computational ethnomusicology: Perspectives and challenges. *Journal of New Music Research*, *42*(2), 111–112. https://doi.org/10.1080/09298215.2013.818038

Holzapfel, A. (2014). Relation between surface rhythm and rhythmic modes in Turkish makam music. *Journal of New Music Research*, *44*(1), 25–38. https://doi.org/10.1080/09298215.2014.939661

Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns*, *4*(9), 100804. https://doi.org/10.1016/j.patter.2023.100804

Karaosmanoğlu, M. K. (2012). A Turkish makam music symbolic database for music information retrieval: SymbTr. *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR)*, 223–228.

Kim, J. W., Salamon, J., Li, P., & Bello, J. P. (2018). CREPE: A convolutional representation for pitch estimation. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 161–165. https://doi.org/10.1109/ICASSP.2018.8461329

Mauch, M., & Dixon, S. (2014). pYIN: A fundamental frequency estimator using probabilistic threshold distributions. *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 659–663. https://doi.org/10.1109/ICASSP.2014.6853678

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19)*, 220–229. https://doi.org/10.1145/3287560.3287596

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*, 0021. https://doi.org/10.1038/s41562-016-0021

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114

Panteli, M., Benetos, E., & Dixon, S. (2018). A review of manual and computational approaches for the study of world music corpora. *Journal of New Music Research*, *47*(2), 176–189. https://doi.org/10.1080/09298215.2017.1418896

Popescu-Judetz, E. (1996). *Meanings in Turkish musical culture*. İstanbul: Pan Yayıncılık.

Rebelo, A., Fujinaga, I., Paszkiewicz, F., Marçal, A. R. S., Guedes, C., & Cardoso, J. S. (2012). Optical music recognition: State-of-the-art and open issues. *International Journal of Multimedia Information Retrieval*, *1*(3), 173–190. https://doi.org/10.1007/s13735-012-0004-6

Signell, K. L. (1977). *Makam: Modal practice in Turkish art music*. Seattle: Asian Music Publications.

Stokes, M. (1992). *The arabesk debate: Music and musicians in modern Turkey*. Oxford: Clarendon Press. https://doi.org/10.1093/oso/9780198273677.001.0001

Sturm, B. L. (2013). Classification accuracy is not enough: On the evaluation of music genre recognition systems. *Journal of Intelligent Information Systems*, *41*(3), 371–406. https://doi.org/10.1007/s10844-013-0250-y

Sturm, B. L. (2014). A simple method to determine if a music information retrieval system is a "horse". *IEEE Transactions on Multimedia*, *16*(6), 1636–1644. https://doi.org/10.1109/TMM.2014.2330697

Şentürk, S., Holzapfel, A., & Serra, X. (2014). Linking scores and audio recordings in makam music of Turkey. *Journal of New Music Research*, *43*(1), 34–52. https://doi.org/10.1080/09298215.2013.864681

Uyar, B., Atlı, H. S., Şentürk, S., Bozkurt, B., & Serra, X. (2014). A corpus for computational research of Turkish makam music. *Proceedings of the 1st International Workshop on Digital Libraries for Musicology (DLfM '14)*, 1–7. https://doi.org/10.1145/2660168.2660174

Ünal, E., Bozkurt, B., & Karaosmanoğlu, M. K. (2014). A hierarchical approach to makam classification of Turkish makam music, using symbolic data. *Journal of New Music Research*, *43*(1), 132–146. https://doi.org/10.1080/09298215.2013.870211

W3C Music Notation Community Group. (2021). *MusicXML 4.0*. https://www.w3.org/2021/06/musicxml40/

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018. https://doi.org/10.1038/sdata.2016.18

Wright, O. (1992). *Words without songs: A musicological study of an early Ottoman anthology and its precursors*. London: School of Oriental and African Studies, University of London.

Turk Muzigi, September 2026
6(3), 251-261, e-ISSN: 2822-3195
turkmuzigidergisi.com

Genç Bilge yayıncılık
gencbilgeyayincilik.com

Research Article

MakamKit: A reliable tool for computational ethnomusicology research in Turkish music

Seyhan Canyakan[2]

Afyon Kocatepe University State Conservatory, Afyonkarahisar, Türkiye

Article Info

Recieved: 21 June 2026
Accepted 30 July 2026
Published: 30 September 2026

Keywords
Computational ethnomusicology
Makam theory
Usul
Comma
Karar pitch
MakamKit

2822-3195 / © 2026 TM
It is published by Genç Bilge Publishing and is an open-access journal distributed under the CC BY-NC-ND license.

Abstract

This article introduces MakamKit, an analysis engine developed for Turkish makam music, reports its measured performance, and discusses a transmission problem that surfaced during those measurements. The engine's most fundamental property is that it carries the pitch world of makam music without distortion: reconstruction of comma information from digital scores was audited over 511 pieces and 129,151 notes and found virtually complete, with an error rate of two per thousand — the music is not rounded to twelve-tone equal temperament as it passes through the system. On this foundation, the engine identifies the makam of more than three quarters of pieces at first candidate on a 180-piece set it had never seen, rising to nearly nine in ten when restricted to the makam names it knows. Usul and form recognition are here measured numerically for the first time: across 2,586 pieces in seventeen usuls the correct usul appears among the top three candidates in about 93% of cases, and across 1,820 pieces in seven forms the correct form is identified at first candidate in about 87%. Makam recognition from audio is reliable in solo taksim, where the seyir is most exposed, and predictably more fragile in ensemble performances woven with modulations. In printed-score recognition, the engine computes which pitch a written sign corresponds to within a one-comma margin in nine notes out of ten, while it cannot reliably name the makam of that page — and rather than concealing this, it refuses to emit a clean score until expert verification is complete. This is the engine's design principle: it stays silent when it does not know, ranks candidates when unsure, and declines entirely when the source is inadequate. That discipline yielded an unexpected finding: two digital copies of the same piece disagree about its final (karar) pitch in more than one case out of four. This is a problem of the corpus rather than of the engine, and it had not previously been measured at this scale; it shows that the digital copy is itself a link in the chain running from meşk to notation to print. The contribution is twofold: to introduce a research tool whose measured performance and source-trust architecture are documented openly, and to bring the transmission finding it enabled onto the field's agenda. The software's internal design is outside the article's scope owing to its licence status.

To cite this article

Canyakan, S. (2026). MakamKit: A reliable tool for computational ethnomusicology research in Turkish music. *Turk Muzigi, 6*(3), 251-261. DOI: https://doi.org/10.5281/zenodo.1

Introduction

The computational study of Turkish makam music has established its own tradition over the past two decades. Bozkurt's (2008) method for extracting pitch from recordings and Gedik and Bozkurt's (2010) pitch-histogram approach laid the measurement foundations of the field; the study by Bozkurt, Yarman, Karaosmanoğlu and Akkoç (2009) showed that the scales given in theory books do not coincide at every point with the pitches actually used in performance. The SymbTr score corpus compiled by Karaosmanoğlu (2012) and the research corpus of Uyar, Atlı, Şentürk, Bozkurt and Serra (2014) made it possible to repeat such investigations over thousands of pieces. Şentürk, Holzapfel and Serra (2014) addressed the alignment of scores with recordings, Holzapfel (2014) the relation between

[2] Assoc.Prof.Dr., Afyon Kocatepe Unievrsity State Conservatory, Afyonkarahisar, Türkiye: Email: scanyakan@aku.edu.tr ORCID: 0000-0001-6373-4245

surface rhythm and *usul*, Ünal, Bozkurt and Karaosmanoğlu (2014) makam recognition from scores, and Bozkurt, Karaosmanoğlu, Karaçalı and Ünal (2014) melodic segmentation guided by *usul* and makam knowledge. The review by Bozkurt, Ayangil and Holzapfel (2014) drew this body of work together.

One thing was missing from this accumulation: a tool that reaches the researcher's hands, works end to end, and openly documents what it does and how accurately it does it. Most published methods focus on a single problem; taking a piece from a score file, measuring its pitches, proposing its makam, estimating its *usul* and form, comparing it with a recording, and keeping all of this within a single record of provenance requires assembling parts that have been solved separately. MakamKit was developed for this gap.

What was decisive in building the tool is the transmission history of Turkish music itself. As Behar (1998) has described in detail, the principal channel of transmission in Ottoman/Turkish music was for centuries not written notation but *meşk*: the piece passed from body to body between master and apprentice, through repeated performance. The entry of Western notation into this tradition was, as Ayangil (2008) traces, late, contested, and never complete. Notation has been able to commit to paper only a part of what *meşk* carried. Signell's (1977) description of *seyir*, Popescu-Judetz's (1996) history of terminology, and Wright's (1992) studies of early sources each show, from a different direction, that every link in this chain involves an interpretive decision. The pitch system is no exception: the Arel–Ezgi–Uzdilek framework that is today the common language of digital corpora (Arel, 1968) is an order theorised in the twentieth century, and the measurements of Bozkurt et al. (2009) have shown that it does not coincide with performance at every point.

This history sets two obligations before any tool. The first is to carry the pitch world of makam without distortion: if a piece loses its commas as it passes through the engine and is rounded to the Western twelve-tone system, the result is no longer about that piece. The second is to know what it does not know: information that notation cannot carry cannot be recovered by computation, and a tool that does not admit this misleads the researcher. MakamKit's design rests on these two obligations.

In measuring the tool, the warnings of the music information retrieval literature have been taken as the baseline. Sturm (2013, 2014) has long reminded us that a program achieving high accuracy has not thereby been shown to have learned anything musical; it may have latched onto incidental regularities in the data. Kapoor and Narayanan (2023) have shown that a substantial share of reported machine performance figures is inflated by a flaw akin to having the exam questions on one's study sheet. Panteli, Benetos and Dixon (2018) repeat the same warning for world-music corpora. For this reason, every figure below was measured either on pieces the engine had never seen or in tests constructed with piece-level integrity preserved; and which test each figure comes from is stated explicitly.

What MakamKit is, and how much of it this article describes

MakamKit performs five tasks: extracting pitch, duration and section information from a score file; extracting the pitch trace from an audio recording and proposing a *karar* pitch and makam candidates; examining *usul* and *aksak* structure; analysing lyrics; and producing similarity, clustering, network and atlas outputs across pieces. Alongside these are a recognition pipeline that turns a printed or scanned score into a review package open to expert inspection, and a score editor that yields no clean output until the expert has approved every decision. Because the editor can play back with comma accuracy, the engine's proposal can also be checked by ear.

This article describes *what the engine does and how accurately it does it.* The internal design of the program — which features it draws on, which model family it uses, how it derives its thresholds, how its files are organised — is *not described*, owing to the software's licence status. By contrast, the construction of every experiment is given in full, because without that information the results cannot be read.

Research Problem

The article addresses the engine's capabilities through seven questions:

- Is the pitch world of makam preserved without distortion as it passes through the engine?
- How accurately does the engine recognise a piece's makam from the score?

- Can it recognise the *usul* from the surface of the score; and is recognising the cycle the same as recognising the weight of that cycle?
- Can it recognise a piece's form; and is it cautious enough to say so when it cannot?
- In recognising makam from an audio recording, in which mode of performance is it reliable?
- In reading a printed score, which of its outputs can be used and which cannot?
- While measuring all of this, what did we learn about the corpus itself?

Common Ground

Data. All score-based experiments were carried out on the SymbTr corpus (Karaosmanoğlu, 2012) and local working copies derived from it. The audio experiments rest on solo *taksim* recordings and on ensemble performances compiled around the CompMusic/OTMM effort (Uyar et al., 2014). The printed-score experiments used a fixed list pairing score images of pieces with their digital counterparts. Use of the corpus is for non-commercial research purposes, and attribution is required.

A caution about the sample. These corpora are not a random cross-section of the repertoire. The pieces in them have survived to the present, been kept alive in the *meşk* chain, been written down, printed, digitised and attributed to a composer; every stage is a selection (Behar, 1998; Ayangil, 2008). None of the results below is a judgement about the repertoire as a whole.

Measuring versus estimating. The results the engine produces are of two different kinds, and this distinction is built into the tool's design.

The first is *measurement:* the frequency of a pitch, the interval between two pitches, the comma deviation of a tone, the duration of a note in the score, the final note of a piece. These are computed, not guessed; they return the same answer for the same file every time. A ruler does not have an accuracy rate.

The second is *judgement:* the makam name, the *usul* name, the form name, what a sign on a scanned page is, where the *karar* pitch lies in a recording. These are estimates and carry a margin of error.

The engine's contract rests on this distinction: it delivers the measurement; it delivers the judgement together with its candidates and its margin; and it does not conflate the two**.** Every answer falls into one of four states — certain, suggestion (ranked candidates), abstention ("I do not know"), refusal (no answer at all). Silently returning a wrong answer is not a defined state in this scheme.

Which examination a figure comes from. The success rate of a judgement is recorded in three separate ways, and these are never used interchangeably: from the path the engine actually gives the user, over pieces of every kind; from the same path, restricted to pieces whose label lies among the names the engine knows; and the mark the model receives in its own internal examination. The third is a real figure, but it does not measure what the user receives, and is therefore kept separate.

Section 1. Is the pitch world of makam carried without distortion?

The question

The first and most basic obstacle in transferring makam music to a computer is pitch. Most common notation software operates on the Western twelve-tone system; it does not distinguish a *bakiye* from a *küçük mücennep*, and rounds the piece to the nearest semitone. Once such rounding occurs, every analysis performed afterwards is about a different piece. The first test of a makam tool is therefore whether it preserves the comma.

Experiment and finding

Comma information is not always written directly in a digital score; it must be reconstructed from a numerical field. The agreement of this reconstruction with the true value was audited note by note over 511 pieces and 129,151 notes and found *virtually complete, with an error rate of two per thousand*. The piece therefore does not lose its pitch identity as it passes through the engine. Even so, the engine records that this is a *reconstruction* rather than a *reading*; the distinction is preserved.

A second measure of robustness is the openability of the corpus. All three thousand files in SymbTr's XML corpus were opened one by one, and *about ninety-eight per cent (98%)* opened without problem. All sixty-seven files that failed are truncated or corrupt at source; the engine produces no result for them and does not silently emit an incomplete output.

Taken together, these two measurements show that the engine offers a foundation that can be built upon: the piece enters undistorted, and a corrupt file does not pass through unnoticed. Every subsequent section stands on this foundation.

Section 2. Makam recognition from the score

The question

In makam music, the identity of a piece depends largely on where it comes to rest. The *karar* pitch is the point reached at the end of the *seyir*, and it is the strongest clue for identifying the makam. Notation carries this information openly: the final note of the piece can usually be read directly. Ünal, Bozkurt and Karaosmanoğlu (2014) showed that makam recognition from scores can be established. The question here is how accurately MakamKit does this.

Experiment

A fixed test set of 180 pieces the engine had never seen before was used (experiment date: 19 June 2026). The model is a fixed model that knows forty makam names and was not altered during the experiment.

Finding

Table 1. Makam recognition on 180 pieces the engine had never seen

Set of pieces	n	First candidate correct	Correct within first three
All	180	77.8%	86.1%
Makam names known to the engine	161	87.0%	96.3%

Among pieces whose label lies within the makam names the engine knows, it identifies about nine out of ten correctly at its first candidate; looking at its first three candidates, the correct makam is on the list in twenty-four out of twenty-five pieces. This is a strong level among results reported for makam music, and it also shows the correct mode of use: the engine produces three candidates, and the expert chooses among them.

Makams the engine does not know are a separate matter. The model knows forty makams; in a randomly drawn SymbTr sample, about one fifth of the pieces fall outside these forty, and for them a correct answer is impossible by construction. The engine does not hide this: with every answer it reports how many makams it knows and which ones, so that the researcher can see, before using a result, whether the makam being sought is on the list. Extending this coverage is an openly stated development goal for the tool.

What this means

The main message of this section lies as much in how the figure is presented as in the figure itself. The engine does not impose a single name; it answers with three candidates, the list of makams it knows, and the success rate measured against that list. Sturm's (2013) warning — that a figure says nothing about the user unless the examination it came from is stated — has here been turned into a design principle. The mark the model receives in its own internal examination is higher than what the user receives; these two figures are kept in separate records, and the internal mark is never displayed as if it were the user's success rate.

Section 3. Usul: a new capability and an unexpected limit

The question

Usul in Turkish music is not the name of the rhythm but its order. To recognise an *usul* is not to know how many beats it has; it is to know where the *düm* and the *tek* fall, which stroke is heavy and which is light. Holzapfel (2014) treated in measurable terms the distance between the rhythm on the surface of the score and the *usul* itself, and Bozkurt et al. (2014) showed that knowledge of *usul* guides melodic segmentation. The question is this: how far can the name of an *usul* be derived from the durational writing of the score?

This question matters for MakamKit in a further respect, because *usul* recognition from scores was established and measured in the engine for the first time with this study; previously the engine produced no *usul* name for a score file.

Experiment

From the SymbTr text corpus, *usul*s with at least thirty pieces were taken: 2,586 usable pieces across seventeen *usul*s. 414 files with a corrupt *usul* field were excluded. The pieces were divided whole into five sets, one set being reserved as the examination each time (experiment date: 20 June 2026).

One point deserves emphasis: only about one per cent of the files in the corpus carry a written time signature. The engine therefore cannot read the *usul* from the time signature; it must derive it from the texture of durations and strokes. This is a condition that makes the task harder, not easier.

Finding

The engine identifies the *usul* correctly at its first candidate in about seventy per cent of pieces, and the correct usul appears among its first three candidates in about ninety-three per cent (93%) of pieces. Obtained without a time signature, from the durational texture alone, this result shows that the tool can produce a usable *usul* proposal.

Looking *usul* by *usul*, cycles such as *aksak*, *aksak semâî* and *semâî* are captured at a high rate. By contrast, the rate drops markedly for *usul*s such as *kapalı curcuna*, *nim sofyan* and *evfer*.

Table 2. Capture rate by *usul* (selected)

Usul	Rate	Usul	Rate
aksak	0.90	düyek	0.70
aksak semâî	0.87	müsemmen	0.67
semâî	0.86	sofyan	0.55
devrihindî	0.80	evfer	0.39
ağır aksak	0.78	nim sofyan	0.29
curcuna	0.74	**kapalı curcuna**	**0.21**

To find the reason for this drop, a second experiment was run: the same predictions were re-scored, this time at the level of *usul* families — *ağır aksak* counted as *aksak*, *kapalı curcuna* as *curcuna*, and *yürük semâî II* as *yürük semâî*.

Table 3. The same predictions evaluated at the level of *usul* families

Family	Rate	Rate of the sub-variant itself
curcuna	0.89	kapalı curcuna: 0.21
aksak	0.92	ağır aksak: 0.78
yürük semâî	0.67	yürük semâî II: 0.51

What this means

Placed side by side, the two tables reveal a clear pattern: the engine recognises the cycle correctly, but not the qualifier of the cycle. *Kapalı curcuna* is identified under its own name in only one piece out of five, whereas the same predictions, counted at the level of the *curcuna* family, are correct in nine pieces out of ten. The engine is in fact saying "curcuna"; where it errs is in *which* curcuna.

This is not a defect of the tool but a limit of notation — and the finding is itself an ethnomusicological result. Qualifiers such as *ağır*, *kapalı* and *yürük* live not in the durational writing of the score but in the *düm–tek* hierarchy of weight, in tempo, and in performance manner. That an *usul* is "*ağır*" is not merely a matter of durations being written long; it is a different distribution of the weight the strokes carry. Paper does not record this weight; *meşk* does. The master teaches the apprentice not how many beats the cycle has but how it is to be struck. Holzapfel's (2014) conceptual distinction is here quantified as the difference between two tables.

In practice this also indicates how the tool should be used: since the engine's first three candidates contain the correct *usul* at a very high rate, the correct use is not to impose a single name but to place three candidates before the expert. This is already how the engine behaves.

Statement of scope

This experiment is the engine's own internal examination; it is not a measurement taken along the path the user follows. At the time of the experiment this capability had not yet been released to users; the figures should be read as ceiling values. The *usul* labels come from the corpus's own file names, and corrupt fields are rejected rather than guessed.

Section 4. Form recognition and caution

The question

Names such as *şarkı*, *türkü*, *peşrev* and *saz semâîsi* describe both a piece's structure and its place in the repertoire. Stokes's (1992) discussion, showing that genre boundaries are as much social as musical, reminds us that these categories are not innocent. The question has two layers: can the engine recognise form, and is it cautious enough to say so when it cannot?

As with *usul*, form recognition was measured in the engine for the first time with this study.

Experiment and finding

From a SymbTr-based working corpus, seven forms with at least fifty pieces were taken: 1,820 pieces. The pieces were divided whole into five sets (experiment date: 21 June 2026).

The engine identifies the form correctly at its first candidate in about eighty-seven per cent (87%) of pieces, and the correct form appears among its first three candidates in nearly every piece.

Table 4. Precision and capture rate by form

Form	Pieces	Correct when it gives this name	Capture of pieces in this form
şarkı	994	0.87	0.92
türkü	285	0.74	0.61
seyir	169	0.97	0.97
kupe	120	0.94	0.98
peşrev	93	0.87	0.82
saz semâîsi	86	0.93	0.95
aranağme	73	0.92	0.78

That classical instrumental forms such as *peşrev* and *saz semâîsi* are recognised at a high rate indicates the tool's usability in repertoire studies. Where the engine errs most is *türkü*, and the error runs largely towards *şarkı*; this is unsurprising, since the boundary between *şarkı* and *türkü* is not always a visible boundary on the surface of the score. The engine's level of caution was also measured. The threshold required for it to say "certain" is calibrated to a target precision of ninety-five per cent; at that threshold the engine speaks with certainty about three quarters of the pieces, presents the rest as a ranked candidate list, and in the most uncertain cases does not speak at all.

Table 5. When the engine says "certain"

	Value
Target precision	0.95
Precision achieved at the threshold	0.95
Share of pieces declared "certain" at the threshold	0.73

What this means

Table 5 puts the tool's most distinctive property into figures: the engine does not promise high accuracy and broad coverage at the same time; it makes the bargain between them openly and reports the outcome. For a researcher this is more useful than a single high rate, because it tells them on which pieces to trust the engine and on which to decide for themselves.

A caution must also be added. The three categories on which the engine performs best — *seyir*, *kupe*, *aranağme* — are the corpus's own editorial units, and whether they count as "forms" in the traditional sense is debatable. This requires keeping in mind the possibility that the machine has learned the corpus's editorial habits rather than musical form; this is the situation Sturm (2014) describes with his well-known "horse" analogy. Dispelling the suspicion requires documenting how corpora construct their own categories (Gebru et al., 2021; Bender & Friedman, 2018).

Section 5. Makam from audio recordings

The question

The mode of performance that displays a makam most openly is the *taksim*: there is no *usul*, no text, and the performer builds the makam's *seyir* freely; whether the makam is ascending or descending, where its resting degrees lie, and how the descent to the *karar* is made are heard here at their most exposed. A *fasıl* performance or an ensemble recording,

by contrast, is a far denser texture with its modulations and its collective manner. The question is in which mode of performance the engine is reliable.

Experiment and finding

Table 6. Makam recognition from audio recordings

Source	Sample	Makams to be recognised	First candidate correct	Correct within first three
Solo *taksim*	280 excerpts	6	75.5%	94.5%
Solo *taksim* (joint decision)	277 excerpts	6	75.0%	95.0%
Ensemble performance	1,000 recordings	20	59.8%	85.1%

The six makams in the narrow scope are: *hicaz*, *nihavend*, *rast*, *saba*, *segâh*, *uşşak*. The engine's path based on the joint decision of several methods is able to say "certain" in about forty per cent of recordings, and in those cases it is correct in nine recordings out of ten — that is, the engine does not speak on every recording, and when it does speak it is largely right.

In the broad twenty-makam scope, the correct makam appears among the first three candidates in over eighty-five per cent of cases; on a sample of a thousand recordings this shows that the tool can be used for preliminary filtering in archive surveys of broad repertoire.

What this means

The figures in the table are not comparable with one another: finding the right answer among six makams is not as hard as finding it among twenty. Read musically, the table says this: the engine is most reliable where the makam's *seyir* is most visible. This accords with Signell's (1977) description of the central place of *seyir* in makam identity — the more openly the *seyir* is heard, the more easily the makam is recognised, for human and machine alike.

One limit should be stated: on the audio side, a fixed test set of pieces the engine has never seen has not yet been established; the figures above are internal-experiment values. Carrying the discipline applied on the score side over to the audio side is an openly stated development goal for the tool.

Section 6. Printed scores: where it can measure and where it cannot read

The question

A large part of the sources of Turkish music consists of printed or manuscript scores. Transferring them to a computer raises the problem of automatically recognising the signs on the page; this is not considered solved even for Western music (Rebelo et al., 2012; Calvo-Zaragoza, Hajič & Pacha, 2020). It is harder in Turkish music, because what must be recognised is not only the position of the note but the accidentals that indicate the comma.

Experiment and finding

A fixed list of 150 score images the engine had never seen, paired with their digital counterparts, was used; 147 could be measured (experiment date: 19 July 2026). None of the pieces on this list is present in the system's pool of visual examples.

The result has two poles. The engine computes which pitch a sign it sees on the page corresponds to, *within a one-comma margin, in nine notes out of ten.* By contrast, it cannot reliably name the makam of the piece on that page. Note-by-note inspection shows where the errors gather: the octave is never mistaken, the note name almost never misread; the bulk of the losses come from notes that are never detected, and the remainder from misreading the comma.

Table 7. Recognition from printed scores (147 pieces)

	Rate
Pitch correct within a one-comma margin	**88.8%**
Karar pitch correct	30.6%
Makam correct within first three candidates	14.3%
Makam correct at first candidate	3.4%

What this means

These two poles are not results of the same kind and cannot be collapsed into a single "recognition accuracy" figure. The engine *measures where it is able to measure, and does not claim to have read where it cannot read.* This is where the tool's design principle is most clearly visible: no result coming from a scan counts directly as a clean score; every sign

remains "unreviewed" until the expert approves it; a clean output is issued only when every sign has been approved, and in the case of incomplete review the system refuses the operation rather than silently producing an incomplete result. This is the trust-side counterpart of Calvo-Zaragoza et al.'s (2020) call to separate recognition from encoding. The practical consequence for the researcher is clear and positive: scan output cannot be used as a source for makam, but — on condition of note-by-note verification — it can be used as a source for comma measurement. For a study seeking to measure the pitch content of manuscripts and old printed scores, this is an already usable capability.

Section 7. What surfaced during measurement: a problem of the corpus itself

Finding

While the experiments of Section 2 were being carried out, an unexpected situation was noticed. Many pieces in the SymbTr corpus exist in two separate digital copies: the corpus's own text file and MusicXML, the common score-interchange format. The same fixed model gives different results when reading the same pieces from these two copies. When the cause was sought, the source was found: *the two copies disagree about a piece's karar pitch in more than one piece out of four.*

Table 8. Agreement between the two copies (80-piece sample, same model)

Quantity compared	n	agreeing	rate
Karar (final) pitch	76	55	**72.4%**
First candidate (makam name)	76	45	59.2%

In some pieces the difference is a fifth, in others a full octave. The consequence is reflected directly in makam recognition: the makam identified correctly in more than three quarters of pieces when read from the text copy drops to half when read from the MusicXML copy.

What this means

This finding is not a defect of MakamKit; the engine processed both copies in the same way, with the same model. The finding belongs to *the corpus itself,* and it has not been measured at this scale in the known literature. It came to light because the tool measures and records the *karar* pitch independently and can compare two sources — that is, here the tool produced knowledge not about itself but about the field it studies.

Its ethnomusicological meaning is this. The *meşk* chain Behar (1998) describes does not end when the piece is written down; the links running from notation to printed edition, and from edition to digital file, are themselves links of transmission, and each produces its own loss. The problems of fit that Ayangil (2008) identified in the entry of Western notation into Turkish music persist today at the level of file formats. For a piece's *karar* pitch to be read as *rast* in one copy and *gerdaniye* in another means that whoever reads it hears it in an entirely different pitch region.

From this follows a concrete recommendation for the field: in a study it is not enough to say "this piece taken from SymbTr"; which copy it was taken from must be stated. This is not bibliographic fastidiousness but methodological necessity. MakamKit therefore carries, with every answer, the type of the source and its trust level; where two copies exist for the same piece, it also reports which one was preferred.

Conclusion and Discussion

MakamKit is a research tool that provides three things together for Turkish makam music: carrying the pitch world of makam without distortion, running the analysis layers end to end, and reporting how reliable each result is together with the result itself.

The measurements bear this out. Comma information is preserved with an error rate of two per thousand over 511 pieces and 129,151 notes; on pieces it had never seen, the engine identifies the makam — among the names it knows — at its first candidate in nearly nine pieces out of ten, and lists the correct makam among its first three candidates in almost every piece. Usul and form recognition were measured for the first time in this study, yielding, at the level of the first three candidates, ninety-three per cent and near-complete accuracy respectively. On the audio side the engine is reliable in solo *taksim*, where the *seyir* is heard most openly, and usable for preliminary filtering in archives of broad repertoire. In printed scores, pitch measurement is usable today.

The tool's second contribution lies in its design. The engine does not fabricate: it stays silent when it does not know, ranks candidates when unsure, gives no answer when the source is inadequate, and issues no clean score until expert verification is complete. Every answer carries the source it came from, how many categories the engine knows, and the measured success rate of that path. The warnings of Sturm (2013, 2014) and of Kapoor and Narayanan (2023) live here not as warnings but as rules of behaviour. This is a concrete attempt to adapt the calls for open and reproducible science (Wilkinson et al., 2016; Munafò et al., 2017; Nosek et al., 2018) to the tools of makam music.

The third contribution consists of two findings the tool made possible; both are about the field rather than about the engine.

The first concerns *usul*. The engine recognises the cycle correctly but cannot recognise qualifiers such as *ağır*, *kapalı* and *yürük* — and the reason is clear: these qualifiers live not in the durational writing of the score but in the *düm–tek* hierarchy of weight and in performance manner. Paper does not record this weight; *meşk* does. Holzapfel's (2014) conceptual distinction is quantified here as the difference between two tables. The ethnomusicological meaning of the finding is this: what *meşk* carried but paper did not cannot be recovered by computation either; to know the weight of an *usul*, one must learn it from someone who has struck it. This describes not the impotence of computational method but a division of labour.

The second concerns the corpus. Two digital copies of the same piece disagree about the *karar* pitch in more than one piece out of four. This is a finding not measured at this scale in the known literature, and it shows that what we call a "dataset" is not a stable ground. The design problems Panteli et al. (2018) identify for world-music corpora reach, in Turkish music, down to the level of a piece's identity. The recommendation that follows is concrete: which digital copy was used is no longer a bibliographic detail but a piece of methodological information.

The engine's place in ethnomusicology follows clearly from this picture: MakamKit is *a measurer, not an interpreter*. It does not decide a piece's makam; it produces the measurements required for that decision, ranks its candidates, says when it does not know, and stays silent when it should. The decision is made by the researcher who knows the source's history, the *meşk* lineage, and the performance tradition. This division of labour is not a shortcoming but the definition of a tool that knows its limits. In the *meşk* tradition, too, knowledge was not accepted without knowing from whom it had been received; MakamKit's source-trust order is the digital counterpart of that old question.

Limitations and Development Goals

Source of the figures and competing interest. All reported figures are taken from the measurement records of the engine developed by the author; none has been re-run by an independent third party. The construction of the experiments is given in full, but independent verification has not been performed.

Date. The figures belong to experiments run between 19 and 21 June 2026; the tool continues to develop.

Limit of transmission. The data are not original historical sources but the last link in a chain running from *meşk* to notation, from notation to print, and from print to digital file. The Arel–Ezgi–Uzdilek framework on which the pitch measurements rest is an order theorised in the twentieth century (Arel, 1968); applying it retrospectively to the early repertoire carries a risk of anachronism (Bozkurt et al., 2009).

Coverage. The score-side makam model knows forty makams, the narrow audio path six makams, the broad audio path twenty makams, the *usul* model seventeen *usul*s, and the form model seven forms. No claim holds for any category outside these scopes. Extending coverage is an openly stated development goal.

Independence of the test set. The test list of Section 2 lies within the pool that a future retraining would use. The figures of today's fixed model are valid; but if the model is retrained, the list loses its status as "unseen". This is a gap that is planned to be closed.

No fixed test set on the audio side. The figures of Section 5 are internal-experiment values; the discipline applied on the score side needs to be carried over to the audio side.

Release to users. The *usul* capability (Section 3) and the form capability (Section 4) had not passed through the path the user follows at the time of the experiments; the figures should be read as ceiling values.

Nature of the labels. The *usul* and form names come from the corpus's own file names; although they originate from expert hands, they have not been independently re-verified. The form field also conflates traditional forms with corpus-internal excerpt labels.

No expert trial. A pre-planned expert trial testing how far the engine's results agree with expert judgement has not yet been run. The experiments measure the machine's consistency, not its musical correctness; this is the tool's next step.

Declarations

Competing interest. The author is the developer of the software introduced in this article. All reported figures, including the low ones, are conveyed unaltered from the engine's own measurement records.

Data and material availability. The SymbTr corpus on which the experiments rest is openly accessible (Karaosmanoğlu, 2012) and open to non-commercial research use with attribution. The construction of the experiments is given in full in the article. The software's source code and internal design are not shared, owing to its licence status; this does not prevent the reported claims about behaviour from being tested with other tools and on open corpora.

Ethics committee approval. The study involves no human or animal participants and therefore requires no ethics committee approval.

Funding. The study received no external funding.

Acknowledgement

MakamKit was conceived with the intellectual contribution of Dr. Hasan Said Tortop. I would like to express my gratitude to him for this valuable contribution. The software, developed by Dr. Seyhan Canyakan, has been exclusively authorized for use and publication in academic journals published by Genç Bilge Publishing and may not be used by any other person or institution for scientific research, analysis, or academic publication without the publisher's written permission.

References

Arel, H. S. (1968). *Türk mûsıkîsi nazariyatı dersleri*. İstanbul: Hüsnütabiat Matbaası.

Ayangil, R. (2008). Western notation in Turkish music. *Journal of the Royal Asiatic Society*, *18*(4), 401–447. https://doi.org/10.1017/S1356186308008651

Behar, C. (1998). *Aşk olmayınca meşk olmaz: Geleneksel Osmanlı/Türk müziğinde öğretim ve intikal*. İstanbul: Yapı Kredi Yayınları.

Bender, E. M., & Friedman, B. (2018). Data statements for natural language processing: Toward mitigating system bias and enabling better science. *Transactions of the Association for Computational Linguistics*, *6*, 587–604. https://doi.org/10.1162/tacl_a_00041

Bozkurt, B. (2008). An automatic pitch analysis method for Turkish maqam music. *Journal of New Music Research*, *37*(1), 1–13. https://doi.org/10.1080/09298210802259520

Bozkurt, B., Ayangil, R., & Holzapfel, A. (2014). Computational analysis of Turkish makam music: Review of state-of-the-art and challenges. *Journal of New Music Research*, *43*(1), 3–23. https://doi.org/10.1080/09298215.2013.865760

Bozkurt, B., Karaosmanoğlu, M. K., Karaçalı, B., & Ünal, E. (2014). Usul and makam driven automatic melodic segmentation for Turkish music. *Journal of New Music Research*, *43*(4), 375–389. https://doi.org/10.1080/09298215.2014.924535

Bozkurt, B., Yarman, O., Karaosmanoğlu, M. K., & Akkoç, C. (2009). Weighing diverse theoretical models on Turkish maqam music against pitch measurements: A comparison of peaks automatically derived from frequency histograms with proposed scale tones. *Journal of New Music Research*, *38*(1), 45–70. https://doi.org/10.1080/09298210903147673

Calvo-Zaragoza, J., Hajič, J., Jr., & Pacha, A. (2020). Understanding optical music recognition. *ACM Computing Surveys*, *53*(4), 1–35. https://doi.org/10.1145/3397499

Cornelis, O., Lesaffre, M., Moelants, D., & Leman, M. (2010). Access to ethnic music: Advances and perspectives in content-based music information retrieval. *Signal Processing*, *90*(4), 1008–1031. https://doi.org/10.1016/j.sigpro.2009.06.020

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé, H., III, & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM*, *64*(12), 86–92. https://doi.org/10.1145/3458723

Gedik, A. C., & Bozkurt, B. (2010). Pitch-frequency histogram-based music information retrieval for Turkish music. *Signal Processing*, *90*(4), 1049–1063. https://doi.org/10.1016/j.sigpro.2009.06.017

Gómez, E., Herrera, P., & Gómez-Martin, F. (2013). Computational ethnomusicology: Perspectives and challenges. *Journal of New Music Research*, *42*(2), 111–112. https://doi.org/10.1080/09298215.2013.818038

Holzapfel, A. (2014). Relation between surface rhythm and rhythmic modes in Turkish makam music. *Journal of New Music Research*, *44*(1), 25–38. https://doi.org/10.1080/09298215.2014.939661

Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns*, *4*(9), 100804. https://doi.org/10.1016/j.patter.2023.100804

Karaosmanoğlu, M. K. (2012). A Turkish makam music symbolic database for music information retrieval: SymbTr. *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR)*, 223–228.

Kim, J. W., Salamon, J., Li, P., & Bello, J. P. (2018). CREPE: A convolutional representation for pitch estimation. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 161–165. https://doi.org/10.1109/ICASSP.2018.8461329

Mauch, M., & Dixon, S. (2014). pYIN: A fundamental frequency estimator using probabilistic threshold distributions. *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 659–663. https://doi.org/10.1109/ICASSP.2014.6853678

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19)*, 220–229. https://doi.org/10.1145/3287560.3287596

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*, 0021. https://doi.org/10.1038/s41562-016-0021

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114

Panteli, M., Benetos, E., & Dixon, S. (2018). A review of manual and computational approaches for the study of world music corpora. *Journal of New Music Research*, *47*(2), 176–189. https://doi.org/10.1080/09298215.2017.1418896

Popescu-Judetz, E. (1996). *Meanings in Turkish musical culture*. İstanbul: Pan Yayıncılık.

Rebelo, A., Fujinaga, I., Paszkiewicz, F., Marçal, A. R. S., Guedes, C., & Cardoso, J. S. (2012). Optical music recognition: State-of-the-art and open issues. *International Journal of Multimedia Information Retrieval*, *1*(3), 173–190. https://doi.org/10.1007/s13735-012-0004-6

Signell, K. L. (1977). *Makam: Modal practice in Turkish art music*. Seattle: Asian Music Publications.

Stokes, M. (1992). *The arabesk debate: Music and musicians in modern Turkey*. Oxford: Clarendon Press. https://doi.org/10.1093/oso/9780198273677.001.0001

Sturm, B. L. (2013). Classification accuracy is not enough: On the evaluation of music genre recognition systems. *Journal of Intelligent Information Systems*, *41*(3), 371–406. https://doi.org/10.1007/s10844-013-0250-y

Sturm, B. L. (2014). A simple method to determine if a music information retrieval system is a "horse". *IEEE Transactions on Multimedia*, *16*(6), 1636–1644. https://doi.org/10.1109/TMM.2014.2330697

Şentürk, S., Holzapfel, A., & Serra, X. (2014). Linking scores and audio recordings in makam music of Turkey. *Journal of New Music Research*, *43*(1), 34–52. https://doi.org/10.1080/09298215.2013.864681

Uyar, B., Atlı, H. S., Şentürk, S., Bozkurt, B., & Serra, X. (2014). A corpus for computational research of Turkish makam music. *Proceedings of the 1st International Workshop on Digital Libraries for Musicology (DLfM '14)*, 1–7. https://doi.org/10.1145/2660168.2660174

Ünal, E., Bozkurt, B., & Karaosmanoğlu, M. K. (2014). A hierarchical approach to makam classification of Turkish makam music, using symbolic data. *Journal of New Music Research*, *43*(1), 132–146. https://doi.org/10.1080/09298215.2013.870211

W3C Music Notation Community Group. (2021). *MusicXML 4.0*. https://www.w3.org/2021/06/musicxml40/

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018. https://doi.org/10.1038/sdata.2016.18

Wright, O. (1992). *Words without songs: A musicological study of an early Ottoman anthology and its precursors*. London: School of Oriental and African Studies, University of London.