

Araştırma Makalesi

Feldman'ın Osmanlı-Türk müziği kuramlarının hesaplamalı etnomüzikolojik açıdan sınanması

Seyhan Canyakan¹

Afyon Kocatepe Üniversitesi Devlet Konservatuvarı, Afyonkarahisar, Türkiye

Makale Bilgisi

Geliş: 16 Haziran 2026

Kabul: 27 Temmuz 2026

Online Yayın: 30 Ekim 2026

Anahtar Kelimeler

Walter Feldman
Hesaplamalı etnomüzikoloji
Makam
Osmanlı-Türk müziği
Sembolik nota analizi
Seyir
SymbTr
Usul

Öz

Bu makale, Feldman'ın Osmanlı-Türk müziğine dair altı ayrı kuramının açık erişimli SymbTr sembolik nota külliyyatından (Karaosmanoğlu, 2012) alınan 149 peşrev ve saz semâîsi ile üç Mevlevî ayini örnekleminde, MakamKit adlı hesaplamalı müzikoloji analiz motoruyla sınamayı amaçlamaktadır. Her çalışma, Feldman'ın metninden önceden yön belirtilmiş hipotezler türetmiş olup, Spearman sıra korelasyonu, besteci-kümelî bootstrap, karma etkili modeller, kör kümeleme, besteci-gruplu çapraz doğrulama ve benzerlik sıralaması gibi yöntemlerle, çoklu karşılaştırma düzeltmeleri altında sınamıştır. Bulgular, altı çalışma boyunca ortak bir örüntü sergiler. Feldman'ın öngördüğü yönde kararlı bir kanıt hiçbir sınamada ortaya çıkmamış, buna karşılık en tutarlı ve tekrar eden sapma, terkidenden seyre geçişin iki bağımsız sınamasında (Çalışma 1 ve Çalışma 2) ortaya çıkan karar-yakınsama ölçütünün beklenenin tam tersi yönde hareket etmesidir. Bu ters yönlü sinyalin dışında kalan sınamalar ise ya hiçbir kararlı örüntü üretmemiş ya da hiç sınanamamıştır. Usul-ezgi eklemlenmesinin dönemsel değişimi (Çalışma 3) modern nota yüzeyinde iz bırakmamış; model-nazire ilişkisi (Çalışma 4) doğrulanmış bir keşif değil, yalnızca uzman incelemesine aday bir öncelik sırası üretmiş; dört dönemlik periyodizasyon (Çalışma 5) doğal bir küme yapısı olarak geri kazanılmamış ve yalnızca kaba bir erken/geç ayrımı için sınırlı istatistiksel kanıt bulunmuş; Birinci Selâm'daki biçim-üslup ilişkisi (Çalışma 6) ise kaynak nota eksikliği nedeniyle hiç sınanamamış ve bir Aşama 1 kayıtlı protokol olarak sunulmuştur. Sonuç olarak, bu altı çalışma Feldman'ın etnomüzikolojik tarihsel tezini çürütmez. Bulgular, modern sembolik aktarım örneklemin ve seçilen ölçüm vekillerinin sınırları içinde okunmalıdır. Araştırmamızın en önemli katkısı Osmanlı-Türk müziği kuramlarını denetlenebilir, yeniden üretilebilir bir hesaplamalı etnomüzikolojik çerçeveye taşımak ve bu çerçevenin nerede sessiz kaldığını dürüstçe belgelemesidir.

2822-3195 / © 2026 TM

Genç Bilge Yayıncılık tarafından
yayınlanmakta olup, açık erişim,
CC BY-NC-ND lisansına sahiptir.



Makaleye atıf için

Canyakan, S. (2026). Feldman'ın Osmanlı-Türk müziği kuramlarının hesaplamalı etnomüzikolojik açıdan sınanması. *Türk Müziği*, 6(Walter Feldman Özel Sayısı), 71-94. DOI: <https://doi.org/10.5281/zenodo.21673679>

Giriş

Walter Feldman'ın *Music of the Ottoman Court: Makam, Composition and the Early Ottoman Instrumental Repertoire* (2024) adlı çalışması, Osmanlı-Türk müziğinin biçimsel tarihine dair alan yazınındaki en kapsamlı ve en ayrıntılı anlatılardan birini sunar. Feldman'a göre erken dönem çalgısal repertuarı, özellikle peşrev ve saz semâîsi gibi formlar; kısa, tekrarlı ve görece kendi içine kapanan ezgi birimlerinden, yani terkidlerden kuruludur. Bu birimler eser içinde yer değiştirebilir, gerektiğinde başka eserlere de taşınabilir. On yedinci yüzyılın ikinci yarısından itibaren bu modüler örgü, Feldman'ın anlatısında yerini makamın kendine özgü melodik yürüngesi olan *seyir* içinde bütünleşen,

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

tek bir yay gibi gerilen bir sürekliliğe bırakır. Bu dönüşüm Feldman'ın metninde hiçbir zaman tek boyutlu bir eğri olarak sunulmaz: tonal merkezlerin sırası, usul devriyle ezginin eklemelenmesi, hane mülazime (teslim) düzeni ve nazire geleneği, bir eserin üslûbunca yeni bir eser yazma pratiğiyle birlikte örülen, çok bileşenli bir tarihsel süreç olarak kurulur. Feldman'ın 2022 tarihli *From Rumi to the Whirling Dervishes: Music, Poetry, and Mysticism in the Ottoman Empire* adlı çalışması ise bu tarihsel-biçimsel okumayı dinî-tören repertuvarına, özellikle Mevlevi ayininin Birinci Selâm bölümüne taşır; burada da biçimsel uzunluk ile bestecinin tarihsel konumu arasında yönlü bir ilişki öngörülür.

Feldman'ın çalışması, tarihsel müzikoloji ile etnomüzikolojinin kesişiminde durur. Onun yöntemi büyük ölçüde filolojik ve karşılaştırmalı-yazmasaldır: on yedinci ve on sekizinci yüzyıl edvâr literatürü, güfte mecmuaları ve az sayıda hayatta kalan erken nota kaynağı üzerinden kurulan bir yakın okuma pratiğine dayanır. Bu yöntem, Osmanlı musikisinin uzun süre yazılı notadan çok meşk (usta-çırak ilişkisi içinde bedenden bedene aktarım) yoluyla yaşadığı bir gelenekte, biçimsel değişimin izini sürebilmenin neredeyse tek yoludur. Etnomüzikolojinin sözlü/icra-merkezli bakışı ile tarihsel müzikolojinin belge-merkezli bakışı, Feldman'ın elinde birbirini dışlamaz. Yazılı kaynağın seyrekliği onu icra geleneğinin (meşk, taksim, gazel) kendi iç mantığına yakından bakmaya zorlarken, elindeki az sayıdaki erken nota da bu okumayı büsbütün spekülâtif olmaktan çıkarır. Ancak bu zenginlik, aynı zamanda bir sınama güçlüğü de doğurur. Feldman'ın önerdiği dönüşüm tezleri büyük ölçüde nesir kanıtına ve seçilmiş örnek çözümlemelerine dayanır. Tezlerin geniş bir repertuar örnekleme üzerinde sistematik, denetlenebilir ve yeniden üretilebilir biçimde sınanıp sınamayacağı sorusu literatürde büyük ölçüde açık kalmıştır. Bu açık soru, tek bir eserin yakın okumasıyla değil, çok sayıda eserin karşılaştırmalı ve ölçülebilir bir çerçevede incelenmesiyle cevaplanabilecek türdendir. Tam da bu noktada hesaplamalı yöntemler, filolojik okumanın yerine değil, onun yanına, tamamlayıcı bir araç olarak girer.

Hesaplamalı müzikoloji, son on yılda Türk makam musikisi üzerine giderek genişleyen bir yöntem dağarcığı geliştirmiştir. Bozkurt, Ayangil ve Holzapfel'in (2014) alanın durumunu gözden geçiren çalışması, bu birikimin ağırlıklı olarak makam tanıma, perde/koma sistemi çözümlemesi ve yüzeysel benzerlik ölçümü gibi sorular etrafında yoğunlaştığını, buna karşılık tarihsel dönüşüm iddialarının sistematik biçimde sınanmasına yönelik yöntemlerin sınırlı kaldığını göstermektedir. Aynı çalışma, sembolik veriden çıkarılabilecek bilginin icra pratiğinin taşıdığı bilginin ancak bir kesiti olduğunu da vurgular. Bu uyarı, el alınan altı çalışmanın her birinde yöntem ve tartışma bölümlerinin ortak bir referans noktasıdır. Bu makale, tam da bu boşlukta dururarak tarihsel etnomüzikolojinin büyük anlatısını, hesaplamalı etnomüzikolojinin ölçüm araçlarıyla, denetlenebilir bir sınama zincirine dönüştürmeyi dener.

Bu sınanmanın veri kaynağı, SymbTr olup, Karaosmanoğlu'nun (2012) tanıttığı, Türk makam müziği için hazırlanmış açık erişimli sembolik nota külliyyatıdır. SymbTr, geleneksel repertuvarıdan çok sayıda eseri, perde, süre, usul ve bölüm yapısı bilgisiyle birlikte makine-okunur biçimde barındırır ve makam müziği bilgi erişimi ile hesaplamalı çözümleme araştırmalarının önemli bir bölümüne temel oluşturmuştur. Kullanım koşulları ticari olmayan araştırma amaçlı olup atıf zorunludur. Bu makalede kullanılan örneklem, SymbTr türevi bir koleksiyondan seçilen peşrev ve saz semâîleriyle sınırlıdır ve olasılıklı bir tarihsel örneklem değildir. Günümüze ulaşma, meşk zincirinde yaşatılma, notaya aktarılma, edisyona girme, dijitalleştirilme ve besteciye atfedilme gibi art arda gelen süzgeçlerden geçmiş, seçilmiş bir aktarım örnekleimidir. Özgün tarihsel yazmaların kendisi değil, bu zincirin son halkasıdır. Bu temel uyarı burada bir kez ayrıntılı biçimde kurulmaktadır. Ele alınan altı çalışma, bu ortak zemine kendi örnekleme özgü ek sınırlılıkları ekleyerek tartışmaktadır. Bu külliyyattan Çalışma 1-5 için ortak biçimde süzülen örneklem 149 peşrev ve saz semâîsidir; eserler besteci atıflarından türetilen dört kaba dönem katmanına (1650 öncesi, 1650-1750, 1750-1850, 1850 sonrası) yerleştirilmiştir. Bu dört katman, Feldman'ın kendi metninde sekiz alt döneme kadar inceltilecek bir dönemlendirmenin kasıtlı bir kabalaştırmasıdır. Katman sınırları ile Feldman'ın anlatısındaki dönem geçişleri arasında sayfa düzeyinde birebir bir eşleme bu makalede kurulmamıştır, dolayısıyla dönem katmanları üzerinden yapılan her sınama Feldman'ın dönemlendirmesinin ancak kaba bir yaklaşıklıkla okunmalıdır. Bu katmanlaşma eserlerin kesin bestelenme tarihini değil, besteci yaşam aralıklarından türetilen bir yaklaşık atfı temsil eder ve örneklemin ağırlık merkezi geç döneme (1850 sonrası) kaymıştır. Bu dengesizlik, altı çalışmanın ortak bir yöntemsel sınırlılığıdır. Çalışma 6 ise bu külliyyatın dışına çıkar: Mevlevi ayini repertuarı SymbTr'nin kapsamında olmadığından, kaynak-izlenebilir bir nota temini süreci ayrıca tanımlanmıştır.

Bu külliyat üzerinde çalışan hesaplamalı altyapı MakamKit'tir: Türk halk ve sanat musikisi için geliştirilmiş, makam-farkında, yeniden kullanılabilir bir Python analiz motorudur. MakamKit beş katmandan oluşur. Sembolik melodik analiz, ses tabanlı perde analizi (53 komalı histogramdan tonik ve makam kestirimine), usul/aksak analizi, güfte-duygu analizi ve benzerlik-kümeleme-ağ çıktıları ve bu makaledeki altı çalışmanın tamamı, ortak bir hesaplamalı sözlük ve tekrarlanabilir bir işlem hattı sağlaması için bu motor üzerinden yürütülmüştür. Altı çalışmanın hepsi motorun yalnız sembolik katmanını kullanmıştır: nota dosyalarından perde, süre, ölçü konumu ve bölüm sınırı bilgisini çıkaran, bu ham veriden makam-usul ilişkisine dair sayısal göstergeler türeten katmandır. Motorun ses tabanlı perde/makam katmanı, usul/aksak katmanı ve güfte-duygu katmanı bu makale kapsamında kullanılmamıştır. Bu, elde SymbTr'nin sunduğu ses kaydı olmamasından kaynaklanan bir veri sınırındır, yöntemsel bir tercih değildir. MakamKit'in bu makaledeki rolü yorumlayıcı değil ölçücüdür: motor, nota dosyalarından tanımlanmış vekil ölçütler türetir; ölçütlerin Feldman'ın kuramsal kavramlarıyla (seyir, terhib, nazire) örtüşme derecesi her çalışmada ayrıca ve açıkça tartışılmaktadır.

Araştırma Problemi

Bu çerçevede makalenin araştırma problemi, Feldman'ın geniş dönüşüm anlatısından türetilen altı ayrı, sınırlı ve yanlışlanabilir önermenin sınanmasıdır. Birinci ve ikinci önerme, modüler terhib örgüsünden seyir-temelli sürekliliğe geçişi iki farklı vekil kümesiyle ele alır: motif tekrarı, cümle yayılımı ve karar-yakınsama eksenini (Çalışma 1) ile beş yönlü, önceden yön belirtilmiş yakınsak seyir göstergesi eksenini (Çalışma 2). Üçüncü önerme, usul devri ile ezginin eklemlenme biçiminin -ölçü çizgisinin aşılma sıklığı, vuruş fazı entropisi gibi göstergelerle-dönemler boyunca değiştiği tezidir. Dördüncü önerme, model-nazire ilişkisinin, yani bir eserin başka bir esere öykünerek bestelenmesi pratiğinin, besteciler arası bir görgü ağı olarak hesaplamalı benzerlik ölçümüyle önceliklendirilebilir olduğu tezidir. Beşinci önerme, Feldman'ın dört-dönemlik biçimsel periyodizasyonunun, dönem etiketi modele hiç verilmeden, eserlerin ölçülebilir özelliklerinde bağımsız olarak (kör kümeleme ile) geri kazanılabilir olduğu tezidir. Altıncı önerme ise Mevlevî ayınının Birinci Selâm bölümünde biçimsel uzunluk ile bestecinin tarihsel konumu arasındaki ilişki tezidir; bu önerme, aşağıda ayrıntılandırılacağı gibi, bu makalede bir sonuç değil bir sınama protokolü olarak sunulmuştur. Bu altı önerme, Feldman'ın kendi anlatısında birbirinden ayrı değildir. Ancak her biri farklı bir ölçüm mantığı, farklı bir örneklem gereksinimi ve farklı bir sınama tasarımı gerektirdiğinden, makale bunları altı ayrı çalışma olarak ele almayı tercih etmiştir.

Makalenin geri kalanı şu şekilde düzenlenmiştir. Çalışma 1 ve Çalışma 2, terhibden seyre geçiş tezini iki farklı vekil kümesiyle -sırasıyla motif tekrarı/karar-yakınsama eksenini ve beş yönlü yakınsak seyir göstergesi ekseniniyle- sınar. Çalışma 3, usul-ezgi eklemlenmesinin dönemselleşimini beş göstergeyle inceler. Çalışma 4, model-nazire ilişkisini bir hipotez sınaması değil bir aday önceliklendirme denemesi olarak çerçeveler. Çalışma 5, dört dönemlik periyodizasyonu kör kümeleme ve besteci-gruplu sınıflandırma ile sınar. Çalışma 6 ise diğer beşinden yöntemsel olarak ayrılır: SymbTr örnekleme değil, kaynak-izlenebilir Mevlevî ayını notaları gerektiren bir doğrulayıcı protokol tasarlar ve bu protokolün veri hazırlığı aşamasında durduğunu, tarihsel bir sonuç üretmediğini açıkça beyan eder. Her çalışma bölümü, kuramın Feldman'ın metnindeki yerini, kuramın testi için kullanılan örneklem ve yöntemi ve bulguları aynı sabit şablonla, kendi bulgu tablolarıyla birlikte sunar; sınırlılıkların konsolide değerlendirmesi ve bulguların Feldman'ın kendi anlatısı ile alan yazınıyla tartışılması, ayrı ayrı çoğaltılmaz, altı çalışmayı birlikte ele alan tek bir Sonuç ve Tartışma bölümünde toplanır. Makale, altı çalışmanın toplu bir okumasıyla, Feldman'ın kuramsal mirasının bu sınama zincirinden nasıl çıktığına dair bir genel değerlendirmeye ve tüm çalışmaların ortak kaynaklarını tek bir yerde toplayan bir Kaynakça ile kapanır.

Çalışma 1. Feldman'ın modüler yapıdan seyir-temelli sürekliliğe geçiş tezi, eşleştirme ve karma modeller sonrasında destekleniyor mu?

Kuram 1: Terhibden seyre -modüler örgünün çözülüşü

Walter Feldman, *Music of the Ottoman Court* adlı çalışmasında Osmanlı saray çalgısal repertuvarının biçimsel tarihine dair geniş bir dönüşüm anlatısı kurar (Feldman, 2024). Bu anlatıya göre erken dönem eserleri özellikle peşrev ve saz semâîsi gibi çalgısal formlar- kısa, tekrarlı ve görece bağımsız ezgici bloklardan, yani terhiblerden kuruludur. Terhib,

eser içinde yer değiştirebilen, başka eserlere de taşınabilen hazır bir ezgici birimdir; erken repertuarın ekonomisi hem birebir yinelenmeye hem de aralık kalıbı korunarak başka perdeden yeniden başlatılmaya (sekans) dayanır -terkib tek bir sabit kalıp değil, form ve bağlama göre değişen bir morfolojik dağarcıktır. Bu konum rastlantısal değildir: peşrev, fasıl denen sıralı icra bütününün çalgısal açılış parçasıdır; terkibin taşınabilirliği bu çerçeveyici işlevle de örtüşür. Feldman’a göre 17. yüzyılın ikinci yarısından itibaren bu modüler örgü gevşemeye başlar: ezgici bloklar, makamın kendine özgü melodik yolu olan seyir içinde daha uzun erimli, tek bir yay gibi gerilen bir sürekliliğe bütünleşir. Seyir; bir makamın çıkıcı, inici ya da çıkıcı-inici karakterini, asma duraklarını ve nihayetinde karara -makamın temel perdesine- bağlanış yönünü belirleyen melodik davranış kalıbıdır.

Feldman’ın anlatısı bu dönüşümü tek bir eksene indirgemez; perde merkezlerinin sırasını, usul ile ezginin eklemelenmesini, hane-mülazime (teslim) düzenini bu çalışmada mülazime ve teslim terimleri hanenin ortak dönüş bölümünü adlandırmak üzere eşanlamli kullanılmıştır- ve nazire geleneğini birlikte kapsayan çok bileşenli bir tarihsel süreç öngörür. Bu çalışma, o geniş anlatının yalnızca “terkib-seyir eksenini”ni -modüler tekrar ekonomisinden seyir-temelli sürekliliğe geçiş önermesini” sınamayı hedefler. Nazire ağları, usul-ezgi eklemelenmesi ve bölüm sınırlarının dokusu, başka ölçüt ve veri haznesi gerektirdiği için kapsam dışı bırakılmıştır. Bozkurt, Ayangil ve Holzapfel (2014), makam müziğinin hesaplamalı çözümlemesinde kullanılan ölçütleri ve açık sorunları derli toplu bir biçimde ortaya koymuştur. Karaosmanoğlu (2012) ise bu tür ölçekli çözümlemelere zemin hazırlayan açık erişimli sembolik nota koleksiyonu SymbTr’yi tanıtmıştır. Bu çalışma, tarihsel etnomüzikolojinin büyük anlatısı ile hesaplamalı etnomüzikolojinin ölçüm araçlarını tek bir soruda buluşturur:

- Feldman’ın modüler yapıdan seyir-temelli sürekliliğe geçiş tezi, modern sembolik aktarımda gözlenebilen biçimsel izler bırakmış mıdır?

Kuramın Testi

Örneklem, SymbTr türevi açık sembolik aktarım koleksiyonundan (Karaosmanoğlu, 2012) seçilen 149 çalgısal eserden oluşur: 82 peşrev ve 67 saz semâisi. Bu eserler özgün tarihsel yazmalar değil, modern notaya aktarılmış temsillerdir. İncelenen veri, meşkten notaya, notadan edisyona, edisyondan dijitalle uzanan aktarım zincirinin son halkasıdır. Yapısal olarak bozuk satırlar, sıfır süreli işaretler ve nota taşımayan usul-meta satırları içe aktarma aşamasında ayıklanmıştır. En az 20 geçerli nota içermeyen kayıtlar, besteci eşlemesi bulunmayanlar ve dört dönem katmanından birine yerleştirilemeyenler analiz dışı bırakılmıştır. Dönem katmanları; 1650 öncesi, 1650–1750, 1750–1850 ve 1850 sonrası, besteci atıflarından türetilmiştir; kesin bestelenme tarihi değildir. İlk kesim (1650) Feldman’ın kendi kronolojisinden gelir ve yazar modüler örgünün çözülme dönüm noktasını 17. yüzyılın ikinci yarısına yerleştirir; sonraki iki kesim (1750, 1850) ise aktarımda yoğunlaşan olgun ve geç katmanları ayırmak içindir ve Feldman’ın metnindeki daha ayrıntılı alt-dönemlendirmeye birebir örtüştüğü gösterilmemiştir. Dönem şeması bu ölçüde bir yaklaşıklıktır. 149 eser 77 besteciye dağılır ve etkin, bağımsız gözlem birimi eser değil bu 77 kişilik besteci kümesidir. Örneklem olasılıklı bir tarihsel örneklem değildir: günümüze ulaşma, notaya alınma, edisyona girme, dijitalleşme ve besteciye atfedilme koşullarının hepsinden geçmiş, seçilmiş bir aktarım örneklemdir. Erken dönem hücreleri (1650 öncesi ve 1650–1750) 5-11 eser taşır; 1750–1850 hücreleri de 7-14 arasındadır. Bu küçüklük yokluk kanıtını özellikle kırılğan kılar.

Çözümleme, bu araştırma dizisinin ortak hesaplamalı altyapısı olan MakamKit ile yürütülmüştür. MakamKit, sembolik nota dosyalarından her eserin perde ve süre dizisini kurar, perdeleri eserin son perdesine göre görelî hâle getirir ve makam-ilişkisel özellik ölçütleri üretir. Bu çalışmada beş vekil ölçüt tanımlanmıştır: (1) *doğrudan tekrar*, 3-4 aralıklı motif pencerelerinde mutlak perde dizisinin birebir yinelenme oranı; (2) *taşınmış (transpoze) tekrar*, aralık kalıbı aynı, başlangıç perdesi farklı motif yinelemelerinin oranı; (3) *cümle yayılımı*; bölüm işaretleriyle ayrılan birimlerin ölçü cinsinden genişliği; (4) *yörünge pürüzlülüğü*; eser 12 eşit nota-sayı zaman dilimine bölünüp ardışık görelî perde-merkezi farklarının ortalama mutlak büyüklüğü; (5) *karar-yakınsama*; son yüzde yirmilik nota kesitinde nihai perdeye ortalama mutlak uzaklık d alınıp $1/(1+d)$ dönüşümüyle 0-1 aralığına çekilmiş hâli; 1’e yaklaşmak, eserin son kesitinin karar perdesi çevresinde daha sıkı toplandığını, yani daha güçlü bir yakınsamayı gösterir — ölçüt bir mesafe değil, bu mesafenin tersine çevrilmiş bir yakınsama skorudur. Bulgulardan önce yazılmış beş hipotez, dönemlerle birlikte

doğrudan ve taşınmış tekrarın azalmasını, cümle yayılımının ve karar-yakınsamanın artmasını, yörünge pürüzlülüğünün azalmasını öngörür. Önceden tanımlanmış birleşik destek ölçütü şudur: beş göstergenin en az üçü beklenen yönde olmalı, bunlardan en az ikisinin besteci-kümeli bootstrap %95 güven aralığı sıfırı dışlamalıdır.

Birincil eğilim ölçüsü Spearman sıra ilişkisidir (Spearman, 1904); dayanıklılık, besteci-kümeli bootstrap ile denetlenmiştir (2000 yineleme, sabit tohum; Efron & Tibshirani, 1994). Form bileşiminin karıştırıcı etkisini ayıklamak için iki ek denetim kurulmuştur: makam×usul×form tabakalarında en yakın uzunluklu erken-geç eser çiftlerinin Wilcoxon işaretli sıra testiyle (Wilcoxon, 1945) karşılaştırılması, yalnızca 12 çift kurulabilmektedir. Ve besteci rastgele kesişimli, makam/usul varyans bileşenli bir karma etkili model (Laird & Ware, 1982). Çoklu karşılaştırmalar Benjamini-Hochberg yöntemiyle (%5 yanlış keşif oranı) düzeltilmiştir (Benjamini & Hochberg, 1995); form-ayrı eğilim tablosunda peşrev ve saz semâisi ayrı beşer testlik düzeltme ailesi olarak ele alınmıştır.

Bulgular

Tablo 1. Dönem ve forma göre örneklem büyüklükleri ile beş ölçütün medyan değerleri.

Dönem	Form	n	Doğrudan tekrar (medyan)	Taşınmış tekrar (medyan)	Cümle yayılımı, ölçü (medyan)	Yörünge pürüzlülüğü (medyan)	Karar-yakınsama (medyan)
1650 öncesi	Peşrev	9	0,497	0,069	2,243	10,842	0,043
1650 öncesi	Saz semâisi	5	0,730	0,038	2,188	12,856	0,049
1650-1750	Peşrev	11	0,553	0,036	1,632	10,135	0,062
1650-1750	Saz semâisi	7	0,697	0,034	2,241	9,410	0,051
1750-1850	Peşrev	14	0,527	0,080	0,614	10,817	0,046
1750-1850	Saz semâisi	7	0,506	0,052	2,300	11,709	0,042
1850 sonrası	Peşrev	48	0,483	0,085	0,883	12,953	0,039
1850 sonrası	Saz semâisi	48	0,626	0,043	3,215	9,583	0,036

Örneklemin ağırlık merkezi geç dönemdedir — 149 eserin 96'sı 1850 sonrası katmanına düşer — ve erken hücrelerin küçüklüğü (1650 öncesi/1650-1750'de 5-11, 1750-1850'de 7-14 eser) bu tablodaki hiçbir farkın tek başına eğilim kanıtı sayılamayacağını gösterir.

Tablo 2. Havuzlanmış dönem eğilimi ve besteci-kümeli bootstrap sonuçları (n=149).

Ölçüt	Spearman ρ	p	Kümeli eğim	%95 GA	Bootstrap p	BH q
Doğrudan tekrar	-0,089	0,278	-0,009	[-0,030; 0,017]	0,458	0,630
Taşınmış tekrar	0,058	0,483	-0,002	[-0,009; 0,008]	0,660	0,630
Cümle yayılımı	0,055	0,504	0,888	[-0,926; 2,669]	0,302	0,630
Yörünge pürüzlülüğü	0,037	0,656	0,336	[-0,446; 1,229]	0,386	0,656
Karar-yakınsama	-0,325	<0,001	-0,005	[-0,010; -0,002]	<0,0005	<0,001

Havuzlanmış örnekleme dört ölçütün güven aralığı sıfırı kapsar; tek kararlı eğilim karar-yakınsamada görülür, fakat yön H5'in öngördüğü artışın tersinedir (dönem ilerledikçe eserin son kesiti karara daha gevşek yaklaşmaktadır).

Tablo 3. Form-ayrı dönem eğilimi (peşrev n=82, saz semâisi n=67; her form kendi içinde beşer testlik düzeltme ailesi).

Ölçüt	Form	Spearman ρ	p	%95 GA	BH q
Doğrudan tekrar	Peşrev	-0,150	0,179	[-0,043; 0,010]	0,224
Taşınmış tekrar	Peşrev	0,195	0,079	[-0,010; 0,016]	0,196
Cümle yayılımı	Peşrev	-0,108	0,335	[-3,514; 0,676]	0,335
Yörünge pürüzlülüğü	Peşrev	0,174	0,117	[-0,198; 1,849]	0,196
Karar-yakınsama	Peşrev	-0,269	0,015	[-0,010; -0,001]	0,073
Doğrudan tekrar	Saz semâisi	-0,145	0,243	[-0,039; 0,017]	0,404
Taşınmış tekrar	Saz semâisi	-0,039	0,757	[-0,009; 0,007]	0,757
Cümle yayılımı	Saz semâisi	0,292	0,017	[1,210; 6,427]	0,041
Yörünge pürüzlülüğü	Saz semâisi	-0,061	0,622	[-1,529; 1,824]	0,757
Karar-yakınsama	Saz semâisi	-0,419	<0,001	[-0,011; -0,002]	0,002

Peşrev tarafında hiçbir eğilim %5 yanlış keşif oranını aşarak anlamlı kalmaz. Saz semâisinde ise iki hareket görülür: karar-yakınsamadaki azalış düzeltme sonrası sağlam kalır (q=0,002), cümle yayılımındaki artış da kendi ailesi içinde %5 eşliğini geçer (q=0,041) — tezle uyumlu tek kısmi ve forma özgü sinyal budur.

Tablo 4. Eşleşmiş 12 çiftte geç-erken farkları (makam×usul×form tabakalarında uzunluğu en yakın erken-geç eser çiftleri).

Ölçüt	Erken ort.	Geç ort.	Ort. fark	Wilcoxon W	p	BH q
Doğrudan tekrar	0,654	0,635	-0,019	39,0	1,00	1,00

Taşınmış tekrar	0,048	0,045	-0,003	38,0	0,970	1,00
Cümle yayılımı	2,156	2,975	0,819	22,0	0,204	0,339
Yörünge pürüzlülüğü	9,844	8,483	-1,361	22,0	0,204	0,339
Karar-yakınsama	0,055	0,041	-0,014	7,0	0,009	0,046

En sıkı denetimde — makam, usul ve formu sabit tutan eşleştirmede — düzeltme sonrası tek anlamlı fark yine karar-yakınsamadır ($q=0,046$); tekrar ölçütlerinde ortalama ve medyan işaretleri ters yönlüdür, bu da 12 çiftlik küçük örneklemde uç değer duyarlılığına işaret eder.

Tablo 5. Besteci rastgele kesişimli karma etkili modelin yakınsama durumu ($n=149, 77$ besteci).

Ölçüt	Yakınsama	Dönem katsayısı β
Doğrudan tekrar	Hayır	—
Taşınmış tekrar	Hayır	—
Cümle yayılımı	Hayır	—
Yörünge pürüzlülüğü	Hayır	—
Karar-yakınsama	Hayır	—

Karma model, beş ölçütün beşinde de “No module named ‘statsmodels.formula’” hatasıyla yani istatistik yazılımı katmanında eksik kalan bir bağımlılık nedeniyle hiç çalışmamış ve hiçbir dönem katsayısı kestirilememiştir. Bu yüzden karma model, besteci ile makam/usul çevresini ayrı birer etki olarak sınavan bağımsız bir kanıt kaynağı olarak kullanılamamaktadır. Terkib–seyir ekseninde geriye kalan üç bağımsız sına, havuzlanmış eğilim, form-ayrı eğilim, eşleşmiş çiftler birleştirildiğinde tutarlı biçimde ortaya çıkan tek örüntü, karar-yakınsamanın dönemlerle beklenenin tersine azalmasıdır; buna saz semâisine özgü, sınırlı bir cümle-yayılımı sinyali eşlik eder.

Özetle, beş vekil ölçütten yalnızca karar-yakınsama üç bağımsız sınamada (havuzlanmış eğilim, form-ayrı eğilim, eşleşmiş çiftler) tutarlı ve düzeltme sonrası anlamlı kalmış, fakat öngörülen yönün tersine hareket etmiştir; saz semâisinde cümle yayılımı ise forma özgü, kısmi bir ek sinyal göstermiştir.

Çalışma 2. Feldman’ın seyir örgütlenmesinin tarihsel bütünleşmesi tezi, beş yönlü sayısal gösterge ile sınıandığında destekleniyor mu?

Kuram 2: Seyrin bütünsel tarihsel kodifikasyonu

Onyedinci yüzyılın ikinci yarısından itibaren erken dönemin bağımsız terkiplerinin, eserin başından sonuna uzanan bir makam-seyri örgütlenmesinin parçalarına dönüştüğünü öne sürer. Bu dönüşüm, terki, seyir, usul-ezgi eklemlenmesi, hane-mülazime düzeni ve nazire ilişkilerinin (bir eserin başka bir esere öykünerek bestelenmesi) birlikte yer değiştirdiği çok bileşenli bir süreç olarak kurulur; tek boyutlu bir “ilerleme” değildir. Bu çalışma, bu geniş tezin yalnızca bir bileşenini, seyir yörüngesinin örgütlenmesinin besteci dönemleri boyunca modern aktarımda izlenebilir olup olmadığını sınamayı hedefler.

Bu tarihsel önermeden beş yönlü, önceden yön belirtilmiş bir hipotez seti türetilir. Seyir bütünleştikçe (H1) yörüngenin pürüzlülüğünün ezginin pencereden pencereye tonal sıçramalarının azalması, (H2) yön değiştirme sayısının seyrekleşmesi, (H3) eserin son bölümünün karar perdesi çevresinde daha sıkı örgütlenmesi (karar-yakınsama), (H4) hane/mülazime sınırlarının iki yanındaki hareket yönünün daha sık örtüşmesi ve (H5) sınırın iki yanındaki perde dağılımlarının birbirine daha çok benzemesi beklenir. Bu beş hipotezin tek tek değil, birlikte aynı yönü göstermesi “destek” ölçütü olarak tanımlanmıştır: çoklu yakınsak geçerlilik ilkesi, herhangi bir göstergenin rastlantıyla anlamlı çıkmasının geniş bir “seyir tarihsel olarak bütünleşmiştir” hükmüne dönüşmesini engellemek için bilinçli bir tercihtir. Bu çerçeve, seyri “çıkıcı/inici” gibi statik bir sınıflandırma etiketinden çıkarıp eser boyunca izlenen bir zaman dizisine dönüştürerek, Feldman’ın tarihsel iddiasını denetlenebilir bir sınamaya açar.

Kuramın Testi

Örneklem, SymbTr türevi açık sembolik nota külliyatından (Karaosmanoğlu, 2012) alınan 149 eserden peşrev ve saz semâisinden, Osmanlı saray çalgısal repertuvarının iki temel formundan oluşur; bu eserler 77 besteciye atfedilmiştir. Analiz birimi tekil nota değil eserdir; en az yirmi geçerli nota içermeyen, besteci eşlemesi bulunmayan veya dört dönem katmanından birine yerleştirilemeyen kayıtlar dışlanmıştır. Dönem, eserin kesin bestelenme tarihi değil, bestecilerin

yaşam aralıklarından türetilen dört atıf katmanıdır: 1650 öncesi (14 eser), 1650-1750 (18 eser), 1750-1850 (21 eser) ve 1850 sonrası (96 eser). Bu üç kesim noktası (1650, 1750, 1850) Feldman'ın metnindeki daha ayrıntılı alt-dönemlendirmesinin (bkz. Çalışma 5) analitik bir sadeleştirmesidir; kaynak makalede bu sınırların Feldman'ın hangi sayfa/pasaj düzeyindeki dönem geçişine karşılık geldiği ayrıca gösterilmemiştir, dolayısıyla dönem katmanları burada yaklaşık bir çerçeve olarak okunmalıdır. Hane/mülazime/serhane sınırına dayanan iki gösterge (H4, H5) için alt örneklem 132 eserdir; 17 eser yapısal bölüm işareti taşımadığından bu göstergelerin dışında kalmıştır. Örneklemin seçilmiş-aktarım niteliği yukarıda (Giriş) tartışılmıştır; burada yalnızca hatırlatılmalıdır ki örneklemin yaklaşık üçte ikisi 1850 sonrasına atıflıdır.

Analiz, MakamKit adlı hesaplamalı motorla yapılmıştır. Makam-farkında, sembolik ve ses tabanlı katmanları bulunan, açık kaynaklı bir Python kütüphanesidir. Bu çalışmada yalnızca sembolik katmanı kullanılmıştır: eser, eşit nota sayılı on iki pencereye bölünmüş, her pencerede notaların süreleriyle ağırlıklandırılmış ortalama perde, 53 komalı sistemde ve eserin son yazılı perdesine göreli olarak hesaplanmıştır. Bu ortak temsil üzerinden beş gösterge türetilmiştir: yörünge pürüzlülüğü (ardışık pencere merkezleri arasındaki mutlak farkların ortalaması, birim koma), yön değiştirme sayısı (ardışık merkez farklarının işaret değişimi sayısı), karar-yakınsama skoru son beşte birlik kesitteki notaların son yazılı perdeye ortalama mutlak uzaklığı d olmak üzere, $1/(1+d)$ formülüyle 0-1 aralığına dönüştürülen bir *Yakınsama* skorudur (mesafenin kendisi değil, mesafenin tersinden türetilir); skor 1'e yaklaştıkça ezgi finale sıkı kenetlenir, 0'a yaklaştıkça finalden uzak kalır, bölüm sınırında yön sürekliliği (hane/mülazime/serhane sınırının iki yanındaki sekizer notalık bağlamın hareket yönü açısından örtüşme oranı) ve bölüm sınırında perde dağılımı benzerliği (sınırın iki yanındaki perde kullanım profillerinin benzerliği).

İstatistiksel strateji üç katmanlıdır. Birincil sinama, sıralı dört dönem ile her gösterge arasındaki Spearman sıra korelasyonudur (Spearman, 1904). Bu katsayı, dönem ilerledikçe göstergenin düzenli bir yön izleyip izlemediğini -1 ile +1 arasında özetler; doğrusallık varsayımı taşımaz. Etki yönünün sağlamlığı, Çalışma 1'de tanımlanan besteci-kümeli bootstrap prosedürüyle denetlenmiştir (Efron ve Tibshirani, 1994; gösterge başına 2.000 yinleme, sabit tohum 20260720, %95 güven aralığı %2,5-%97,5 yüzdeliklerinden) bu, üretken bir bestecinin çok sayıdaki eserinin yapay bir kesinlik yaratmasını önlemek içindir. Beş test aynı anda yapıldığından çoklu karşılaştırma denetimi uygulanmıştır: Bonferroni düzeltmeli eşik $\alpha = 0,01$ (0,05/5) olarak belirlenmiştir; ayrıca Benjamini-Hochberg (1995) yanlış-keşif-oranı düzeltmesi hesaplanmış ve aşağıda Tablo 2.1'e BH q sütunu olarak eklenmiştir (karar-yakınsama skoru Çalışma 1 ve Çalışma 3'te de ölçülmüştür; her çalışma burada kendi beş-testlik düzeltme ailesini ayrı tanımlamaktadır, bu nedenle aynı gösterge için çalışmalar arasında farklı q değerleri görülebilir). En küçük dönem hücrelerinde %5 anlamlılık ve %80 güç düzeyinde saptanabilecek asgari etki büyüklüğü, hücre büyüklüğüne göre değişir: standart Fisher z yaklaşımıyla, 1650 öncesi hücresinde ($n=14$) $|p| \approx 0,69$, 1650-1750 hücresinde ($n=18$) $|p| \approx 0,62$ olarak hesaplanmıştır. Bu, erken dönemde orta büyüklükte bir eğilim gerçekten var olsa bile mevcut tasarımın onu göremeyeceği, yalnızca çok güçlü eğilimleri yakalayabileceği anlamına gelir. Ana göstergelerin tek bir ölçüm parametresine (on iki pencere, son beşte birlik kesit) bağlı olup olmadığını denetlemek için ayrıca bir duyarlılık analizi yapılmıştır: üç pencere sayısı (8, 12, 16) ve üç son-kesit oranı (%10, %20, %30) altında dokuz kombinasyonda yeniden hesaplama.

Bulgular

Tablo 6. Beş seyir göstergesinin dönem eğilimi ($n = 149$; bölüm göstergeleri $n = 132$)

Gösterge	n	ρ	p	Eğim (dönem başına)	%95 GA alt	%95 GA üst	p_boot	BH q	Beklenen yön	Gözlenen yön uyumlu mu
Yörünge pürüzlülüğü	149	0,037	0,656	0,336	-0,418	1,319	0,371	0,820	azalış (H1)	Hayır
Yön değiştirme sayısı	149	0,085	0,304	0,150	-0,027	0,409	0,100	0,507	azalış (H2)	Hayır
Karar-yakınsama skoru	149	- 0,325	<0,0 01	-0,005	-0,010	-0,002	<0,0005	0,000 259	artış (H3)	Hayır*
Bölüm sınırında yön sürekliliği	132	- 0,017	0,849	-0,019	-0,060	0,055	0,569	0,849	artış (H4)	Hayır
Bölüm sınırında perde dağılımı benzerliği	132	- 0,127	0,146	-0,027	-0,078	0,010	0,144	0,365	artış (H5)	Hayır

*Yalnızca bu satırda tersinme istatistiksel olarak anlamlıdır ($p < 0,001$; BH q = 0,000259); diğer dört satırda nokta tahmini beklenenin tersi yönde olmakla birlikte anlamlılık düzeyine ulaşmaz.

Bu tablo, beş göstergenin dördünde dönemler arasında kararlı bir eğilim bulunmadığını gösterir: güven aralıkları sıfırın iki yanına taşar. Son iki sütun ise ayrı bir örüntüye işaret eder: beş hipotezin (H1-H5) beşi de, noktasal ρ tahmini düzeyinde, Feldman'ın öngördüğü yönün tam tersi işarete sahiptir. Bu beşli tersinme yalnızca karar-yakınsama skorunda istatistiksel olarak anlamlıdır ($\rho = -0,325$, $p < 0,001$) ve bu gösterge hem Bonferroni düzeltmeli eşğin ($\alpha = 0,01$) hem de Benjamini-Hochberg yanlış-keşif-oranı düzeltmesinin (BH $q = 0,000259$) altında kalarak anlamlılığını korur; diğer dört göstergenin BH q değerleri (0,365-0,849 aralığında) düzeltme sonrası da anlamsız kalır. Bu sonuç, tek bir göstergenin rastlantıyla ters çıkmasından çok, örüntünün tamamının aynı yönde durduğunu göstermektedir; ayrıntısı Sonuç ve Tartışma'da ele alınacaktır.

Tablo 7. Dönem $\times$ form hücrelerinde gösterge medyanları

Dönem	Form	n	Pürüzlülük	Yön değişimi	Karar-yakınsama	Yön sürekliliği	Dağılım benzerliği
1650 öncesi	peşrev	9	10,842	6,0	0,043	0,286	0,443
1650 öncesi	saz semâisi	5	12,856	6,0	0,049	0,429	0,370
1650-1750	peşrev	11	10,135	5,0	0,062	0,310	0,374
1650-1750	saz semâisi	7	9,410	6,0	0,051	0,429	0,434
1750-1850	peşrev	14	10,817	6,0	0,046	0,050	0,495
1750-1850	saz semâisi	7	11,709	6,0	0,042	0,292	0,373
1850 sonrası	peşrev	48	12,953	6,0	0,039	0,268	0,445
1850 sonrası	saz semâisi	48	9,583	6,0	0,036	0,357	0,326

Bu tablo, havuzlanmış sonucun tek bir formdan kaynaklanmadığını denetler: karar-yakınsama medyanı hem peşrevde hem saz semâisinde en yüksek değerine 1650-1750 hücresinde ulaşır ve 1850 sonrasında düşer, düşüş iki formda da ortaktır ama bu, dönemler arası doğrusal bir azalma değil, ortada bir tepe yapıp sonra gerileyen bir örüntüdür.

Tablo 8. En az dört eserle temsil edilen makamlarda gösterge medyanları

Makam	n	Pürüzlülük	Yön değişimi	Karar-yakınsama	Yön sürekliliği	Dağılım benzerliği
acemâşiran	6	14,970	5,5	0,031	0,667	0,380
beyati	5	6,841	7,0	0,048	0,375	0,577
bûselik	4	9,220	5,0	0,047	0,482	0,382
ferahfeza	5	20,643	6,0	0,021	0,227	0,415
hicaz hümâyün	4	9,641	6,0	0,050	0,364	0,450
hicazkâr	6	10,476	5,5	0,041	0,488	0,514
hüseyini	7	8,402	6,0	0,047	0,264	0,450
hüzzam	4	10,315	7,0	0,049	0,411	0,672
kürdilihicazkâr	5	9,934	6,0	0,033	0,000	0,494
mahur	7	13,675	6,0	0,031	0,400	0,370
muhayyer	4	10,563	5,0	0,043	0,000	0,042
nevâ	4	8,446	5,5	0,043	0,238	0,364
nihavent	6	13,006	6,5	0,037	0,375	0,326
rast	4	14,606	6,0	0,037	0,000	0,521
sabâ	6	9,889	6,5	0,050	0,429	0,265
segâh	5	9,440	5,0	0,064	0,294	0,545
uşşak	5	7,242	6,0	0,065	0,191	0,284

Dönemler arası makam bileşiminin sabit olmadığı bir külliyatta, Tablo 8'deki havuzlanmış eğilimin dönemin mi yoksa makam karışımının mı eseri olduğu sorusuna içeriden bakar: pürüzlülük medyanı burada 6,841 (beyati) ile 20,643 (ferahfeza) arasında değişir bu aralık, Tablo 9'daki dönem hücreleri arasındaki farktan çok daha geniştir; bu da makam kimliğinin, gösterge değerlerini dönem kimliğinden daha güçlü ayırt ettiğini düşündürür.

Tablo 9. Pencere sayısı ve son-kesit oranına göre duyarlılık analizi (ρ ve p ; $n = 149$)

Pencere	Son kesit	Pürüzlülük ρ (p)	Yön değişimi ρ (p)	Karar-yakınsama ρ (p)
8	%10	0,083 (0,314)	0,060 (0,466)	-0,438 (<0,001)
8	%20	0,083 (0,314)	0,060 (0,466)	-0,325 (<0,001)
8	%30	0,083 (0,314)	0,060 (0,466)	-0,318 (<0,001)
12	%10	0,037 (0,656)	0,085 (0,304)	-0,438 (<0,001)
12	%20	0,037 (0,656)	0,085 (0,304)	-0,325 (<0,001)
12	%30	0,037 (0,656)	0,085 (0,304)	-0,318 (<0,001)
16	%10	0,063 (0,444)	0,025 (0,765)	-0,438 (<0,001)
16	%20	0,063 (0,444)	0,025 (0,765)	-0,325 (<0,001)
16	%30	0,063 (0,444)	0,025 (0,765)	-0,318 (<0,001)

Bu duyarlılık tablosunun dokuz satırının tamamı için ayrı bir ham veri kaydı sonuç dosyasında (results JSON) bulunmamaktadır; tablo yalnızca kaynak makale metnine dayanılarak aktarılmıştır ve bağımsız olarak yeniden hesaplanamamıştır. Doğrulanabilen tek satır “12 pencere / %20 son kesit” kombinasyonudur bu, tanım gereği Tablo 2.1’in ana sonuçlarıyla birebir örtüşür ve iç tutarlılığı teyit eder. Geri kalan sekiz satırda bildirilen değerler bu doğrulanmış çapa noktasıyla tutarlı bir örüntü sergilese de, ham hesaplamaları burada bağımsız olarak sıranamamıştır; dolayısıyla “karar-yakınsamanın ters yönlü eğilimi dokuz kombinasyonun tamamında ayaktadır” biçimindeki dayanıklılık okuması, doğrulanmış tek nokta ötesinde ihtiyatla değerlendirilmelidir. Pürüzlülük ile yön değiştirme göstergelerinin hiçbir kombinasyonda anlamlı bir eğilim üretmediği ise, dokuz satırın tamamında p-değerinin 0,05’in üzerinde kalmasıyla tutarlıdır ve bu noktada ayrıca bir dayanıklılık iddiası öne sürülmemektedir.

Özetle, beş göstergenin dördü (pürüzlülük, yön değiştirme, bölüm sınırında yön sürekliliği, bölüm sınırında perde dağılımı benzerliği) dönemler arasında kararlı bir eğilim göstermezken, istatistiksel olarak anlamlı kalan tek gösterge olan karar-yakınsama skoru beklenen yönün tersine hareket etmektedir; beş hipotezin tamamı noktasal tahmin düzeyinde Feldman’ın öngördüğü yönün tersi işarete sahiptir.

Çalışma 3. Feldman’ın usul–melodi ilişkisinin tarihsel dönüşümü tezi, modern nota yüzeyinde beş göstergeyle sınıandığında destekleniyor mu?

Kuram 3: Usul devri ile ezginin eklemelenme biçiminin dönemsel değişimi

Feldman’a (2024) göre erken dönem peşrev ve saz semâîlerinde ezgi usul devrine sıkı sıkıya bağlıyken, on yedinci yüzyılın ikinci yarısından itibaren bu bağ gevşer ve ezgi, usul devrini aşan, makamın seyri içinde bütünleşen daha uzun soluklu bir düşünceye yerini bırakır. Bu çalışma tezin yalnızca bu bileşenini sınar; tonal merkez sırası, hane–mülazime sürekliliği, nazire ilişkileri ve repertuarın dönemlere ayrılması gibi diğer bileşenler kapsam dışındadır.

Feldman’ın anlatısında usul yalnızca bir sayma düzeni değildir; düm ve tek vuruşlarının ağırlık hiyerarşisiyle yaşayan, meşk yoluyla bedenden bedene aktarılan bir zaman örgüsüdür. Erken dönem ezgisi bu örgüye sıkı sıkıya yaslanırken, geç dönem ezgisi örgünün üzerinde daha serbestçe süzülür. Bu araştırma sorusu şu biçimde kuruldu: melodik fikirlerin usul döngüsünü ve ölçü sınırlarını aşma biçimi, eserlerin besteci atıflarından türetilen tarihsel dönem katmanlarıyla birlikte değişiyor mu? Bu soruyu sayısal olarak sınayabilmenin güçlüğü baştan kabul edilmelidir: tarihsel usul, kaydedilmemiş bir vuruş ağırlığıyla yaşarken, elimizdeki veri yalnızca modern notanın ölçü çizgisi yüzeyini verir. Bozkurt, Ayangil ve Holzapfel (2014), hesaplamalı makam araştırmalarının genel bir sınırına işaret eder: sembolik notadan çıkarılabilen bilgi, icra pratiğinin taşıdığı bilginin ancak bir kesitidir. Bu çalışmada bu sınır özellikle usul boyutunda keskinleşir — düm ve tek vuruşlarının göreceli ağırlığı, yani usulü bir sayma düzeninden bir zaman örgüsüne dönüştüren şey, hiçbir sembolik nota biçiminde kodlanmaz; yöntem bu boşluğu baştan kabul ederek kurulmuştur.

Kuramın Testi

Örnekleme, modern notaya aktarılmış açık sembolik nota külliyatından, SymbTr koleksiyonunun (Karaosmanoğlu, 2012) türevinden oluşur. İncelenen iki çalgısal form peşrev ve saz semâîsi repertuarın çekirdeğini oluşturur ve hane–mülazime mimarisi sayesinde usul–melodi ilişkisinin izlenebileceği doğal bir alan sunar. En az 20 geçerli nota içermeyen, besteci eşlemesi bulunmayan veya dört atıf katmanından birine yerleştirilemeyen kayıtlar analiz dışı bırakıldı; içe aktarma katmanı ayrıca yapısal olarak bozuk satırları, sıfır süreli işaretleri ve nota taşımayan usul üstbilgi satırlarını müzikal olaylardan ayıklamıştır. Bu süzgeçten sonra kalan örneklem 149 eserdir: besteci atıflarına göre 1650 öncesi 14, 1650–1750 arası 18, 1750–1850 arası 21, 1850 sonrası 96 eser; toplamda 77 farklı besteci temsil edilmektedir. Külliyattaki 67 saz semâîsinin tamamı Aksak Semai usulündedir bu form–usul yapışıklığı yöntem bölümünde ayrıca ele alınacaktır. Örneklemin seçilmiş-aktarım niteliği yukarıda tartışılmıştır.

Analiz, MakamKit adlı hesaplamalı motorla yürütüldü. MakamKit, sembolik nota dosyalarını (SymbTr türevi) okuyarak perde, süre, ölçü konumu ve bölüm sınırı bilgisini çıkaran, bu ham verilerden makam ve usul ilişkisine dair sayısal göstergeler üreten bir analiz aracıdır. Bu çalışma için beş gösterge (G1–G5) tanımlandı. G1, ölçü çizgisini aşan bağlantı oranıdır. Ardışık ses olayları arasındaki bağlantıların ölçü sınırını kapsama oranı. G2, nihai perde–ölçü sonu örtüşmesidir. Eserin ana perdesi üzerindeki ses olaylarının ölçü sonuna denk gelme oranı. G3, vuruş başlangıcı faz entropisidir. Nota başlangıçlarının ölçü içi fazının, en yüksek olası değere normalize edilmiş Shannon entropisi. G4,

cümle yayılımıdır — ardışık cümle/bölüm sınırları arası uzaklığın ölçü cinsinden ortalaması. G5, bölüm sınırı kapanış oranıdır. Hane, mülazime ve serhane sınırlarında kapanış davranışının oranı. Bu son gösterge yalnızca bölüm sınırı işareti taşıyan eserlerde hesaplanabildiğinden örneklem büyüklüğü 149 değil 132'dir.

Tezden beş yönlü hipotez türetildi, hepsi 1650 öncesinden 1850 sonrasına doğru beklenen yönde: H1 (bağlantı oranı artar), H2 (nihai perde-ölçü sonu örtüşmesi azalır), H3 (faz entropisi artar), H4 (cümle yayılımı genişler), H5 (bölüm sınırı kapanış oranı azalır). Kuramsal destek için önceden tanımlanan ölçüt şuydu: beş göstergenin en az üçü, beklenen yönde ve besteci etkisi hesaba katıldığında kararlı bir dönem eğilimi göstermelidir. Birincil eğilim testi, sıralı dört dönem katmanını ile her gösterge arasındaki Spearman sıra korelasyonuydu (Spearman, 1904). Eğilimin dayanıklılığı, yukarıda tanımlanan besteci-kümeli bootstrap prosedürüyle (2.000 yinleme, sabit tohum 20260720, %95 güven aralığı yüzdelik yöntemiyle [%2,5–%97,5 dilimler]) sınıandı (Efron ve Tibshirani, 1994). Beş ayrı sıralı test yapıldığından p-değerleri Benjamini–Hochberg yöntemiyle (Benjamini ve Hochberg, 1995) düzeltildi. Karar okumasında hiçbir ölçüt tek başına belirleyici sayılmadı: yön, bootstrap eğimi, güven aralığının sıfırı kapsayıp kapsamadığı, bootstrap p-değeri ve düzeltilmiş q-değeri birlikte değerlendirildi.

Bulgular

Tablo 10. Beş usul–ezgi göstergesi için dönem eğilimi sınamaları (havuzlanmış, n=149; bölüm kapanışı için n=132).

Gösterge	n	Spearman ρ	p	Bootstrap eğimi	%95 güven aralığı	Bootstrap p	Düzeltilmiş q (BH)
Ölçü çizgisini aşan bağlantı oranı	149	0,000	0,999	–0,019	[–0,053, 0,025]	0,352	0,999
Nihai perde–ölçü sonu örtüşmesi	149	0,001	0,991	–0,000	[–0,021, 0,014]	0,997	0,999
Vuruş başlangıcı faz entropisi	149	0,130	0,114	0,004	[–0,001, 0,011]	0,123	0,570
Cümle yayılımı (ölçü)	149	0,055	0,504	0,888	[–0,832, 2,691]	0,287	0,840
Bölüm sınırı kapanış oranı	132	–0,089	0,313	–0,023	[–0,104, 0,034]	0,510	0,782

Tablo, beş göstergenin hiçbirinde dönemler boyunca kararlı bir yön bulunmadığını göstermektedir; ilk iki göstergede ilişki pratik olarak sıfır düzeyindedir, en güçlü aday olan faz entropisinde ($\rho=0,130$) beklenen yönde hafif bir kıpırtı görülsede güven aralığı sıfırı kapsamakta ve çoklu sınama düzeltmesinden sonra hiçbir bulgu ayakta kalmamaktadır (tüm q-değerleri $\geq 0,570$).

Tablo 11. Dönem $\times$ form hücrelerinde beş göstergenin medyanları.

Dönem	Form	n	Bağlantı oranı	Nihai perde–ölçü sonu	Faz entropisi	Cümle yayılımı	Bölüm kapanışı
1650 öncesi	Peşrev	9	0,829	0,152	0,989	2,243	0,400
1650 öncesi	Saz semâisi	5	0,721	0,060	0,962	2,188	0,714
1650–1750	Peşrev	11	1,000	0,163	0,993	1,632	0,633
1650–1750	Saz semâisi	7	0,644	0,048	0,961	2,241	0,714
1750–1850	Peşrev	14	0,571	0,112	0,992	0,614	0,429
1750–1850	Saz semâisi	7	0,645	0,125	0,961	2,300	0,500
1850 sonrası	Peşrev	48	0,645	0,118	0,995	0,883	0,429
1850 sonrası	Saz semâisi	48	0,792	0,096	0,979	3,215	0,545

Bu betimsel çözümleme, havuzlanmış testin gizlediği dalgalanmayı ortaya koyar: 1650–1750 arası peşrevlerde bağlantı oranı medyanı tam 1,0'a ulaşırken 1750–1850 arasında belirgin biçimde geriler ve 1850 sonrasında yeniden yükselir; aynı dönemde cümle yayılımı da en düşük değerine iner. Bu “çukur” tek dönemliktir ve iki ucu eski düzeye dönen bir V biçimi çizdiğinden, kalıcı bir dönüşüm sinyali olarak okunamaz dönemlerin usul bileşimindeki farklılıkla açıklanabilecek bir dalgalanmadır.

Tablo 11. Usul başına eser sayısı ve beş göstergenin medyanları²

Usul	n	Bağlantı oranı	Nihai perde–ölçü sonu	Faz entropisi	Cümle yayılımı	Bölüm kapanışı
Aksak Semai	67	0,745	0,091	0,976	2,695	0,571
Devr-i Kebir	25	0,636	0,118	0,992	0,598	0,317
Muhammes	13	0,750	0,091	0,994	0,770	0,300
Fahte	8	0,540	0,136	0,995	0,818	0,429
Hafif	8	0,571	0,127	0,989	0,729	0,321
Cenber	4	0,911	0,193	0,996	1,938	0,514
Çifte Düyek	4	0,663	0,089	0,992	0,777	0,750
Düyek	4	0,818	0,221	0,993	4,950	0,429
Ağır Düyek	3	0,582	0,095	0,986	2,234	0,500
Darb-ı Fetih	3	0,667	0,136	0,992	0,277	0,250

² $n \geq 3$ olan usuller; tek eserli usuller makalede ayrıca listelenmiştir

Tablo 11 göstergelerin gerçekte neyi ölçtüğüne dair bir ipucu taşır: bağlantı oranı Cenber’de (0,911) Aksak Semai’ye (0,745) göre, Aksak Semai de Devr-i Kebir (0,636) ve Fahte’ye (0,540) göre daha yüksektir. Bu sıralama, usulün döngü uzunluğuyla örtüşür — uzun döngülü usullerde ezginin çizgiyi aşmadan genişleyecek alanı daha fazladır. Dolayısıyla göstergenin dönemden çok usul döngü uzunluğunu yansıtıyor olması, aşağıda tartışılacak alternatif bir açıklamadır.

Ana bulgu şudur: beş göstergenin hiçbirisi dönemler boyunca kararlı bir yön göstermemiştir; en güçlü eğilim adayı olan G3 dahi ($p=0,130$, $p=0,114$, düzeltilmiş $q=0,570$) belirsizlik sınırları içinde kalmıştır ve güven aralığı sıfırı kapsamaktadır. Çoklu sınama düzeltmesinden sonra hiçbir bulgu ayakta kalmamıştır. Dolayısıyla havuzlanmış eğilim testleri, betimsel tablolarda gözlenen dönem ve usul bazlı dalgalanmalara karşın, dört dönem katmanı boyunca istatistiksel olarak kararlı bir yön ortaya koymamıştır.

Çalışma 4. Feldman’ın model–nazire ilişkisi tezi, hesaplamalı benzerlik sıralamasıyla uzman incelemesine aday üretebilir mi?

Kuram 4: Model-nazire ilişkisinin hesaplamalı önceliklendirilmesi

Walter Feldman, *Music of the Ottoman Court* adlı çalışmasında erken Osmanlı çalgısal repertuvarının dönüşümünü tek çizgilik bir “ilerleme” olarak resmetmez (Feldman, 2024). Onun anlatısında kısa, tekrarlı ve görece bağımsız terimler — birkaç ölçülük, kendi içinde kapalı ezgi birimleri — zamanla daha uzun süreli bir makam seyri içinde bütünleşir; bu süreçte usul devriyle ezginin bağı, hane–mülazime düzeni ve besteciler arası model–nazire ilişkileri birlikte değişir. Nazire, bir eserin üslubunca yeni bir eser yazma pratiğidir; aynen kopya ile tamamen özgün yorum arasındaki geniş bir alanı kaplar. Nazire yazar besteci modelin makamını, usulünü, çoğu zaman biçim planını ve seyir görgüsünü paylaşır, ama ezgi malzemesini kendi cümleleriyle yeniden kurar. Feldman’ın anlatısında nazire yalnızca bir benzerlik etiketi değil, repertuvarın besteciler arasında nasıl dolaşıma girdiğini gösteren ilişkisel bir kanıttır.

Nazirenin kökeni burada anılmaya değer bir gerilim taşır: pratik, edebiyat tarihinde güfteye ve aruz veznine bağlı bir Osmanlı şiir geleneği olarak doğmuştur — bir şairin başka bir şairin gazeline aynı vezin, kafiye ve mazmun düzeniyle yeni bir şiir yazmasıdır. Feldman bu terimi güftesiz, saf çalgısal formlara (peşrev, saz semâîsi) taşırken, edebi nazirenin sözcük-vezin-kafiye eksenini makam-usul-seyir eksenine dönüştürür (Feldman, 2024). Bu aktarımın kendisi Feldman’ın metninde sorgusuz bırakılmaz, ama burada da tartışılmadan geçilmemelidir: bir şiire yazılan nazire dinleyicide doğrudan tanınabilir sözel/vezinsel bir yankı bırakırken, çalgısal bir nazirenin izi yalnızca yapısal ve seyirsel dokuda aranabilir — dolayısıyla ölçülen “yapısal parmak izi benzerliği”nin nazireyi zayıf yakalaması, yöntemin başarısızlığından çok, edebi kavramın çalgısal alana aktarılırken kendi doğrulama ölçütünü de değiştirmiş olmasından kaynaklanıyor olabilir.

Bu kuramın hesaplamalı müzikoloji içinde test edilmiş bir yeri neredeyse yoktur. Bozkurt, Ayangil ve Holzapfel (2014), Türk makam müziği üzerine sembolik veriyle yapılan hesaplamalı araştırmaların son on yılda makam tanıma, usul çözümleme ve yüzeysel benzerlik ölçümünde ilerleme kaydettiğini, ama bu birikimin nazire gibi çok katmanlı — biçim planı, seyir yürüyüşü ve doku görgüsünü birlikte gerektiren — bir ilişkiyi önceliklendirme aracına çevirmediğini belirtir. Çalışma 4’ün araştırma sorusu buradan doğar: makam, usul, biçim ve seyir benzerliği üzerinden olası nazire çiftleri, uzman incelemesi için önceliklendirilebilir mi? Bu soru bilinçli olarak bir hipotez sınaması değil, bir aday önceliklendirme denemesi olarak çerçevelenir; “yeni nazireler bulundu” iddiası taşımaz. Doğrulanmış nazire etiketi hiçbir modern külliyatta yoktur; dolayısıyla test edilebilecek olan Feldman’ın tarihsel tezinin kendisi değil, bu tezin notada iz bırakması beklenen yüzeysel vekilleridir: aynı makam–usul–form paylaşımı, benzer yapısal doku istatistikleri ve benzer seyir yürüyüşleri.

Kuramın Testi

Örneklem, SymbTr türevi açık sembolik koleksiyondaki (Karaosmanoğlu, 2012) peşrev ve saz semâîleriyle sınırlıdır; analiz birimi nota değil eserdir. Üç dışlama kuralı uygulanmıştır: en az 20 geçerli nota içermeyen kayıtlar, besteci eşlemesi bulunmayan kayıtlar ve dört atıf katmanından birine yerleştirilemeyen kayıtlar elenmiştir. Kalan örneklem 149 eser ve 77 bestecidir; eserler besteci atfına göre dört dönem katmanına ayrılmıştır (1650 öncesi 14, 1650–1750 arası 18, 1750–1850 arası 21, 1850 sonrası 96 eser).¹ Bu dört kesim noktası (1650, 1750, 1850), besteci yaşam aralıklarından

türetilmiş, çalışmalar arasında ortak bir konvansiyondur; Feldman'ın kendi anlatısındaki daha ayrıntılı dönem geçişlerinin kaba bir yaklaşıklığıdır ve metnindeki kesin yıl sınırlarının birebir karşılığı olduğu iddia edilmemektedir. Örneklem, olasılıklı bir tarihsel örnekleme değildir; seçilmiş-aktarım niteliği ve bunun hedef popülasyonu nasıl daralttığı Giriş'te ayrıntılandırılmıştır.

Analizi yürüten MakamKit, SymbTr türevi sembolik notayı okuyup makam, usul ve form üstverisiyle eşleyen, eser düzeyinde yapısal ve seyirsel öznitelikler çıkaran bir hesaplama analiz motorudur; bu çalışmada iki katmanlı bir benzerlik ölçümü için kullanılmıştır. Birinci katman, yapısal parmak izi benzerliğidir: doğrudan tekrar, aktarılmış tekrar, seyir pürüzlülüğü, karar-yakınsama, bölüm kenarı benzerliği, bölüm yön sürekliliği, ölçü çizgisi bağlantı oranı, ezgisel yoğunluk ve ses alanı (koma) olmak üzere dokuz özellik z-standartlaştırılıp (eksikler örneklem ortancasıyla doldurularak) kosinüs benzerliğiyle karşılaştırılmıştır. İkinci katman, yörünge benzerliğidir: her eser on iki eşit dilime bölünmüş, her dilimde süre-ağırlıklı perde ağırlık merkezi hesaplanmış, standartlaştırılmış on iki noktalık eğriler arasında kısıtsız klasik dinamik zaman bükme (DTW) uzaklığı ölçülmüştür (Sakoe ve Chiba, 1978); toplam maliyet $24'e$ bölünüp $e^{(-uzaklık)}$ dönüşümüyle 0–1 aralığına çevrilmiştir. Birleşik puan, %65 parmak izi ve %35 yörünge ağırlığıyla bu iki benzerliğin toplamıdır; ağırlıklar doğrulanmış nazire etiketi bulunmadığı için optimize edilmemiş, keşif tercihi olarak ilan edilmiştir. Adaylar yalnızca aynı makam–usul–form hücreleri içindeki eser çiftlerinde aranmıştır; bu filtre karşılaştırılabilirliği artırırken, farklı usul ya da forma dönüştürülmüş gerçek nazireleri baştan dışlar. Sıralamanın sağlığını denetlemek için tek bir istatistiksel test planlanmıştır: aynı-bestecili çiftlerin puanlarının farklı-bestecili çiftlerden yüksek olup olmadığını sınavan tek-yönlü Mann–Whitney U testi. Ağırlık seçiminin sıralamayı ne ölçüde belirlediğini görmek için ayrıca dört alternatif ağırlık profili (%25/%75, %50/%50, %65/%35, %75/%25) altında sıralama yeniden hesaplanmış ve referans profille (%65/%35) karşılaştırılmıştır. Tüm benzerlik değerleri 3 ondalığa yuvarlanmıştır; sıralama hattı tamamen belirleyicidir, rastgele örnekleme veya model kurma içermez.

Bulgular

Tarama, aynı makam–usul–form hücrelerinde 34 karşılaştırılabilir eser çifti üretmiştir. En yüksek birleşik puan 0,885, en düşük 0,068'dir.

Tablo 12. Birleşik puana göre ilk 10 nazire adayı çifti (34 çiftin tamamı veri paketinde).

Sıra	Makam	Usul	Form	Eser A	Eser B	Parmak izi benzerliği	Yörünge benzerliği	Birleşik puan
1	Segâh	Aksak Semai	saz semâisi	k1807 · Hızır Ağa	k1808 · Nâyi Osman Dede	0,896	0,863	0,885
2	Uşşak	Aksak Semai	saz semâisi	k2123 · Dede Salih Efendi	k2125 · Neyzen Aziz Dede	0,907	0,718	0,841
3	Hüseyni	Aksak Semai	saz semâisi	k0793 · Lavtacı Andon	k0795 · Tatyos Efendi	0,942	0,641	0,837
4	Nihavent	Aksak Semai	saz semâisi	k1454 · Mesut Cemil	k1455 · Vecdi Seyhun	0,877	0,683	0,809
5	Hicaz Hümayun	Aksak Semai	saz semâisi	k0596 · Tanburi İsak	k0597 · Veli Dede	0,746	0,747	0,746
6	Beyati	Aksak Semai	saz semâisi	k0190 · Kanuni Ömer Efendi	k0191 · Nâyi Osman Dede	0,770	0,695	0,744
7	Saba	Aksak Semai	saz semâisi	k1717 · Neyzen Aziz Dede	k1718 · Zaharya	0,741	0,739	0,740
8	Beyati	Aksak Semai	saz semâisi	k0190 · Kanuni Ömer Efendi	k0192 · Mısırlı Udi İbrahim Efendi ("Arap")	0,777	0,660	0,736
9	Uşşak	Aksak Semai	saz semâisi	k2124 · Kantemiroğlu	k2125 · Neyzen Aziz Dede	0,746	0,708	0,733
10	Segâh	Devr-i Kebir	peşrev	k1764 · Nâyi Osman Dede	k1765 · Neyzen Yusuf Paşa	0,635	0,861	0,714

Birleşik puan formülü kalemle doğrulanabilir (birinci satır: $0,65 \times 0,896172 + 0,35 \times 0,862931 = 0,884538 \approx 0,885$). Listedeki en çarpıcı bulgu sütunlarda değil altlıktadır: 34 çiftin hiçbirisi aynı besteciye atfedilmemiştir.

Tablo 13. Benzerlik tanılaması: aynı-bestecili kontrol denetiminin sonucu.

Gösterge	Değer
Karşılaştırılabilir çift sayısı	34
Aynı bestecili çift sayısı	0
Aynı bestecili çiftlerin ortanca puanı	— ^a
Farklı bestecili çiftlerin ortanca puanı	0,589
Farklı bestecili çiftlerin çeyrekler açıklığı (Ç1=0,246; Ç3=0,728)	0,482
Mann–Whitney U	— ^a
Tek yönlü p (aynı besteci daha yüksek)	— ^a

^a Kontrol çifti bulunmadığından hesaplanamamıştır.

Sıralamanın sağlığını denetlemek için planlanan tek test, havuzda aynı besteciye ait tek bir çift bile bulunmadığı için çalıştırılmamıştır; bu durum bir yöntem başarısızlığından çok, saz semâisi türünün besteci başına genellikle tek eser üretme geleneğinin dolaysız bir sonucudur.

Tablo 14. Aday havuzunun yapısı (34 çift üzerinden)

Gösterge	Değer
Makam–usul–form hücresi sayısı	21
Tek çift içeren hücre	16
Üç çift içeren hücre	4
Altı çift içeren hücre	1 (Mahur)
Havuzdaki eser sayısı	48 (külliyyatın %32,2'si)
Hiçbir çifte giremeyen eser	101 (külliyyatın %67,8'i)
Aksak Semai usullü saz semâisi çifti	30 (%88,2)
Devr-i Kebir usullü peşrev çifti	4 (%11,8)
Parmak izi benzerliği aralığı	–0,344 ... 0,942
Yörünge benzerliği aralığı	0,634 ... 0,863
Birleşik puan aralığı	0,068 ... 0,885

Havuzun içi dengesizdir: 149 eserin üçte ikisi hiçbir çifte girmez ve çiftlerin onda dokuzu tek bir usul-form kombinasyonuna, Aksak Semai usullü saz semâisine düşer. İki bileşenin ayırt etme gücü de farklıdır: parmak izi benzerliği neredeyse zıt kutuptan (–0,344) güçlü örtüşmeye (0,942) uzanırken, yörünge benzerliği dar bir banda (0,634–0,863) sıkışmıştır; bu, on iki noktalık yumuşak eğrilerin doğuştan birbirine benzediğine ve yörünge bileşeninin tek başına ayırt edici gücünün sınırlı kaldığına işaret eder.

Tablo 15. Ağırlık profiline duyarlılık: referans profil (%65 parmak izi / %35 yörünge) ile karşılaştırma.

Ağırlık profili (parmak izi/yörünge)	Sıra korelasyonu (referansla)	İlk 10 örtüşmesi
%25 / %75	0,924	0,8
%50 / %50	0,992	1,0
%65 / %35 (referans)	1,0	1,0
%75 / %25	0,997	1,0

Sıralama, ağırlık seçimine görece dayanıklıdır: yörünge bileşenine aşırı ağırlık verildiğinde (%75) bile sıra korelasyonu 0,92'nin üzerinde kalır ve ilk-10 listesinin en az sekizi değişmez; parmak izi ağırlığı arttıkça (%50 ve üzeri) ilk-10 listesi tamamen sabitlenir. Bu, listenin tepesindeki adayların tek bir keyfi ağırlık seçiminin ürünü olmadığını, iki bileşenin çoğunlukla aynı yönde işaret ettiğini gösterir; ama sıralamanın dayanıklı olması, sıralamanın doğru olduğu anlamına gelmez ve Tablo 13'teki boş kontrol hücreleri bu ayrımı hatırlatır.

Özetle, tarama 149 eserlik külliyyattan 34 karşılaştırılabilir çift üretmiş, bu çiftlerin hiçbirinde aynı besteci yer almamış ve sıralama dört farklı ağırlık profili altında büyük ölçüde kararlı kalmıştır³.

Çalışma 5. Feldman'ın dört-dönem dönemlendirmesi, tarih ve besteci adı gizlenmiş müzikal özelliklerden kör kümeleme ile geri kazanılabiliyor mu?

Kuram 5: Dört dönemlik biçimsel periyodizasyon

Walter Feldman, *Music of the Ottoman Court: Makam, Composition and the Early Ottoman Instrumental Repertoire* (2024) adlı eserinde, Giriş'te özetlenen terkiib-seyir dönüşümünü tek bir doğrusal “ilerleme” olarak değil, tonal merkezlerin sırası, usul devriyle ezgi arasındaki bağ, hane–mülazime sürekliliği ve nazire ilişkileri gibi birbirleriyle eşzamanlı ilerleyen bileşenlerin ortak evrimi olarak ele alır. Çalışma 5'i doğrudan ilgilendiren veçhe, bu dönüşümün nihayetinde repertuarın dönemlere ayrılabilirliğini de içermesidir: Feldman'ın kendi dönemlendirmesi sekiz alt döneme kadar inceliklidir; ancak kuramın temel iddiası, bu dönemlerin yalnız arşiv kronolojisinde değil, eserlerin kendi biçimsel dokusunda da okunabilir izler bırakmış olmasıdır.

Bu iddia, tarihsel müzikolojide sıkça karşılaşılan bir epistemik soruna işaret eder: üslup dönemlendirmesi çoğunlukla belge, biyografi ve gelenek aktarımına dayanır; eserin kendi iç yapısının bu dönemlendirmeyi bağımsız olarak doğrulayıp doğrulamadığı ayrıca sınırlıdır. Feldman'ın kuramı, eğer doğruysa, dönem bilgisi saklandığında dahi

³ Dönem kırılımı ve besteci sayısı (77), sonuç veri dosyasında ayrı bir alan olarak bulunmadığından yalnızca toplamı (149) üzerinden dolaylı biçimde doğrulanabilmektedir.

eserlerin ölçülebilir özelliklerinin (motif, seyir, bölüm yapısı, usul-melodi eklemlenmesi) döneme göre ayrışabilir bir örüntü sergilemesini öngörür. Çalışma 5 bu öngörü, Feldman'ın kuramının kendisini değil, ondan türetilen sınırlı ve yanlışlanabilir bir alt-önermeyi sınavarak ele alır: *tarih ve besteci adı gizlendiğinde, müzikal özellikler dönem atıflarını yeniden kurabiliyor mu?*

Bu formülasyonun iki önemli sınırı baştan belirtilmelidir. Birincisi, kullanılan veri özgün 17.-18. yüzyıl yazmaları değil, modern notaya aktarılmış açık sembolik nota külliyatıdır; bağımsız değişken eserlerin kesin bestelenme tarihi değil, besteci atıflarından türetilen dört geniş dönem katmanıdır. İkincisi, Feldman'ın "Osmanlı-Türk üslubunun Persianate çevreden ayrıştığı" tezi bu çalışmada doğrudan test edilmemiştir; böyle bir sınama İran ve Transoksanya pişrev repertuarından bir karşılaştırma korpusu gerektirir ve bu korpus veri setinde bulunmamaktadır. Sınanan, bu geniş tezin çok daha dar bir önkoşuludur:

- Modern aktarımdaki çalgısal repertuar, atfedilen dönemleri taşıyan ölçülebilir bir biçimsel yapı gösteriyor mu?

Kuramın Testi

Örneklem. Analiz, 149 peşrev ve saz semaisinden oluşan bir örneklem üzerinde yürütülmüştür. Eserler, modern notaya aktarılmış, açık biçimde erişilebilen bir sembolik nota külliyatından (Karaosmanoğlu, 2012'nin tanımladığı türde bir kaynak; ticari olmayan araştırma kullanımı, atıf zorunlu) alınmıştır. En az yirmi geçerli nota içermeyen, besteci eşlemesi bulunmayan veya dört atıf katmanından birine yerleştirilemeyen kayıtlar analiz dışı bırakılmıştır. Analiz birimi bağımsız nota değil, eserin bütünüdür. Nihai örneklem dört dönem katmanına şöyle dağılmaktadır: 1650 öncesi 14 eser, 1650–1750 arası 18 eser, 1750–1850 arası 21 eser ve 1850 sonrası 96 eser; toplam 77 farklı bestecinin eserlerinden oluşmaktadır. Bu dağılımın belirgin dengesizliği, özellikle erken dönem hücrelerinin küçüklüğü (14 ve 18 eser) ileride ayrı bir sınırlılık olarak tartışılmaktadır. Bu dört sınırın Feldman'ın kendi anlatısındaki hangi dönem geçişlerine birebir karşılık geldiği kaynak metinde sayfa düzeyinde gösterilmemektedir; kullanılan dört dönem katmanı, Feldman'ın sekiz alt döneme kadar uzanan ince taneli dönemlendirmesinin kaba bir yaklaşıklığıdır ve birebir bir eşleme olarak sunulmamaktadır.

Bu örneklemin Feldman'ın kastettiği tarihsel evrenin olasılıklı bir örnekleme olmadığı giriş bölümünde ayrıntılı tartışılan seçilmiş-aktarım niteliğini burada da taşıdığı, özellikle vurgulanmalıdır: sonuçların hedef kütlesi "bütün Osmanlı çalgısal repertuarı" değil, tanımlanan aktarım zincirinde gözlemlenebilen eserlerdir.

MakamKit ile özellik çıkarımı. MakamKit, Türk makam müziği için geliştirilmiş, sembolik ve ses tabanlı çok katmanlı bir hesaplamalı analiz motorudur; bu çalışmada yalnız sembolik katmanı kullanılmıştır. Her eserden şu özellik aileleri çıkarılmıştır: (i) aralık motifleri: üç ve dört aralıklı pencerelerde doğrudan ve transpoze tekrarların ayrımı; (ii) seyir yörüngesi: eser eşit nota-sayılı on iki pencereye bölünerek her pencerede süre-ağırlıklı göreceli perde merkezi, ardışık pencereler arası yörünge pürüzlülüğü, yön değişim sayısı ve son beşte birlik kesitte nihai perdeye yakınsama derecesi hesaplanır; (iii) bölüm sınırları:hane, mülazime/teslim ve serhane olarak işaretlenen geçişlerde sekiz notalık bağlamaların histogram benzerliği ve yön sürekliliği; (iv) usul-melodi eklemlenmesi:nota başlangıcının ölçü içi fazı, ölçü çizgisini aşan nota olayları, nihai perdenin ölçü sonuna denk gelişi ve bölüm aralıklarının ölçü cinsinden yayılımı. Bu ölçümlerin hiçbirisi tarihsel usul vuruş ağırlıklarının veya özgün icra tavrının doğrudan karşılığı değildir; yalnız modern notalama yüzeyinin ölçülebilir zaman ve perde örgüsünü temsil eder.

Kör tasarım. Kümeleme modeline besteci adı, dönem etiketi, eser başlığı veya tarih hiçbir biçimde verilmemiştir; yalnız motif, seyir, bölüm, usul-melodi, yoğunluk, ambitus ve uzunluk özellikleri kullanılmıştır. Özellikler standartlaştırıldıktan sonra $k=2$ 'den $k=8$ 'e kadar KMeans, elli farklı rastgele başlangıçla çalıştırılmıştır. Küme iç sıklığı silhouette katsayısı (Rousseeuw, 1987) ile; kümelerin dört dönem etiketiyle dış uyumu düzeltilmiş Rand indeksi ARI (Hubert ve Arabie, 1985) ve normalize karşılıklı bilgi NMI ile ölçülmüştür. Bu iki ölçüm ayrı sorulara cevap verir: silhouette, verinin kendi içinde doğal adalar oluşturup oluşturmadığını; ARI ve NMI ise (bu adalar varsa) onların tarihsel dönemlerle örtüşüp örtüşmediğini gösterir.

Sınıflandırma. Kümelemeden bağımsız olarak, müzikal özelliklerin dönem bilgisini taşıyıp taşımadığı besteci-gruplu beş katlı çapraz doğrulamayla ayrıca sınanmıştır. Aynı bestecinin eserleri hiçbir zaman aynı anda hem eğitim hem test katında yer almamıştır; bu önlem, modelin besteci imzasını (kişisel üslup) dönem sinyali sanmasını engellemek içindir. Dört dönemin tümü için dengeli doğruluk raporlanmış; ayrıca erken iki dönemi (1650 öncesi ve 1650-1750) geç iki dönemden (1750-1850 ve 1850 sonrası) ayıran daha kaba bir ikili sınıflandırma için rastgele orman (random forest; Breiman, 2001) tabanlı ROC-AUC hesaplanmıştır, 500 ağaç, yaprak başına en az iki örnek ile. Bu ikili ayırım, Feldman'ın sekiz dönemlik özgün dönemlendirmesinin tam bir testi değildir; yalnız kaba bir erken/geç gradyanının varlığını sınar.

Belirsizlik denetimi. Tüm tanılamalar için besteci-bloklı permütasyon testi (1.000 yineleme, sabit rastgele tohum 20260720) ve besteci-bootstrap %95 güven aralıkları hesaplanmıştır (Efron ve Tibshirani, 1994); her ikisinde de aynı besteciye atfedilen eserler birlikte yeniden örneklenmiş, böylece üretken bestecilerin yapay kesinlik yaratması önlenmiştir. Permütasyon testinde her yinelemede karıştırılan yalnızca besteci→dönem atamasıdır (bir besteciye ait tüm eserler o yinelemede aynı rastgele dönem etiketini alır); kümeleme etiketleri ile çapraz-doğrulanmış sınıflandırma tahminleri ve olasılıkları sabit tutulur — yani KMeans ve rastgele orman modelleri her yinelemede yeniden eğitilmez, yalnızca gözlenen (sabit) model çıktısı rastgele dönem atamalarıyla yeniden karşılaştırılır. Bu tasarım, sabit model çıktısının rastgele bir dönem atamasından beklenenden daha yüksek uyum gösterip göstermediğini sınar; modelin kendisinin rastgele etiketlerle yeniden eğitildiğinde benzer bir uyumu üretilip üretemeyeceğini sınamaz. Bu ayırım, aşağıdaki Sınırlılıklar bölümünde raporlanan p-değerlerinin olası iyimserliği bakımından ayrıca ele alınmaktadır.

Bulgular

Tablo 16. Küme sayısı taraması, k=2–8

k	Silhouette	ARI (dört dönem)	NMI (dört dönem)
2	0,1349	0,0164	0,0338
3	0,1136	0,0522	0,0651
4	0,1149	0,0462	0,0746
5	0,1146	0,0150	0,0720
6	0,1149	0,0299	0,0702
7	0,1174	0,0289	0,0812
8	0,1110	-0,0036	0,0813

En yüksek iç sıklık iki kümeli çözümde görülmekte, fakat 0,1349 düzeyindeki bu değer düşük kalmaktadır ve özellik uzayında belirgin doğal adalar bulunmadığına işaret eder; hiçbir k değerinde ARI dört dönemle güçlü bir örtüşmeye yaklaşmamaktadır.

Tablo 17. Çapraz doğrulama karışıklık matrisi, dört dönem (n=149)

Gerçek dönem	Tahmin: 1	Tahmin: 2	Tahmin: 3	Tahmin: 4
1 (1650 öncesi, n=14)	0	1	0	13
2 (1650–1750, n=18)	0	4	3	11
3 (1750–1850, n=21)	1	3	0	17
4 (1850 sonrası, n=96)	2	2	3	89

Model, dört dönemden hiçbirini güvenilir biçimde ayırt edememekte, eserlerin büyük çoğunluğunu dönem etiketinden bağımsız olarak dördüncü sınıfa (1850 sonrası) yönlendirmektedir; bu eğilim, en kalabalık sınıfın (n=96) diğer üçüne oranla yapısal ağırlığından kaynaklanıyor olabilir.

Tablo 18. Dönemlendirme tanılama özeti (n=149, besteci sayısı=77)

Ölçüt	Değer
Dört dönem dengeli doğruluk	0,2873
Erken/geç ikili ayırım ROC-AUC	0,6976
ARI (k=4)	0,0462
NMI (k=4)	0,0746
Silhouette (k=4)	0,1149
Permütasyon tekrar sayısı	1.000
Permütasyon tohumu	20260720

Dört dönem sınıflandırmasının dengeli doğruluğu (0,2873), rastlantı tabanı olan 0,25'in yalnız sınırlı ölçüde üzerindedir; buna karşılık erken/geç ikili ayrımın ROC-AUC değeri (0,6976) daha belirgin bir sinyale işaret etmektedir, Tablo 19'daki belirsizlik denetimi bu ayrımın istatistiksel olarak da tutarlı olduğunu gösterir.

Tablo 19. Belirsizlik ve permütasyon denetimi

Tanımlama	Gözlenen	%95 besteci-bootstrap GA	Permütasyon p
Dört kümeli ARI	0,0462	[0,0076, 0,1194]	0,0020
Dört kümeli NMI	0,0746	[0,0418, 0,1711]	0,0040
Dört dönem dengeli doğruluk	0,2873	[0,2392, 0,3844]	0,1159
Erken/geç ROC-AUC	0,6976	[0,5807, 0,8019]	0,0060

Permütasyon p-değerleri, ARI ve NMI'nin rastlantısal etiket dağılımlarından istatistiksel olarak ayrıldığını ($p=0,0020$; $p=0,0040$) fakat bu ayrılığın mutlak büyüklüğünün küçük kaldığını gösterir; dört dönem dengeli doğruluğu için permütasyon testi anlamlılık eşiğini geçmemektedir ($p=0,1159$), oysa erken/geç ROC-AUC istatistiksel olarak tutarlı bir sinyal taşımaktadır ($p=0,0060$) — bu, çalışmanın “kaba bir tarihsel gradyana sınırlı destek” sonucunun sayısal temelidir.

Özetle, bulgular ince taneli bir dönemsel biçimlenmenin modern notada bulunmadığını, buna karşılık daha kaba bir erken/geç tarihsel gradyanın izinin sürülebildiğini göstermektedir.

Çalışma 6. Feldman'ın Mevlevi ayini Birinci Selâm'da biçimsel uzunluk ile tarihsel üslup arasındaki ilişki tezi, kaynak-ankrajlı bir doğrulama protokolüyle sınamaya hazır mı?

Kuram 6: Birinci Selâm'da biçim-üslup ilişkisi

Walter Feldman, *From Rumi to the Whirling Dervishes: Music, Poetry, and Mysticism in the Ottoman Empire* (2022) adlı kitabında Mevlevi ayininin Birinci Selâm bölümünü, bestecinin teknik birikimi ile müzikal vizyonunun en yoğun biçimde gözlemlenebildiği bir alan olarak okur. Feldman'a göre Birinci Selâm'ın biçimsel uzunluğu ile ait olduğu bestecinin tarihsel konumu arasında yönlü bir ilişki vardır: kadim veya erken örneklerde anonim Dügâh Ayini gibi mısra ve terennüm bölümleri birkaç devr-i revân çevrimi içinde derli toplu düzenlenirken, daha geç dönem bestecilerinde melodik cümlelerin çevrim sınırlarını aşması ve makam bölgesinin genişlemesi beklenir. Bu okuma, yalnızca ölçü sayısının artmasına indirgenebilecek yüzeysel bir gözlem değildir; Feldman'ın vurguladığı, uzun erimli melodik düşüncenin usul çevrimini kapsayarak yeniden örgütlenmesi, yani söz-usul-seyir ilişkisinin bestecinin elinde daha karmaşık bir mimariye kavuşmasıdır.

Kuram, üç unsurun eş zamanlı izlenmesini gerektirir: cümle sınırının nerede bittiği, bu sınırın devr-i revân çevrimiyle ilişkisi ve makam bölgesinin (geçkinin) kaç kez değiştiği. Feldman'ın (2022) basılı metni bu unsurları Dügâh ve Beyati ayinleri için kısmi biçimsel açıklamalarla destekler. Saba Ayini içinse notaların kendisini basmaz; okuru web ekine ve İrsoy'un *Mevlevi Ayinleri* (cilt 11, s. 565–568) kaynağına yönlendirir. İrsoy'un bu derlemesi için makalenin ve kaynak dokümanların hiçbirinde tam bir bibliyografik künye (yazarın tam adı, yayın yılı, yayınevi) bulunmadığından, bu çalışma kaynağı yalnızca Feldman'ın (2022) atfı üzerinden, cilt ve sayfa bilgisiyle anılmaktadır. Feldman'ın 2024 tarihli *Music of the Ottoman Court* eseri de bu çalışma paketinin kaynakçasında yer almakla birlikte, Birinci Selâm'a ilişkin tez doğrudan 2022 kitabından türetilmiştir.

Bu çerçeveden hareketle araştırma sorusu şöyle kurulmuştur: kaynakla doğrulanmış Dügâh, Beyati ve Saba Birinci Selâm notaları aynı ölçüm sözleşmesiyle kodlandığında, kronolojik sıra (a) cümlelerin devr-i revân çevrimlerini aşması ve (b) bağımsız makam bölgelerinin çoğalması yönünde betimsel bir örüntü üretir mi? Üç önceden tanımlı hipotez bu soruyu işler hale getirir. H1, kronolojik sırada uzman-kodlu ortalama cümle uzunluğunun (devr-i revân cinsinden) arttığını öngörür; H2, en az bir usul çevrimi sınırını kapsayan cümlelerin oranının arttığını öngörür; H3 ise bağımsız uzman makam-bölgesi etiketlerinden hesaplanan bölge/geçki sayısının arttığını öngörür. Karar kuralı önceden sabitlenmiştir: üç gösterge de beklenen yöndeysse sonuç descriptive_support, en az biri ters yöndeysse descriptive_counterexample olarak kaydedilecektir. Üç vaka söz konusu olduğundan p-değeri veya nüfus genellemesi üretilmesi zaten planlanmamıştır. Bu, Feldman'ın önerdiği ilişkiyi seçilmiş-vaka düzeyinde bir replikasyon denemesi olarak konumlandırır, olasılıklı bir örneklem iddiası olarak değildir.

Kuramın Testi

Örnekleme, kronolojik ve kuramsal karşıtlık gözetilerek üç vaka ile sınırlı tutulmuştur: anonim ve kadim kabul edilen Dügâh Ayini (erken dönem), Köçek Mustafa Dede'ye atfedilen Beyati Ayini (orta dönem) ve İsmail Dede Efendi'nin 1823 tarihli Saba Ayini (geç dönem, kesin tarihli). Bu üçlü, Feldman'ın tezinin ima ettiği erken-geç karşıtlığını hem kronolojik hem kuramsal olarak temsil edecek şekilde seçilmiştir; ancak Osmanlı Mevlevi ayin repertuvarının olasılıklı bir örnekleme değildir. Makalenin kendisi ve bu bölüm bunu açık bir sınırlılık olarak kabul eder.

Analiz, Giriş'te tanıtilen beş katmanlı (sembolik melodik analiz, ses tabanlı perde analizi, usul/aksak analizi, güfte/duygu analizi ve benzerlik-kümeleme) MakamKit motorunun sembolik katmanı üzerinden yürütülmesi planlanan bir protokol olarak tasarlanmıştır. Bu çalışmada motorun rolü doğrulayıcı bir aracınkiyle sınırlıdır. Dört adımdan oluşan bu doğrulama şöyle işler: kaynak-izlenebilir MusicXML (W3C Music Notation Community Group, 2021) dosyalarını şema doğrulamasından geçirmek; uzman tarafından işaretlenmiş cümle sınırlarını ilgili devr-i revân çevrim uzunluğuna bölerek cümle uzunluğunu hesaplamak; açık aralıkta kalan usul sınırlarını sayarak çevrim aşımını belirlemek; bağımsız uzman makam-bölgesi etiketlerinden bölge/geçki sayısını türetmek. Motor, uzman sınırı bulunmayan durumlarda süre-temelli sezgisel bir bölümlenme üretebilir; ancak bu çıktı yalnızca yazılım tanılması içindir, phrase_source: heuristic olarak işaretlenir ve doğrulayıcı H1/H2 kararına giremez. Aynı ilke makam genişlemesi için de geçerlidir: algoritmik perde histogramı uzman etiketlerinin yerine geçmez; eksik değer sıfıra çevrilmez, n_modulations: null olarak bırakılır. Bu tasarım tercihi, ölçülemeyen bir değişkeni sıfır sanma hatasını önlemek içindir.

Doğrulayıcı koşunun açılabilmesi dört veri-hazırlık kapısının aynı anda karşılanmasına bağlanmıştır: (1) üç vakanın kaynak-izlenebilir MusicXML dosyalarının bulunması, (2) her dosyaya uzman phrase_end_qL dizisinin eklenmiş olması, (3) bağımsız region_labels etiketlerinin sağlanması, (4) nota/cümle/bölge kodlarının kaynak sayfaya karşı çift kontrolden geçirilmiş olması. Şema doğrulaması bilinmeyen açıklama alanlarını reddedecek şekilde sıkı tutulmuştur. Kalite güvencesi ayrıca aktarımı yapan kişi ile ikinci denetçinin ayrı tutulmasını, anlaşmazlıkların kör uzlaşısı öncesinde korunmasını ve güvenilirliğin ayrıca raporlanmasını öngörür; uzlaşısı veri seti ile anlaşmazlıkları-sınır-çıkarılmış duyarlılık koşusu ayrı ayrı sunulacak, sonuç işaret değiştirirse “kodlama duyarlılığı” olarak sınıflandırılacaktır. Üç göstergeli birleşik karar kuralı dışında resmi bir çoklu karşılaştırma düzeltmesi (örn. Bonferroni) uygulanmamıştır, çünkü üç vaka nedeniyle zaten hiçbir p-değeri üretilmemektedir.

Bulgular

Bu bölüm, olağan anlamıyla bir “olumlu sonuç” raporu değildir: aşağıda sunulan bulgular, kuramın sınanıp desteklendiğini veya çürütüldüğünü değil, sınama koşullarının henüz hazır olmadığını belgeler. Doğrulayıcı örnekleme büyüklüğü sıfırdır; confirmatory_ready=false ve historical_test_status=not_run_missing_real_scores değerleri makine-okunur sonuç dosyasında sabittir. H1, H2 ve H3'ün üçü de test edilmemiştir. Bu, hipotezlerin çürütüldüğü değil, sınanmadığı anlamına gelir. Tablo 20 ve Tablo 21'in ve sonuç dosyasındaki durum alanlarının, sonuç JSON'u ile çapraz kontrol edilmiş birebir aktarımıdır.

Tablo 20. Vaka bazında veri hazırlığı

Vaka	Gerçek nota hazır	Yayımlanmış ankraj	Kanıt kapsamı
Dügâh	Hayır	6	Kısmi basılı örnek ve metinsel çözümleme
Beyati	Hayır	7	Kısmi basılı örnek ve metinsel çözümleme
Saba	Hayır	1	Yerel kitap kaynağı tanımlar; nota web eki veya İrsoy cildindedir

Tablo 20, hiçbir vaka için kaynak-izlenebilir gerçek notanın hazır olmadığını, buna karşın Feldman'ın (2022) basılı metninden doğrulanabilir 6, 7 ve 1 adet yayımlanmış önermenin (ankraj) çıkarılabildiğini göstermektedir. Ankraj, Feldman'ın basılı metninde belirli bir bölüme atfedilen, kaynak sayfasıyla denetlenebilir somut bir önermeyi ifade eder (örn. “Dügâh Bend III'te üçüncü devirde Segâh geçkisi”); bu ankrajlar veri edinimini denetlemek için kullanılabilir, ancak kendileri hipotez testi için yeterli veri niteliği taşımaz.

Tablo 21. Doğrulayıcı karar kapıları

Kapı	Zorunlu	Güncel durum
Kaynak-izlenebilir üç nota	Evet	Karşılanmadı
Uzman cümle sınırları	Evet	Karşılanmadı
Bağımsız makam-bölgesi etiketleri	Evet	Karşılanmadı
Çift kaynak kontrolü	Evet	Karşılanmadı
Doğrulayıcı analiz	Evet	Çalıştırılmadı

Tablo 21, dört veri-hazırlık kapısının tümünün ve bunlara bağlı doğrulayıcı analiz adımının henüz karşılanmadığını göstermekte; bu durum, tezin yürütüldüğüne değil, sınama koşullarının henüz oluşmadığına işaret etmektedir.

Tablo 22. Protokol durum bayrakları

Gösterge	Değer
Gerçek nota sayısı	0
Eksik hedef sayısı (Dügâh, Beyati, Saba)	3
confirmatory_ready	Hayır
historical_test_status	not_run_missing_real_scores
protocol_status	Aşama 1 kayıtlı rapor, gerçek nota eksik
Figürler	Yok

Tablo 22, protokolün kendi durum bayraklarını özetlemekte ve makalede sunulan hiçbir sayısal veya görsel bulgunun tarihsel bir sonuç iddiası taşımadığını teyit etmektedir; sentetik veri üzerinde çalışan yazılım öz-testlerinin geçmesi de bu değerlendirmeyi değiştirmez, çünkü bu testler yalnızca ayırıştırma, şema reddi, çevrim sayımı ve eksik-değer davranışını sınamaktadır, tarihsel kanıt üretmemektedir. Özetle, üç vakanın hiçbirisi için dört veri-hazırlık kapısı henüz karşılanmamış, doğrulayıcı analiz çalıştırılmamış ve H1–H3 test edilmemiştir; gerçek nota sayısı sıfır, eksik hedef sayısı üçtür.

Sonuç ve Tartışma

Bu makale, Walter Feldman'ın Osmanlı saray çalgı musikisine dair geniş dönüşüm anlatısından altı ayrı, sınırları önceden çizilmiş önerme türetmiş; bunları ortak bir hesaplama altyapı (MakamKit), ortak bir veri kaynağı (SymbTr türevi sembolik nota külliyatı) ve şeffaf, çoklu karşılaştırma düzeltmeli istatistiksel protokollerle sınamıştır. Altı çalışmanın toplu okuması tek bir cümleye indirgenemez, ama iç tutarlılığı yüksek bir tablo ortaya koyar: beş çalışma modern nota yüzeyinde ya destek bulamamış ya da yalnızca kaba, sınırlı bir sinyal yakalamış; altıncısı ise henüz sınanabilir hâle gelmemiş bir kuramın veri hazırlık aşamasını belgelenmiştir.

Tablo 23. Altı çalışmanın toplu epistemik statüsü.

Çalışma	Kuram	Statü	Ana bulgu
Çalışma 1	Terkibden seyre geçiş (motif tekrarı, cümle yayılımı, karar-yakınsama)	Desteklenmiyor; bileşene ve modele duyarlı	Birleşik destek ölçütü ($\geq 3/5$ gösterge) karşılanmadı; tek kararlı gösterge (karar-yakınsama) beklenenin tersi yönde; saz semâisinde cümle yayılımı form-içi kısmi destek verdi
Çalışma 2	Seyrin bütünsel tarihsel kodifikasyonu (beş yakınsak gösterge)	Desteklenmiyor	Dört göstergede kararlı eğilim yok; karar-yakınsama ters yönde ($\rho = -0,325$, $p < 0,001$); beş hipotezin beşi de noktasal tahminde tezin öngördüğü yönün tersi işaretli
Çalışma 3	Usul devri ile ezginin eklemlenme biçiminin dönemsel değişimi	Modern nota yüzeyinde iz bulunamadı	Beş göstergenin hiçbirisi düzeltme sonrası anlamlı eğilim üretmedi (tüm $q \geq 0,570$)
Çalışma 4	Model-nazire ilişkisinin hesaplama önceliklendirilmesi	Doğrulanmış keşif yok; yalnız aday sıralaması	149 eserlik külliyattan yalnız 34 karşılaştırılabilir çift üretilebildi; sıfır aynı-bestecili kontrol çifti
Çalışma 5	Dört dönemlik biçimsel periyodizasyon	Doğal küme olarak geri kazanılmıyor; kaba gradyana sınırlı kanıt	Dengeli doğruluk 0,2873 (perm. $p = 0,1159$); erken/geç ROC-AUC 0,6976 ($p = 0,0060$)
Çalışma 6	Birinci Selâm'da biçim-üslup ilişkisi	Sınanmadı — Aşama 1 kayıtlı protokol	Gerçek nota sayısı sıfır; confirmatory_ready=false

Bu tablonun ilk okunuşu, Feldman'ın kuramsal mirasının modern sembolik aktarım örnekleminde “kaybolduğu” biçiminde olabilir; bu okuma aceleci olur. Altı çalışmanın hiçbirisi kuramın kendisini yürütmemektedir. Her biri kuramın notaya indirgenmiş, dar bir vekil kümesiyle ölçülen bir izdüşümünü sınamıştır. Bu ayırımın gerekçesi burada bir kez, güçlü biçimde kurulmalı ve aşağıdaki altı tartışmanın hepsine zemin oluşturmalıdır: MakamKit'in girdisi olan SymbTr türevi külliyat (Karaosmanoğlu, 2012), on yedinci-on sekizinci yüzyıl icrasının olasılıklı bir örnekleme değildir; günümüze ulaşma, meşk zincirinde yaşatılma, notaya aktarılma, edisyona girme ve besteciye atfedilme süzgeçlerinden geçmiş, seçilmiş bir aktarım örneklemdir. Bu süzgecin her halkası bir seçimdir ve icranın esnekliğini süslemeleri, perde

ayarlamalarını, tempo serbestliğini, düm-tek vuruş ağırlığını sabit işaretlere indirger. Buna ek olarak, MakamKit'in perde ölçümlerinin dayandığı 53 komalı Arel-Ezgi-Uzdilek (AEU) dizgesi yirminci yüzyılda kuramsallaştırılmış bir çerçevedir; bu çerçevenin erken dönem eserlerine geriye dönük uygulanması, aktarım zincirinin taşıdığı belirsizlikten ayrı, ölçümün kavramsal temeline içkin kendine özgü bir anakronizm riski taşır. Bu iki sınır, aktarım süzgeci ve kavramsal anakronizm altı çalışmanın tamamında ortaktır ve aşağıda her çalışma yeniden anılmayacak, yalnızca her bulgunun bu ortak zemin üzerinde nasıl konumlandığı tartışılacaktır. Sonuçların büyük bölümünün “desteklenmiyor” ya da “sınırlı kanıt” biçiminde çıkması, bu nedenle Feldman'ın tarihsel önermesinin bir çürütülmesi değil, o önermenin modern aktarım yüzeyindeki izinin ve bu izin seçilen vekillerle ne ölçüde yakalanabildiğinin bir haritasıdır.

Terkibden seyre: bütünleşme denemesi (Çalışma 1 ve Çalışma 2)

Feldman'ın (2024) merkezi tezi, on yedinci yüzyılın ikinci yarısından itibaren peşrev ve saz semâisi repertuvarının, terki denen modüler hane-mülazime yapısından, eserin bütününü kucaklayan bir seyir mantığına doğru kaydığıdır. Çalışma 1 bu geçişi motif tekrarı, cümle yayılımı ve karar-yakınsama gibi göstergelerle; Çalışma 2 ise aynı geçişin tonal ağırlık merkezine dayalı beş göstergesiyle (pürüzlülük, yön değiştirme, bölüm sınırında yön sürekliliği, bölüm sınırında perde dağılımı benzerliği, karar-yakınsama) sınamıştır. İki çalışmanın da ortaya çıkardığı örüntü çarpıcı biçimde örtüşür: beklenen yönde hareket eden gösterge sayısı birleşik destek eşliğinin altında kalmış, buna karşılık karar-yakınsama her iki çalışmada da istatistiksel olarak anlamlı ama beklenen yönün tam tersine hareket etmiştir (Çalışma 2'de $\rho=-0,325$, $p<0,001$). Çalışma 2'de beş hipotezin beşi de noktasal tahminde ters işaretlidir; bu, tek bir göstergenin tesadüfen ters çıkmasından çok, tutarlı bir örüntüye işaret eder.

Bu tutarlı tersinme, Feldman'ın metniyle yüzleştirildiğinde iki farklı biçimde okunabilir. Birincisi, karar-yakınsama ölçütünün kendisi kaba bir vekildir: son notayı mekanik biçimde karar sayar, oysa teslim (mülazime) bazı formlarda asma bir kalışla kapanabilir; ölçüt bu inceliği ayırt etmez. Bu durumda tersinme, seyrin tarihsel olarak zayıfladığını değil, vekilin teslim bölgesini yanlış temsil ettiğini gösteriyor olabilir. İkincisi, ve Çalışma 2'nin sınırlılıklar bölümünün işaret ettiği gibi, karar-yakınsamanın dönem medyanlarındaki seyri tek yönlü değil, 1650-1750 katmanında bir tepe yapıp ardından gerileyen bir dalgalanmadır; bu, H1-H5'in kurduğu doğrusal beklenti çerçevesinin Feldman'ın kendi nüanslı, çok boyutlu anlatısından daha kısıtlayıcı olabileceğini düşündürür. Feldman'ın argümantasyonu zaten terki den seyre geçişi tek eksenli bir “ilerleme” olarak değil, perde merkezlerinin sırası, usul-ezgi eklemlenmesi, hane-mülazime düzeni ve nazire ilişkisinin birlikte yer değiştirdiği çok boyutlu bir süreç olarak kurar (Feldman, 2024); bu makalenin iki çalışması o geniş örgüden yalnızca tonal-yapısal iki ipliği çözmüştür. Bir ipliğin elde tutulmaması dokunun kendisinin var olmadığı anlamına gelmez. Ama iki bağımsız çalışmanın, iki farklı gösterge kümesiyle, aynı ters yönlü karar-yakınsama örüntüsünde buluşması, bu tekilliğin bir vekil kusuru olmaktan öte, aktarım zincirinin sinyali sistematik biçimde bir yönde çarpıttığı ihtimalini de güçlendirir. Bozkurt, Ayangil ve Holzapfel'in (2014) hesaplamalı makam araştırmaları için işaret ettiği genel sınır, sembolik ölçütlerin tarihsel-kuramsal kavramların ancak gölgesini yakalayabilmesi, burada iki bağımsız sınamada aynı yönde somutlaşmıştır. Bulgu tamamen boş da değildir: Çalışma 1'de saz semâisinde cümle yayılımı kendi form-içi düzeltme ailesinde tezle uyumlu, forma özgü kısmi bir sinyal vermiştir; bu, sonucun tek bir düz “yok” yargısına değil, bileşene ve forma duyarlı bir olumsuz/karma yargıya işaret ettiğini gösterir.

Usul devri ile ezginin eklemlenmesi (Çalışma 3)

Feldman'ın (2024) tezinin usul boyutu, düm-tek vuruşlarının ağırlık hiyerarşisi ile ezginin bu hiyerarşiye göre nasıl eklemlendiği üzerine kuruludur. Çalışma 3, ezginin usul devrini giderek daha uzun soluklu ve daha sık aşmasını beş göstergeyle sınamış; beşinin de düzeltme sonrası anlamlı bir dönem eğilimi üretmediğini (tüm $q\geq 0,570$) bulmuştur. Burada olumsuz sonuç, Çalışma 1 ve 2'dekinden farklı bir mekanizmayla açıklanabilir: kaynak veride düm-tek vuruş şiddetleri kodlanmadığından çalışma bir olay-fazı vekili kullanmak zorunda kalmıştır ve hiçbir gösterge usulün ağırlık hiyerarşisini doğrudan ölçmez. Ölçü çizgisinin kendisi, tarihsel usul devrinin değil, modern aktarımcının editoryal kararının bir izidir. Bozkurt vd.'nin (2014) genel uyarısı burada özellikle keskin bir biçim alır: sembolik veri zamanı

verir, vuruş ağırlığını vermez. Feldman'ın metnindeki terkiib-seyir dönüşümü meşk zincirinde bedenden bedene aktarılan bir bilgiye dayanırken, bu çalışmanın ölçtüğü şey o bilginin modern nota kâğıdına düşen gölgesidir.

Çalışma 3'ün sınırlılıklar bölümünde işaret edilen bir karıştırıcı, bu olumsuz sonucun okunuşunu daha da temkinli kılar: saz semâilerinin tamamı Aksak Semai usulündedir, dolayısıyla form ile usul etkisi bu veriyle ayrıştırılamaz; göstergeler usul döngü uzunluğuna duyarlıdır ve dönemlerin usul bileşimi farklılaştığından, havuzlanmış test bu karıştırıcının yükünü taşımaktadır. Ayrıca 1750-1850 dönemine denk düşen tek dönemlik bir “çukur” bağlantı oranı ve cümle yayılımının birlikte düşmesi Feldman'ın öngördüğü pencereye kabaca denk gelse de, kalıcı bir açılma değil V biçimli bir dalgalanmadır; dönüşüm tezi ise tek yönlü ve kalıcı bir eğilim öngörmektedir. Üç olası mekanizma aktarım zincirinin sinyali homojenleştirmesi, usul-döngü uzunluğunun döneme karışması ve dönüşümün doğrusal olmayan bir dalgalanma biçiminde seyretmesi veride ayrı ayrı karşılık bulur ve hiçbirisi bu makale kapsamında birbirinden ayrıştırılamaz. Üç yüzyıllık aktarım sürecinin nota üzerinde bıraktığı iz durağan görünmektedir; ancak bu durağanlığın tarihsel usul pratiğinin kendisinde bir durağanlığa işaret ettiği sonucuna doğrudan varılamaz.

Model-nazire ilişkisi (Çalışma 4)

Feldman'ın nazire kavramı, besteciler arası bir görgü ilişkisini, tek bir motif eşitliğine indirgenemeyecek bir kavram olarak sunar (Feldman, 2024). Çalışma 4, bu ilişkiyi hesaplamalı bir aday önceliklendirme problemine dönüştürmüş ve külliyatta tarihsel nazire etiketi bulunmadığından bir doğrulama değil, yalnızca bir sıralama üretmiştir. Listenin tepesinde duran Hızır Ağa-Nâyi Osman Dede Segâh saz semâisi çifti (0,885), iki bestecinin de on yedinci yüzyıl sonu Osmanlı saray çevresine ait olması ve Nâyi Osman Dede'nin dönemin eski repertuarını derleyip notaya aktaran konumu nedeniyle gerçek bir ilk durak olarak değerlendirilebilir. Ancak bu yüksek puan iki ayrı hikâyeye de uyumludur: gerçek bir model-nazire yakınlığı ya da ortak bir notacının aktarım görgüsü. Bu ayrımı çözmek makinenin değil, kaynak taramasıyla çalışan tarihçinin işidir.

Çalışmanın en öğretici bulgusu, planlanan tek doğrudan sağlık denetiminin (aynı bestecinin aynı hücreye yazdığı ikinci bir eserin kontrol grubu olarak kullanılması) veri yetersizliğinden hiç çalıştırılmamış olmasıdır: 149 eserlik külliyatta olası ikili karşılaştırma sayısı matematiksel olarak 11.026 iken, bunun yalnızca 34'ü aynı makam-usul-form hücrelerinde kalmış ve sıfır aynı-bestecili kontrol çifti üretilmiştir. Bu boşluk, Feldman'ın kuramının değil, modern kanonun kendi türsel mantığının bir yansımasıdır: saz semâisi on dokuzuncu-yirminci yüzyılda standartlaşmış bir form olduğundan, bu dönemin bestecileri neredeyse her makamda tek bir saz semâisi yazmıştır; “her makama bir saz semâisi” görgüsü, aynı bestecinin aynı hücreye ikinci bir eser yazmasını yapısal olarak engellemiştir. Bu, Feldman'ın nazire kuramının yanlışlığını göstermez. Tam tersine, modern kanonun kendi seçici mantığının tarihsel ilişkinin izini nasıl örtebileceğini gösterir. Ağırlık profiline duyarlılık denetimi, sıralamanın tek bir keyfi katsayı seçiminin ürünü olmadığını göstermiştir; ama iç tutarlılık dışsal geçerlik değildir. Bozkurt vd.'nin (2014) genel biçimde işaret ettiği boşluk burada kapatılmamış, yalnızca haritalanmıştır: modern aktarım kanonunun kendi türsel mantığının tarihsel model-nazire ilişkisinin izini ne ölçüde örttüğü, eldeki veriyle yanıtlanamayan, ama gelecekteki bibliyografik incelemenin doğrudan sıayabileceği açık bir sorudur.

Biçimsel dönemlendirme (Çalışma 5)

Feldman'ın (2024) sekiz dönemlik dönemlendirmesinin dörtlü bir indirgemesini sıayan Çalışma 5, dört kümeli bir çözümün dönem etiketleriyle örtüşmediğini ($ARI=0,0462$, $NMI=0,0746$) ve dört dönem sınıflandırmasının rastlantı tabanının yalnız sınırlı üzerinde kaldığını (dengeli doğruluk 0,2873; permütasyon $p=0,1159$, anlamlılık eşliğini geçmiyor) bulmuştur. Buna karşılık daha kaba bir erken/geç ayrımı istatistiksel olarak tutarlı, zayıf-orta düzey bir sinyal taşımaktadır ($ROC-AUC=0,6976$; $p=0,0060$). Bu ikili sonuç ilk bakışta çelişkili görünse de, Feldman'ın kendi kuramsal çerçevesiyle okunduğunda paradoksal biçimde tutarlıdır: Feldman'ın dönemlendirmesi sekiz ayrı, keskin sınırlı blok değil, terkiib biçiminin kısalığından seyir-temelli bütünleşmeye kademeli bir geçişi, örtüşen bir sürekliliği öngörür. Kümeleme yönteminin doğası gereği aradığı ayırık, birbirinden kopuk gruplar, bu türden kademeli bir dönüşümün doğru bir sınaması olmayabilir; erken/geç ikili ayrımın kümelemeden daha güçlü sinyal vermesi de bu okumayla uyumludur. Keskin sınıflandırma yerine kaba bir yön arandığında sinyal artmaktadır.

Bununla birlikte, düşük küme uyumunun alternatif bir okuması da göz ardı edilemez: modern aktarım sürecinin kendisi tarihsel üslup imzalarını homojenleştirmiş olabilir; Bozkurt vd. (2014), sembolik notaya aktarımın icra pratiğine özgü mikro-tonal ve ritmik ayrıntıları büyük ölçüde standartlaştırdığını vurgular. Bu iki okuma, kuramın kendi süreklilik veçhesi ve aktarım artefaktı birbirini dışlamaz, birlikte bulguyu açıklar. Erken/geç ayrımının kısmi başarısı da temkinle okunmalıdır: 0,6976 düzeyindeki bir ROC-AUC rastlantı düzeyinin (0,50) üzerinde olsa da, güven aralığının alt sınırının (0,5807) rastlantı düzeyine yakın seyretmesi, bunu güçlü bir dönemlendirme kanıtı saymayı engeller. Katkı, dönemleri “makineyle keşfetmek” değil, tarih etiketlerini modele hiç vermeden dönemlendirme iddiasını besteci-sızıntısına karşı denetimli bir çerçeveye sınımlamaktır; Feldman’ın kendi argümantasyon tarzı zaten büyük ölçüde filolojik ve karşılaştırmalı-yazmasal kanıta dayandığından (2024), bu çerçeve o kanıtlarla rekabet eden değil, ona tamamlayıcı ve dolaylı bir hesaplamalı sınaama sunmayı hedeflemektedir.

Birinci Selâm: kayıtlı protokol (Çalışma 6)

Çalışma 6, diğer beşinden köklü biçimde farklı bir epistemik konumdadır ve bu fark burada açıkça vurgulanmalıdır: Çalışma 1-5 SymbTr üzerinden derlenmiş modern sembolik aktarım örneklemiyle çalışmış ve betimsel veya çıkarımsal sonuçlar üretebilmiştir; Çalışma 6’da ise SymbTr veri kaynağı değildir, çünkü Mevlevî ayin repertuarı SymbTr korpusunun kapsamı dışındadır. Dügâh, Beyatî ve Saba ayinlerinin hiçbirisi için kaynak-izlenebilir gerçek nota henüz doğrulayıcı veri paketinde bulunmamaktadır; gerçek nota sayısı sıfırdır. Dügâh için 6, Beyatî için 7 ve Saba için 1 yayımlanmış ankraj doğrulanmış olsa da bunlar yalnızca veri edinim sürecini denetleyen kontrol noktalarıdır, sonuç üretme kapasitesine sahip değildir. Bu bölümün amacı bir kuram lehine ya da aleyhine karar bildirmek değildir; sunulabilecek tek dürüst değerlendirme, Feldman’ın Birinci Selâm tezinin desteklendiği ya da çürütüldüğü değil, henüz gözlemlenebilir kılınmadığıdır.

Bu durum tezin zayıflığını göstermez. Feldman’ın Dügâh ve Beyatî’nin bölüm uzunlukları, devr-i revân sayıları ve bazı makam hareketleri hakkındaki gözlemleri basılı, izlenebilir bir nesir kanıtına dayanır; ama nesir kanıtı ile makine-okunur analiz girdisi arasındaki fark, hesaplamalı Osmanlı müzikolojisinin yapısal bir sorununu görünür kılar. Feldman’ın “Dügâh Bend III’te üçüncü devirde Segâh geçkisi” türünden bir gözlemi belirli bir bölümü doğrulamak için kullanılabilir, ama eserin bütün Birinci Selâm’ını olay dizisine dönüştürmez. Bozkurt vd.’nin (2014) durum tespiti burada da geçerlidir: makam müziğinin hesaplamalı analiz altyapısı büyük ölçüde modern, ses veya nota kaydı bulunan repertuvara dayanır; tarihsel ve dinî-tören repertuarının bu altyapıya girişi ayrı bir aktarım emeği gerektirir. Bu bölümün bulgusu, tam olarak bu emeğin henüz sarf edilmediğidir. Çalışmanın asıl katkısı, tarihsel kuramı eksik veri üzerinde taklit etmemesidir: Nosek ve diğerlerinin (2018) kayıtlı rapor mantığı burada bir adım öteye taşınmış, veri toplama tamamlanmadan protokolün kendisi yayımlanmıştır; Munafò ve diğerlerinin (2017) ayırdığı iki katman — yazılımın yeniden üretilebilirliği ile tarihsel çıkarımın geçerliliği — burada özellikle nettir: MakamKit’in sentetik dosyalar üzerindeki öz-testlerinin geçmesi yalnızca yazılımın beklendiği gibi çalıştığını gösterir, Feldman’ın tarihsel savının doğrulandığı anlamına gelmez.

Genel değerlendirme

Feldman’ın kuramsal mirası bu altı sınaamadan nasıl çıkacağına baktığımızda ilk bakışta olumsuz görünen tablo, aslında Feldman’ın kendi yönteminin gücünü dolaylı biçimde teyit eder. Feldman’ın argümantasyonu büyük ölçüde filolojik ve karşılaştırmalı-yazmasal kanıta dayanır (2024; 2022). Bu makalenin altı çalışması, o kanıtlarla rekabet eden değil, ona tamamlayıcı ve dolaylı bir hesaplamalı sınaama sunmayı denemiştir. Sonuçların çoğunluğunun “desteklenmiyor” ya da “sınırlı kanıt” biçiminde çıkması, tam da bu tamamlayıcılığın sınırlarını gösterir. Modern notanın ölçü çizgisi tarihsel usul devrinin vuruş ağırlığı hiyerarşisini taşımaz (Çalışma 3). Son yazılı perde uzman bir kalış etiketiyle özdeş değildir (Çalışma 1, 2). Makam-usul-form paylaşımı nazire için ne gerekli ne yeterlidir (Çalışma 4). Dört dönem sekiz dönemin kaba bir indirgemesi olarak bile doğal küme yapısı üretmez (Çalışma 5). Bu, hesaplamalı yöntemin Feldman’ın kuramına uygulanamaz olduğu anlamına gelmez; tersine, hangi ölçümlerin bu aktarım yüzeyinde iz taşımadığını, hangisinin ters yönde iz taşıdığını (karar-yakınsama örneğinde olduğu gibi, Çalışma 1 ve 2’de bağımsız biçimde) ve hangi kuramsal bileşenin hangi yapısal önkoşulun eksikliğinde sınaanamaz kaldığını (nazire için sıfır kontrol çifti,

Çalışma 4; Birinci Selâm için sıfır gerçek nota, Çalışma 6) açıkça haritalamıştır. Bu haritalama, alanın bir sonraki adımlarını tarif eden somut bir kazanımdır.

Gelecek çalışma için birkaç yön, altı çalışmanın sınırlılıklar bölümlerinde ortaklaşa işaret edilmektedir. İlk olarak, karar-yakınsama ve cümle yayılımı gibi göstergelerin teslim bölgelerine duyarlı biçimde yeniden tanımlanması gerekmektedir; mevcut ölçütler son notayı mekanik biçimde karar sayar, oysa teslim bazı formlarda asma bir kalışla kapanabilir. İkincisi, Çalışma 5'in işaret ettiği özellik-ailesi ablasyonu, negatif kontrol modelleri ve alternatif kümeleme yöntemleriyle model-bağımlılık denetimi, bu makalenin bulduğu zayıf-orta düzey sinyallerin yöntem seçimine ne ölçüde bağlı olduğunu netleştirecektir. Üçüncüsü, Çalışma 4'ün önerdiği gibi, uzman kalış çözümlemesi ve kör uzman değerlendirmesi, hesaplamalı sıralamaların tarihsel-müzikolojik yargıyla nasıl tamamlanabileceğini gösterecektir; benzer biçimde Çalışma 1 ve 2'nin ortak açığı olan uzman doğrulama pilotu, karar-yakınsama gibi ölçütlerin geçerliliğini bağımsız biçimde sınayacaktır. Dördüncüsü ve belki en somutu, Çalışma 6'nın tanımladığı dört veri-hazırlık kapısının kaynak-izlenebilir üç Mevlevi ayini notası, uzman cümle sınırları, bağımsız makam-bölgesi etiketleri, çift kaynak kontrolü karşılanması, Feldman'ın (2022) Birinci Selâm tezini Aşama 2'ye taşıyacak somut bir gündemdir.

Bu gündemin sürdürülmesinde iki ilke belirleyici olmalıdır: hipotezleri ve karar kurallarını veri görülmeden sabitleyen kayıtlı rapor mantığı (Nosek vd., 2018) ve yazılımın yeniden üretilebilirliği ile tarihsel çıkarımın geçerliliğini birbirinden ayıran bir şeffaflık ilkesi (Munafò vd., 2017). Bu iki ilkeye bağlı kalındığında, alan ileride benzer iddiaları “testler geçti” ile “kuram destekleniyor” arasındaki farkı bulanıklaştırmadan değerlendirebilir. Eser kimliği, kaynak sayfası, dönüşüm günlüğü ve türetilmiş ölçümlerin birlikte sürümlenmesini öneren FAIR ilkeleri (Wilkinson vd., 2016) de bu tür bir kaynak-aktarım altyapısının kurulmasında yol gösterici bir çerçeve sunmaktadır.

Sonuç olarak, bu altı çalışma Feldman'ın Osmanlı saray çalgı musikisi kuramlarını çürütmemiş, ama onları da doğrudan doğrulamamıştır. Bulgular, modern sembolik aktarım örnekleminin meşkten notaya, notadan edisyona, edisyondan dijital külliyata uzanan bir zincirin son halkasının seçilen vekillerle taşıdığı sınırlı, kimi zaman beklenenin tersine işaret eden bir izdüşümünü yansıtmaktadır. Bu, Osmanlı saray musikisi tarihyazımı için olumsuz bir sonuç değildir. Tersine, tarihsel etnomüzikolojinin büyük anlatılarıyla hesaplamalı etnomüzikolojinin ölçüm araçlarını dürüstçe yüzleştiren, nerede sessiz kaldığını saklamayan bir sınama zincirinin kayıdır. Feldman'ın mirası bu tabloda, kendi filolojik yönteminin hâlâ vazgeçilmez olduğunu ve hesaplamalı sınamanın ancak bu yöntemle birlikte, onun yerine değil yanında ilerleyebileceğini gösteren bir biçimde çıkmaktadır.

Sınırlılıklar

Altı çalışmanın sınırlılıkları, tek tek tekrarlanmak yerine, burada üç kaynaktan doğan ortak bir yapı olarak konsolide edilmelidir.

Aktarım ve temsil sınırları. Bütün çalışmalarda veri, özgün tarihsel yazmalar değil, meşkten notaya, notadan edisyona, edisyondan SymbTr'nin dijital temsiline uzanan bir zincirin son halkasıdır; bu zincirin her aşaması bir seçimdir ve bulgular yalnızca bu zincirin ürünü hakkında konuşabilir. Bu genel sınıra ek olarak, MakamKit'in perde ölçümlerinin dayandığı 53 komalı AEU dizgesi yirminci yüzyılda kuramsallaştırılmış bir çerçevedir; bu çerçevenin erken dönem eserlerine geriye dönük uygulanması altı çalışmanın tamamında ayrı, kendine özgü bir anakronizm riski olarak yinelenmiştir (Çalışma 1, 2, 3, 4, 5, 6). Besteci atfı da eserin bestelenme tarihiyle özdeş değildir: dönem katmanları besteci atıflarına dayanır, anonim eserler elenmiştir, meşk silsilesi üzerinden sonradan yapılan atıf pratiği denetlenememiştir; bu belirsizlik dönem eksenini bulanıklaştırabilir (Çalışma 1, 3, 4).

Örnekleme büyüklüğü ve dengesizliği. Erken dönem hücreleri sistematik biçimde küçüktür — Çalışma 1'de 1650 öncesi ve 1650-1750'de 5-11 eser, Çalışma 2'de erken dönem $n=14$ ve $n=18$, Çalışma 3'te 14, 18 ve 21 eser, Çalışma 5'te 14 ve 18 eser — ve güç analizleri bu hücrelerde ancak orta-büyük etkilerin (Çalışma 2'de $|p|\approx 0,69$ ve $|p|\approx 0,62$; Çalışma 3'te kabaca $|p|\gtrsim 0,23$) güvenle yakalanabileceğini göstermektedir. Bu nedenle “anamlı değil” biçimindeki sonuçlar “etki yok” olarak okunmamalıdır; küçük-orta ölçekli bir dönüşüm istatistiksel güç yetersizliğinden kaçırılmış olabilir. Buna paralel olarak örneklemin dönem ve makam bileşimi de dengesizdir: Çalışma 2'de eserlerin üçte ikisi 1850

sonrasına aittir; Çalışma 4'te havuzun %88,2'si Aksak Semai usullü saz semâsidir; bu dengesizlikler havuzlanmış testlerde bir bileşim sürüklemesi (kompozisyon karıştırıcısı) riski taşır.

Ölçüm vekillerinin geçerliliği. Hiçbir çalışmadaki gösterge, Feldman'ın kastettiği kuramsal kavramı doğrudan ölçmez; hepsi notadan çıkarılabilen dolaylı vekillerdir. Karar-yakınsama ölçütü (Çalışma 1, 2) teslim bölgelerini ayırt etmez ve son notayı mekanik biçimde karar sayar; bu, en dayanıklı bulgunun bile teslim-duyarlı bir yeniden ölçüm gerektirdiği anlamına gelir. Cümle yayılımı (Çalışma 1, 3) usul-normalizasyonsuzdur ve icradaki nefes ile vurguyla değil notadaki bölüm işaretleriyle sınırlıdır. Çalışma 3'te düm-tek vuruş şiddetleri kaynak veride kodlanmadığından bir olay-fazı vekili kullanılmış, usulün ağırlık hiyerarşisi hiçbir göstergeyle doğrudan ölçülememiştir; saz semâilerinin tamamının Aksak Semai usulünde olması form ile usul etkisinin ayrıştırılmasını da engellemiştir. Çalışma 4'te aynı makam-usul-form paylaşımı nazire için ne gerekli ne yeterli bir koşuldur; hücre filtresi farklı usul ya da forma dönüştürülmüş gerçek nazireleri baştan dışlar, benzerlik ölçümleri ortak notacı etkisini besteci üslubundan ayıramaz ve 48 eserden türeyen 34 çift bağımsız gözlemler değildir. Çalışma 5'te ise özellik ailesi ablasyonu, negatif kontrol modeli ve alternatif kümeleme yöntemleriyle model-bağımlılık denetimi henüz yapılmamıştır; besteci-bloklu permütasyon testi kümeleme etiketlerini ve sınıflandırma tahminlerini sabit tuttuğundan, raporlanan p-değerleri modelin her yinelemede yeniden eğitildiği daha katı bir teste kıyasla iyimser çıkmış olabilir.

Doğrulama eksikliği. Hiçbir çalışmada, ölçüm vekillerinin uzman yargısıyla ne ölçüde örtüştüğünü sınavan önceden planlanmış bir pilot çalışma koşulmamıştır; bu, altı çalışmanın ortak ve en temel metodolojik açığıdır. Çalışma 1'in karma modeli yazılım katmanındaki bir hata nedeniyle yakınsamamış ve bu nedenle üçlü örüntüye (havuzlanmış, form-ayrı, eşleşmiş) ne katkı ne itiraz sunabilmiştir. Çalışma 4'te tarihsel nazire etiketi bulunmadığından duyarlılık ve kesinlik hiç hesaplanamamış, üretilen sıralama bir doğrulama metriği değil yalnızca bir inceleme önceliği olarak kalmıştır. Çalışma 6'nın sınırlılığı ise niteliksel olarak farklıdır: burada eksik olan bir istatistiksel güç ya da vekil geçerliliği değil, analiz için gerekli ham verinin kendisidir — kaynak-izlenebilir gerçek nota, uzman cümle sınırları, bağımsız makam-bölgesi etiketleri ve çift kaynak kontrolü, dördü de henüz karşılanmamış dört ön koşuldur.

Kaynakça

- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- 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>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Efron, B., & Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. Chapman & Hall/CRC. <https://doi.org/10.1201/9780429246593>
- Feldman, W. (2022). *From Rumi to the Whirling Dervishes: Music, Poetry, and Mysticism in the Ottoman Empire*. Edinburgh University Press. <https://doi.org/10.3366/edinburgh/9781474491853.001.0001>
- Feldman, W. (2024). *Music of the Ottoman Court: Makam, Composition and the Early Ottoman Instrumental Repertoire*. Brill. <https://doi.org/10.1163/9789004531260>
- Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of Classification*, 2(1), 193–218. <https://doi.org/10.1007/BF01908075>
- 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*, 223–228. <https://dblp.org/rec/conf/ismir/Karaosmanoglu12>
- Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 38(4), 963–974. <https://doi.org/10.2307/2529876>
- Munafo, M. R. ve diğerleri. (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. *PNAS*, 115(11), 2600–2606. <https://doi.org/10.1073/pnas.1708274114>
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1), 43–49. <https://doi.org/10.1109/TASSP.1978.1163055>

- Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72–101. <https://doi.org/10.2307/1412159>
- W3C Music Notation Community Group. (2021). *MusicXML 4.0*. <https://www.w3.org/2021/06/musicxml40/>
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), 80–83. <https://doi.org/10.2307/3001968>
- Wilkinson, M. D. ve diğerleri. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>



Research Article

Testing Feldman's theories of Ottoman-Turkish music from a computational ethnomusicological perspective

Seyhan Canyakan⁴

Afyon Kocatepe University State Conservatory, Afyonkarahisar, Türkiye.

Article Info

Received: 26 July 2026

Accepted: 31 August 2026

Published: 30 October 2026

Keywords

Computational ethnomusicology
Makam
Ottoman-Turkish music
Seyir
Symbolic music analysis
SymbTr
Usul
Walter Feldman

Abstract

This article aims to test six distinct theories proposed by Feldman concerning Ottoman-Turkish music using the MakamKit computational musicology analysis engine on a sample consisting of 149 peşrevs and saz semâîs drawn from the open-access SymbTr symbolic score corpus (Karaosmanoğlu, 2012), together with three Mevlevî âyins. Each study derived a priori directional hypotheses from Feldman's writings and tested them under multiple-comparison correction using methods including Spearman rank correlation, composer-clustered bootstrap, mixed-effects models, blind clustering, composer-grouped cross-validation, and similarity ranking. Across the six studies, the findings exhibit a common pattern. No test produced consistent evidence in the direction predicted by Feldman. Instead, the most consistent and recurrent deviation was the cadential-convergence measure, which moved in the direction opposite to that predicted in two independent tests of the transition from terkib to seyir (Studies 1 and 2). Apart from this reverse signal, the remaining analyses either produced no consistent pattern or could not be tested at all. The period-specific change in the articulation between usul and melody (Study 3) left no detectable trace on the surface of modern notation; the model–nazire relationship (Study 4) produced not a validated discovery but only a priority ranking for expert examination; the four-period periodization (Study 5) could not be recovered as a natural cluster structure, and only limited statistical evidence was found for a coarse early/late distinction; and the form–style relationship in the First Selâm (Study 6) could not be tested because of the absence of source-traceable scores and was therefore presented as a Stage 1 registered protocol. In conclusion, these six studies do not falsify Feldman's ethnomusicological historical thesis. Rather, the findings should be interpreted within the limitations of the modern symbolic transmission sample and the selected measurement proxies. The principal contribution of this research is to place theories of Ottoman-Turkish music within a testable and reproducible computational ethnomusicological framework while honestly documenting the points at which that framework remains silent.

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.



To cite this article

Canyakan, S. (2026). Testing Feldman's theories of Ottoman-Turkish music from a computational ethnomusicological perspective. *Türk Müziği*, 6(Walter Feldman Special Issue), 71-94. DOI: <https://doi.org/10.5281/zenodo.1>

Introduction

Walter Feldman's *Music of the Ottoman Court: Makam, Composition and the Early Ottoman Instrumental Repertoire* (2024) offers one of the most comprehensive and detailed accounts of the formal history of Ottoman-Turkish music. According to Feldman, the early instrumental repertoire, particularly forms such as the *peşrev* and *saz semâîsi*, consists of short, repetitive, and relatively self-contained melodic units, namely *terkibs*. These units can be reordered within a composition and, when appropriate, transferred to other works. From the second half of the seventeenth century onward, this modular organization gradually gives way, in Feldman's account, to a continuous musical structure integrated through the *seyir*, the characteristic melodic trajectory of the *makam*. Feldman never presents this

⁴ Assoc.Prof.Dr., Afyon Kocatepe University State Conservatory, Afyonkarahisar, Türkiye. Email: scanyakan@aku.edu.tr ORCID: 0000-0001-6373-4245

transformation as a single-dimensional process. Rather, it is constructed as a multi-layered historical development involving the ordering of tonal centres, the articulation between *usul* cycles and melody, the *hâne-mülâzime* (*teslim*) structure, and the *nazire* tradition, namely the practice of composing new works in the style of an existing composition. In *From Rumi to the Whirling Dervishes: Music, Poetry, and Mysticism in the Ottoman Empire* (2022), Feldman extends this historical-formal interpretation to the religious ceremonial repertoire, particularly the First *Selâm* of the Mevlevi *ayin*, where he likewise proposes a directional relationship between formal length and the historical position of the composer.

Feldman's scholarship stands at the intersection of historical musicology and ethnomusicology. His methodology is predominantly philological and comparative, relying on close readings of seventeenth- and eighteenth-century *edvâr* treatises, lyric anthologies, and the limited number of surviving early notated sources. In a musical tradition transmitted primarily through *meşk*, the master-apprentice system of oral transmission, rather than written notation, this approach represents one of the few means of tracing formal change. In Feldman's work, the performance-oriented perspective of ethnomusicology and the document-based perspective of historical musicology complement rather than exclude one another. The scarcity of written sources necessitates close attention to the internal logic of performance practices such as *meşk*, *taksim*, and *gazel*, while the surviving early notated materials prevent the interpretation from becoming entirely speculative. At the same time, this richness creates a methodological challenge. Feldman's historical propositions rely largely on textual evidence and selected analytical examples. Whether these propositions can be systematically, transparently, and reproducibly examined across a large musical corpus remains largely unresolved in the literature. This question cannot be answered through close reading of a single work but requires the comparative and measurable analysis of many compositions. It is precisely at this point that computational approaches enter the discussion, not as a replacement for philological interpretation but as a complementary analytical tool.

Over the past decade, computational musicology has developed an expanding methodological repertoire for the study of Turkish *makam* music. The review by Bozkurt, Ayangil, and Holzapfel (2014) demonstrates that this body of research has focused primarily on *makam* recognition, pitch and microtonal analysis, and surface-level similarity measurement, whereas methods for systematically testing claims of historical transformation remain comparatively limited. The same study also emphasizes that information extracted from symbolic notation represents only a portion of the knowledge embodied in performance practice. This observation constitutes a common methodological reference point throughout the six studies presented here. Positioned within this gap, the present article seeks to translate a broad historical ethnomusicological narrative into a transparent sequence of computational ethnomusicological tests using reproducible analytical methods.

The data source for this investigation is SymbTr, the open-access symbolic music corpus for Turkish *makam* music introduced by Karaosmanoğlu (2012). SymbTr contains a large collection of works from the traditional repertoire in machine-readable form, together with information on pitch, duration, *usul*, and formal sections, and has served as the foundation for a substantial body of research in music information retrieval and computational analysis of *makam* music. Its license permits non-commercial research use with mandatory attribution. The corpus employed in this article is limited to *peşrevs* and *saz semâîs* selected from a SymbTr-derived collection and does not constitute a probabilistic historical sample. Rather, it represents a curated transmission sample that has passed through successive stages of survival, preservation within the *meşk* tradition, transcription into notation, editorial selection, digitization, and composer attribution. It is therefore not the original historical manuscripts themselves but the final stage of this transmission chain. This fundamental limitation is stated explicitly here. Each of the six studies further discusses additional limitations specific to its own sample. For Studies 1-5, the common dataset comprises 149 *peşrevs* and *saz semâîs*, grouped into four broad historical strata derived from composer attributions (pre-1650, 1650-1750, 1750-1850, and post-1850). These four strata intentionally simplify the finer periodization, extending to eight sub-periods, proposed by Feldman. No page-level correspondence is established between these strata and Feldman's historical transitions; consequently, every analysis based on them should be interpreted only as an approximation of his chronology. The strata represent approximate historical attribution inferred from composers' lifespans rather than

exact composition dates, and the corpus is weighted toward the late period (post-1850). This imbalance constitutes a shared methodological limitation across the six studies. Study 6 departs from this corpus entirely. Because the Mevlevi *ayin* repertoire is not included in SymbTr, a separate procedure for obtaining traceable notated sources was defined. The computational infrastructure employed throughout the study is MakamKit, a reusable, *makam*-aware Python analysis engine developed for Turkish folk and art music. MakamKit consists of five modules: symbolic melodic analysis, audio-based pitch analysis (from 53-comma pitch histograms to tonic and *makam* estimation), *usul/aksak* analysis, lyric-sentiment analysis, and similarity, clustering, and network analysis. All six studies were conducted through this engine to ensure a common computational vocabulary and a reproducible analytical pipeline. Only the symbolic module was used in the present article. This module extracts pitch, duration, metrical position, and sectional boundaries from symbolic notation files and derives quantitative indicators of *makam-usul* relationships from these data. The audio-based pitch/*makam* module, the *usul/aksak* module, and the lyric-sentiment module were not employed. This reflects a limitation of the available data, as SymbTr does not provide audio recordings, rather than a methodological preference. Within this study, MakamKit functions as a measurement tool rather than an interpretative one. It derives predefined proxy variables from symbolic notation, while the extent to which these proxies correspond to Feldman's theoretical concepts, including *seyir*, *terkib*, and *nazire*, is discussed explicitly in each study.

Research Problem

Within this framework, the research problem of the present article is to examine six distinct, limited, and falsifiable propositions derived from Feldman's broader account of historical transformation. The first and second propositions address the transition from modular *terkib* organization to *seyir*-based continuity using two different sets of proxies: motif repetition, phrase expansion, and the decision-convergence axis (Study 1), and a five-dimensional set of pre-specified convergent *seyir* indicators (Study 2). The third proposition examines whether the articulation between *usul* cycles and melody changes across historical periods, measured through indicators such as barline crossing frequency and beat-phase entropy. The fourth proposition investigates whether the model-*nazire* relationship, namely the practice of composing a work in imitation of another, can be prioritized through computational similarity analysis as a network of inter-composer influence. The fifth proposition tests whether Feldman's four-period formal periodization can be recovered independently from measurable musical features through unsupervised clustering, without providing period labels to the model. The sixth proposition concerns the relationship between formal length and the historical position of the composer in the First *Selâm* of the Mevlevi *ayin*. As explained below, this proposition is presented here not as an empirical result but as a testing protocol. Although these six propositions form an integrated historical narrative in Feldman's work, each requires a distinct measurement strategy, dataset, and analytical design. Accordingly, they are treated as six separate studies.

The remainder of the article is organized as follows. Studies 1 and 2 examine the transition from *terkib* to *seyir* using two different sets of proxies, namely the motif repetition/decision-convergence axis and the five-dimensional convergent *seyir* indicator framework. Study 3 investigates historical changes in *usul*-melody articulation using five quantitative indicators. Study 4 frames the model-*nazire* relationship not as a hypothesis test but as an exploratory candidate-prioritization exercise. Study 5 evaluates the proposed four-period historical classification through unsupervised clustering and composer-grouped classification. Study 6 differs methodologically from the preceding studies. Rather than relying on the SymbTr corpus, it establishes a confirmatory protocol requiring traceable symbolic scores of Mevlevi *ayins* and explicitly states that the study remains at the data-preparation stage, producing no historical conclusions. Each study follows the same organizational template, presenting Feldman's theoretical argument, the dataset and methods used for testing it, and the corresponding findings with its own result tables. To avoid repetition, the overall evaluation of limitations and the discussion of findings in relation to Feldman's interpretation and the broader literature are consolidated into a single *Discussion and Conclusion* section addressing all six studies collectively. The article concludes with an integrated assessment of what this sequence of investigations reveals about Feldman's theoretical legacy and a unified reference list containing the sources shared across all studies.

Study 1. Does Feldman's Thesis of the Transition from Modular Organization to *Seyir*-Based Continuity Receive Support after Matching and Mixed-Effects Modelling?

Theory 1: From *Terkib* to *Seyir*: The Dissolution of Modular Organization

In *Music of the Ottoman Court*, Walter Feldman presents a comprehensive account of the formal evolution of the Ottoman court instrumental repertoire (Feldman, 2024). According to this account, early instrumental forms, particularly the *peşrev* and *saz semâisi*, are constructed from short, repetitive, and relatively autonomous melodic blocks known as *terkibs*. A *terkib* is a ready-made melodic unit that can be rearranged within a composition and transferred to other works; the compositional economy of the early repertoire relies both on literal repetition and on restarting the same intervallic pattern from a different pitch level (sequence). Thus, *terkib* is not a single fixed pattern but a morphological repertoire that varies according to form and context. This characteristic is not incidental. As the *peşrev* serves as the instrumental opening of the *fasıl*, the ordered performance suite, the portability of *terkibs* is consistent with this framing function. According to Feldman, from the second half of the seventeenth century onward this modular organization gradually weakens, as melodic blocks become integrated into the *seyir*, the characteristic melodic trajectory of the *makam*, producing a longer-range musical continuity that unfolds as a single extended arc. *Seyir* defines the ascending, descending, or ascending-descending character of a *makam*, its intermediate resting points, and ultimately the directed motion toward the final cadence on the principal pitch.

Feldman does not reduce this transformation to a single dimension. Instead, he describes it as a multidimensional historical process involving the ordering of tonal centres, the articulation between *usul* and melody, the *hâne-mülâzime* (*teslim*) structure, in this study the terms *mülâzime* and *teslim* are used synonymously to denote the common returning section following each *hâne*, and the *nazire* tradition. The present study isolates only the *terkib-seyir* dimension of this broader narrative, namely the proposed transition from a modular economy of repetition to *seyir*-based continuity. The *nazire* network, *usul*-melody articulation, and the structural organization of sectional boundaries are excluded because they require different measurements and different datasets. Bozkurt, Ayangil, and Holzapfel (2014) provide a comprehensive overview of the measures and unresolved issues in the computational analysis of *makam* music, while Karaosmanoğlu (2012) introduced SymbTr, the open-access symbolic music corpus that enables large-scale computational studies. The present study brings together historical ethnomusicology and computational ethnomusicology through a single research question:

- Does Feldman's thesis of the transition from modular organization to *seyir*-based continuity leave observable formal traces in the modern symbolic transmission?

Testing the Theory

The sample consists of 149 instrumental works selected from a SymbTr-derived open symbolic corpus (Karaosmanoğlu, 2012), including 82 *peşrevs* and 67 *saz semâîs*. These works are not original historical manuscripts but modern symbolic transcriptions. The analyzed data therefore represent the final stage of a transmission chain extending from *meşk* to notation, from notation to critical editions, and from editions to digital representation. During data import, structurally corrupted rows, zero-duration events, and *usul* metadata rows containing no musical notes were removed. Records containing fewer than twenty valid notes, lacking composer attribution, or not assignable to one of the four historical strata were excluded from the analysis. The historical strata, pre-1650, 1650-1750, 1750-1850, and post-1850, were derived from composer attributions rather than documented composition dates. The first boundary (1650) follows Feldman's own chronology, in which the transition from modular organization is located in the second half of the seventeenth century. The later boundaries (1750 and 1850) were introduced to distinguish the mature and late layers most strongly represented in the transmission and have not been demonstrated to correspond directly to Feldman's more detailed historical subdivisions. Consequently, the adopted periodization should be regarded as an approximation.

The 149 compositions represent the work of 77 composers, making the effective independent unit of observation the composer rather than the individual composition. The dataset is not a probabilistic historical sample but a selected transmission sample that has survived successive processes of preservation, notation, editorial selection, digitization,

and composer attribution. The early historical strata (pre-1650 and 1650-1750) contain only 5-11 works, while the 1750-1850 strata contain between 7 and 14 works. These limited sample sizes make any evidence of absence particularly fragile.

The analysis was conducted using MakamKit, the common computational framework developed for this research programme. MakamKit reconstructs the pitch and duration sequences of each composition from symbolic notation files, converts pitches into relative values referenced to the final pitch of the work, and derives *makam*-related quantitative features. Five proxy measures were defined for this study: (1) literal repetition, the proportion of exact repetitions of absolute pitch sequences within three- and four-note motif windows; (2) transposed repetition, the proportion of repeated intervallic patterns beginning on different pitch levels; (3) phrase span, the length of sectionally defined musical units measured in bars; (4) trajectory roughness, computed by dividing each composition into twelve equal note-count segments and calculating the mean absolute difference between successive relative pitch centres; and (5) cadential convergence, calculated by measuring the mean absolute distance (d) from the final pitch within the last 20% of the note sequence and transforming it using the function $1/(1 + d)$, yielding values between 0 and 1. Values approaching 1 indicate that the final section of a composition is more tightly concentrated around the final pitch, reflecting stronger cadential convergence. The measure therefore represents not *a priori* but the inverse transformation of that distance into a convergence score. Five hypotheses were specified *a priori*: literal repetition and transposed repetition were expected to decrease across historical periods, phrase span and cadential convergence to increase, and trajectory roughness to decrease. The predefined composite support criterion required that at least three of the five indicators move in the predicted direction, with at least two of these exhibiting composer-clustered bootstrap 95% confidence intervals excluding zero.

The primary trend statistic was Spearman's rank correlation (Spearman, 1904), and robustness was assessed through composer-clustered bootstrap resampling (2,000 iterations with a fixed random seed; Efron & Tibshirani, 1994). To control for potential confounding by formal type, two additional analyses were performed. First, early and late compositions were matched by nearest duration within identical *makam* $\times$ *usul* $\times$ *form* strata and compared using the Wilcoxon signed-rank test (Wilcoxon, 1945); only twelve matched pairs could be constructed. Second, a mixed-effects model incorporating random composer intercepts and variance components for *makam* and *usul* was fitted (Laird & Ware, 1982). Multiple comparisons were corrected using the Benjamini-Hochberg procedure with a 5% false discovery rate (Benjamini & Hochberg, 1995). In the form-specific trend analyses, *peşrev*s and *saz semâî*s were treated as separate correction families of five tests each.

Results

Table 1. Sample sizes and median values of the five measures by historical period and form

Period	Form	<i>n</i>	Literal repetition (median)	Transposed repetition (median)	Phrase span, bars (median)	Trajectory roughness (median)	Cadential convergence (median)
Pre-1650	<i>Peşrev</i>	9	0.497	0.069	2.243	10.842	0.043
Pre-1650	<i>Saz semâîsi</i>	5	0.730	0.038	2.188	12.856	0.049
1650-1750	<i>Peşrev</i>	11	0.553	0.036	1.632	10.135	0.062
1650-1750	<i>Saz semâîsi</i>	7	0.697	0.034	2.241	9.410	0.051
1750-1850	<i>Peşrev</i>	14	0.527	0.080	0.614	10.817	0.046
1750-1850	<i>Saz semâîsi</i>	7	0.506	0.052	2.300	11.709	0.042
Post-1850	<i>Peşrev</i>	48	0.483	0.085	0.883	12.953	0.039
Post-1850	<i>Saz semâîsi</i>	48	0.626	0.043	3.215	9.583	0.036

The sample is heavily weighted toward the late period, with 96 of the 149 compositions belonging to the post-1850 stratum. The small size of the early strata (5-11 works in the pre-1650 and 1650-1750 groups, and 7-14 works in the 1750-1850 group) indicates that none of the differences reported in this table should be interpreted, by themselves, as evidence of historical trends.

Table 2. Pooled historical trends and composer-clustered bootstrap results ($n = 149$)

Measure	Spearman's ρ	<i>p</i>	Clustered slope	95% CI	Bootstrap <i>p</i>	BH <i>q</i>
Literal repetition	-0.089	0.278	-0.009	[-0.030, 0.017]	0.458	0.630
Transposed repetition	0.058	0.483	-0.002	[-0.009, 0.008]	0.660	0.630
Phrase span	0.055	0.504	0.888	[-0.926, 2.669]	0.302	0.630

Trajectory roughness	0.037	0.656	0.336	[-0.446, 1.229]	0.386	0.656
Cadential convergence	-0.325	<0.001	-0.005	[-0.010, -0.002]	<0.0005	<0.001

In the pooled sample, the confidence intervals of four measures include zero. The only robust historical trend is observed for cadential convergence; however, its direction is opposite to that predicted by Hypothesis 5. Rather than increasing across historical periods, the final sections of compositions exhibit progressively weaker convergence toward the final pitch.

Table 3. Form-specific historical trends (*peşrev* $n = 82$, *saz semâîsi* $n = 67$; each form treated as a separate five-test correction family)

Measure	Form	Spearman's ρ	p	95% CI	BH q
Literal repetition	<i>Peşrev</i>	-0.150	0.179	[-0.043, 0.010]	0.224
Transposed repetition	<i>Peşrev</i>	0.195	0.079	[-0.010, 0.016]	0.196
Phrase span	<i>Peşrev</i>	-0.108	0.335	[-3.514, 0.676]	0.335
Trajectory roughness	<i>Peşrev</i>	0.174	0.117	[-0.198, 1.849]	0.196
Cadential convergence	<i>Peşrev</i>	-0.269	0.015	[-0.010, -0.001]	0.073
Literal repetition	<i>Saz semâîsi</i>	-0.145	0.243	[-0.039, 0.017]	0.404
Transposed repetition	<i>Saz semâîsi</i>	-0.039	0.757	[-0.009, 0.007]	0.757
Phrase span	<i>Saz semâîsi</i>	0.292	0.017	[1.210, 6.427]	0.041
Trajectory roughness	<i>Saz semâîsi</i>	-0.061	0.622	[-1.529, 1.824]	0.757
Cadential convergence	<i>Saz semâîsi</i>	-0.419	<0.001	[-0.011, -0.002]	0.002

Within the *peşrev* repertoire, none of the observed trends remained significant after controlling the false discovery rate at 5%. In the *saz semâîsi* repertoire, two patterns emerged. The decline in cadential convergence remained robust after correction ($q = 0.002$), while the increase in phrase span also exceeded the 5% threshold within its own correction family ($q = 0.041$). This represents the only partial and form-specific finding consistent with Feldman's thesis.

Table 4. Early-late differences in 12 matched pairs (nearest-duration early-late pairs within identical *makam* $\times$ *usul* $\times$ *form* strata)

Measure	Early mean	Late mean	Mean difference	Wilcoxon W	p	BH q
Literal repetition	0.654	0.635	-0.019	39.0	1.000	1.000
Transposed repetition	0.048	0.045	-0.003	38.0	0.970	1.000
Phrase span	2.156	2.975	0.819	22.0	0.204	0.339
Trajectory roughness	9.844	8.483	-1.361	22.0	0.204	0.339
Cadential convergence	0.055	0.041	-0.014	7.0	0.009	0.046

Under the strictest control, matching compositions with identical *makam*, *usul*, and formal type, cadential convergence again remained the only statistically significant difference after correction ($q = 0.046$). For the repetition measures, the directions of the mean and median differences were inconsistent, indicating sensitivity to extreme values within the small sample of twelve matched pairs.

Table 5. Convergence status of the mixed-effects model with random composer intercepts ($n = 149$, 77 composers)

Measure	Model convergence	Period coefficient (β)
Literal repetition	No	—
Transposed repetition	No	—
Phrase span	No	—
Trajectory roughness	No	—
Cadential convergence	No	—

The mixed-effects model failed for all five measures because of the software dependency error "*No module named 'statsmodels.formula'*". Consequently, no period coefficients could be estimated. The mixed-effects model therefore cannot be regarded as an independent source of evidence for separating composer-level effects from *makam* and *usul* effects. Across the remaining three independent analyses of the *terkib-seyir* dimension, namely the pooled trend analysis, form-specific trend analysis, and matched-pair comparison, the only consistent pattern is the decline in cadential convergence across historical periods, contrary to theoretical expectation. This finding is accompanied by a limited increase in phrase span specific to the *saz semâîsi* repertoire.

In summary, among the five proxy measures, only cadential convergence remained consistent and statistically significant after correction across three independent analyses, namely the pooled trend analysis, form-specific analysis, and matched-pair comparison, but its direction was opposite to that predicted. Phrase span exhibited an additional, partial, and form-specific signal only within the *saz semâîsi* repertoire.

Study 2. Does Feldman's Thesis of the Historical Integration of *Seyir* Organization Receive Support When Tested with a Five-Dimensional Quantitative Indicator Framework?

Theory 2: The Historical Codification of *Seyir*

Feldman argues that, from the second half of the seventeenth century onward, the relatively independent *terkib* characteristic of the early repertoire became integrated into a *makam seyir* organization extending from the beginning to the end of a composition. This transformation is described as a multidimensional historical process in which *terkib*, *seyir*, *usul*-melody articulation, the *hâne-mülâzime* structure, and *nazire* relationships, namely the practice of composing one work in imitation of another, evolve together rather than representing a simple linear progression. The present study examines only one component of this broader historical thesis by testing whether the organization of the *seyir* trajectory can be detected across composer periods within the modern symbolic transmission.

From this historical proposition, a five-dimensional set of directional hypotheses was derived *a priori*. As *seyir* becomes increasingly integrated, it is expected that (H1) trajectory roughness, reflecting tonal jumps between successive melodic windows, will decrease; (H2) the number of directional changes will become less frequent; (H3) the final section of a composition will become more tightly organized around the final pitch (cadential convergence); (H4) melodic direction will more frequently align across the boundaries between *hâne* and *mülâzime*; and (H5) the pitch distributions on either side of these boundaries will become increasingly similar. Support for Feldman's thesis was defined not by any single indicator but by the convergence of all five indicators in the predicted direction. This criterion of multiple convergent validity was adopted deliberately to prevent an isolated statistically significant result from being interpreted as evidence that *seyir* became historically integrated. Within this framework, *seyir* is treated not as a static label such as ascending or descending but as a dynamic temporal trajectory unfolding throughout an entire composition, thereby rendering Feldman's historical proposition amenable to empirical testing.

Testing the Theory

The sample consists of 149 *peşrevs* and *saz semâîs*, the two principal forms of the Ottoman court instrumental repertoire, selected from the open-access SymbTr symbolic music corpus (Karaosmanoğlu, 2012), and attributed to 77 composers. The unit of analysis is the musical composition rather than the individual note. Records containing fewer than twenty valid notes, lacking composer attribution, or not assignable to one of the four historical strata were excluded. Historical period was defined not by the documented date of composition but by four attribution strata derived from composers' lifespans: pre-1650 (14 works), 1650-1750 (18 works), 1750-1850 (21 works), and post-1850 (96 works). These three temporal boundaries (1650, 1750, and 1850) represent an analytical simplification of Feldman's more detailed historical periodization (see Study 5). Because the original study does not explicitly map these boundaries to specific pages or passages in Feldman's text, the resulting historical strata should be interpreted as approximate analytical categories. For the two indicators based on *hâne-mülâzime-serhâne* boundaries (H4 and H5), the effective sample was reduced to 132 works, as 17 compositions contained no structural section markers. The selected-transmission character of the corpus has already been discussed in the Introduction; it should simply be recalled here that approximately two-thirds of the sample belong to the post-1850 attribution layer.

The analyses were performed using the MakamKit computational engine, an open-source Python library incorporating both *makam*-aware symbolic and audio-based modules. Only the symbolic module was employed in this study. Each composition was divided into twelve windows containing equal numbers of notes, and within each window the duration-weighted mean pitch was calculated in the 53-comma tuning system relative to the final notated pitch of the composition. Five indicators were derived from this common representation: trajectory roughness (the mean absolute difference between successive window centres, measured in commas); the number of directional changes (the number of sign changes in successive window-centre differences); the cadential convergence score, calculated as the inverse transformation $1/(1 + d)$, where d is the mean absolute distance of notes in the final fifth of the composition from the final written pitch, yielding values between 0 and 1, with higher values indicating stronger convergence toward the cadence; directional continuity across section boundaries (the proportion of agreement in melodic direction within the eight-note contexts immediately before and after each *hâne-mülâzime-serhâne*

boundary); and pitch-distribution similarity across section boundaries (the similarity between pitch usage profiles on either side of each boundary).

The statistical strategy consisted of three components. The primary analysis employed Spearman's rank correlation (Spearman, 1904) between the ordered historical strata and each indicator. This coefficient summarizes monotonic trends across historical periods without assuming linearity. The robustness of effect direction was evaluated using the composer-clustered bootstrap procedure introduced in Study 1 (Efron & Tibshirani, 1994; 2,000 iterations per indicator, fixed random seed 20260720, 95% confidence intervals from the 2.5th and 97.5th percentiles), thereby preventing prolific composers from artificially inflating statistical certainty. Because five simultaneous tests were conducted, multiple-comparison correction was applied. The Bonferroni-adjusted significance threshold was set at $\alpha = 0.01$ ($0.05/5$). In addition, Benjamini-Hochberg (1995) false-discovery-rate correction was calculated and is reported in Table 6 as the BH q value. Although the cadential convergence score was also examined in Studies 1 and 3, each study defines its own five-test correction family, and therefore q -values may differ across studies. Power analysis indicated that, owing to the small early-period samples, only very strong effects could be detected reliably. Using the standard Fisher z approximation, the minimum detectable effect size at 80% power and $\alpha = 0.05$ was approximately $|\rho| = 0.69$ for the pre-1650 group ($n = 14$) and $|\rho| = 0.62$ for the 1650-1750 group ($n = 18$). Consequently, moderate historical trends may remain undetected within the current design. Finally, a sensitivity analysis examined whether the principal indicators depended on the selected computational parameters by recalculating the analyses under nine combinations of three window numbers (8, 12, and 16) and three terminal-segment proportions (10%, 20%, and 30%).

Results

Table 6. Historical trends of the five *seyir* indicators ($n = 149$; section-boundary indicators $n = 132$).

Indicator	n	ρ	p	Slope (per period)	95% CI lower	95% CI upper	p_{boot}	BH q	Expected direction	Observed direction consistent?
Trajectory roughness	149	0.037	0.656	0.336	-0.418	1.319	0.371	0.820	Decrease (H1)	No
Number of directional changes	149	0.085	0.304	0.150	-0.027	0.409	0.100	0.507	Decrease (H2)	No
Cadential convergence score	149	-0.325	<0.001	-0.005	-0.010	-0.002	<0.0005	0.000259	Increase (H3)	No*
Directional continuity across section boundaries	132	-0.017	0.849	-0.019	-0.060	0.055	0.569	0.849	Increase (H4)	No
Pitch-distribution similarity across section boundaries	132	-0.127	0.146	-0.027	-0.078	0.010	0.144	0.365	Increase (H5)	No

Only the reversal observed for the cadential convergence score is statistically significant ($p < 0.001$; BH $q = 0.000259$). The remaining four indicators exhibit point estimates opposite to the predicted direction but do not reach statistical significance.

This table shows that four of the five indicators exhibit no stable trend across historical periods, as their confidence intervals include zero. The final two columns reveal a different pattern: all five hypotheses (H1-H5) display point estimates opposite in sign to those predicted by Feldman. This fivefold reversal reaches statistical significance only for the cadential convergence score ($\rho = -0.325$, $p < 0.001$), which remains significant after both Bonferroni correction ($\alpha = 0.01$) and Benjamini-Hochberg false-discovery-rate adjustment (BH $q = 0.000259$). The remaining four indicators remain non-significant after correction (BH $q = 0.365$ -0.849). Rather than reflecting an isolated contradictory result, the findings indicate that all indicators point in the same opposite direction, an issue examined further in the Discussion and Conclusion section.

Table 7. Median indicator values by historical period and form

Period	Form	n	Roughness	Direction changes	Cadential convergence	Directional continuity	Distribution similarity
Pre-1650	<i>Peşrev</i>	9	10.842	6.0	0.043	0.286	0.443
Pre-1650	<i>Saz semâisi</i>	5	12.856	6.0	0.049	0.429	0.370
1650-1750	<i>Peşrev</i>	11	10.135	5.0	0.062	0.310	0.374
1650-1750	<i>Saz semâisi</i>	7	9.410	6.0	0.051	0.429	0.434
1750-1850	<i>Peşrev</i>	14	10.817	6.0	0.046	0.050	0.495
1750-1850	<i>Saz semâisi</i>	7	11.709	6.0	0.042	0.292	0.373
Post-1850	<i>Peşrev</i>	48	12.953	6.0	0.039	0.268	0.445
Post-1850	<i>Saz semâisi</i>	48	9.583	6.0	0.036	0.357	0.326

Table 7 confirms that the pooled results are not driven by a single musical form. The median cadential convergence score reaches its highest value in the 1650-1750 stratum for both *peşrevs* and *saz semâis* before declining in the post-

1850 period. This decline is common to both forms, although the pattern reflects a mid-period peak followed by regression rather than a strictly linear historical decrease.

Table 8. Median values of the indicators for *makams* represented by at least four compositions

<i>Makam</i>	<i>n</i>	Trajectory roughness	Direction changes	Cadential convergence	Directional continuity	Distribution similarity
<i>Acemâşiran</i>	6	14.970	5.5	0.031	0.667	0.380
<i>Beyati</i>	5	6.841	7.0	0.048	0.375	0.577
<i>Bûselik</i>	4	9.220	5.0	0.047	0.482	0.382
<i>Ferahfeza</i>	5	20.643	6.0	0.021	0.227	0.415
<i>Hicaz Hümayun</i>	4	9.641	6.0	0.050	0.364	0.450
<i>Hicazkâr</i>	6	10.476	5.5	0.041	0.488	0.514
<i>Hüseyni</i>	7	8.402	6.0	0.047	0.264	0.450
<i>Hüzzam</i>	4	10.315	7.0	0.049	0.411	0.672
<i>Kürdilibicazkâr</i>	5	9.934	6.0	0.033	0.000	0.494
<i>Mabur</i>	7	13.675	6.0	0.031	0.400	0.370
<i>Muhayyer</i>	4	10.563	5.0	0.043	0.000	0.042
<i>Nevâ</i>	4	8.446	5.5	0.043	0.238	0.364
<i>Nihavent</i>	6	13.006	6.5	0.037	0.375	0.326
<i>Rast</i>	4	14.606	6.0	0.037	0.000	0.521
<i>Sabâ</i>	6	9.889	6.5	0.050	0.429	0.265
<i>Segâh</i>	5	9.440	5.0	0.064	0.294	0.545
<i>Uşşak</i>	5	7.242	6.0	0.065	0.191	0.284

Table 8 provides an internal perspective on whether the pooled historical trend reflects historical period or differences in *makam* composition within a corpus whose *makam* distribution is not constant across periods. The median trajectory roughness ranges from 6.841 (*Beyati*) to 20.643 (*Ferahfeza*), a substantially wider interval than the differences observed between the historical period strata in Table 9. This suggests that *makam* identity discriminates indicator values more strongly than historical period.

Table 9. Sensitivity analysis by window number and terminal-segment proportion (ρ and p ; $n = 149$)

Windows	Final segment	Trajectory roughness ρ (p)	Direction changes ρ (p)	Cadential convergence ρ (p)
8	10%	0.083 (0.314)	0.060 (0.466)	-0.438 (<0.001)
8	20%	0.083 (0.314)	0.060 (0.466)	-0.325 (<0.001)
8	30%	0.083 (0.314)	0.060 (0.466)	-0.318 (<0.001)
12	10%	0.037 (0.656)	0.085 (0.304)	-0.438 (<0.001)
12	20%	0.037 (0.656)	0.085 (0.304)	-0.325 (<0.001)
12	30%	0.037 (0.656)	0.085 (0.304)	-0.318 (<0.001)
16	10%	0.063 (0.444)	0.025 (0.765)	-0.438 (<0.001)
16	20%	0.063 (0.444)	0.025 (0.765)	-0.325 (<0.001)
16	30%	0.063 (0.444)	0.025 (0.765)	-0.318 (<0.001)

The complete nine-row sensitivity table is not accompanied by a separate raw-data record in the results JSON file. Consequently, it has been reproduced directly from the source article and cannot be independently recalculated. The only verifiable row is the 12-window / 20% terminal-segment configuration, which, by definition, corresponds exactly to the principal results reported in Table 6 and therefore confirms internal consistency. Although the remaining eight rows display a pattern consistent with this validated reference point, their underlying calculations could not be independently verified. Accordingly, the interpretation that the inverse trend in cadential convergence remains stable across all nine parameter combinations should be treated with caution beyond the single validated configuration. By contrast, the absence of significant trends for trajectory roughness and directional changes is consistent across all nine parameter combinations, with p -values exceeding 0.05 throughout, and no additional robustness claim is made for these indicators.

In summary, four of the five indicators, namely trajectory roughness, the number of directional changes, directional continuity across section boundaries, and pitch-distribution similarity across section boundaries, exhibit no stable trend across historical periods. The only statistically significant indicator, the cadential convergence score, changes in the direction opposite to that predicted. At the level of point estimates, all five hypotheses display signs contrary to those proposed by Feldman.

Study 3. Does Feldman's Thesis of the Historical Transformation of the *Usul-Melody* Relationship Receive Support When Examined through Five Indicators in Modern Symbolic Notation?

Theory 3: Historical Change in the Articulation between *Usul* Cycles and Melody

According to Feldman (2024), melody in early *peşrevs* and *saz semâts* is closely bound to the *usul* cycle, whereas from the second half of the seventeenth century onward this relationship gradually loosens, allowing melody to extend beyond the boundaries of the *usul* cycle and develop into a longer-span musical continuity integrated within the *seyir* of the *makam*. The present study examines only this component of Feldman's broader thesis. Other dimensions, including the ordering of tonal centres, *hâne-mülâzime* continuity, *nazire* relationships, and the historical periodization of the repertoire, remain outside its scope.

In Feldman's account, *usul* is not merely a metrical counting scheme but a temporal organization shaped by the hierarchical weighting of *düm* and *tek* strokes and transmitted through the embodied tradition of *meşk*. Whereas melody in the early repertoire adheres closely to this temporal framework, later melody moves more freely across it. The research question is therefore formulated as follows:

- Does the manner in which melodic ideas extend beyond *usul* cycles and bar boundaries vary across the historical strata derived from composer attributions?

The difficulty of addressing this question quantitatively must be acknowledged from the outset. Historical *usul* existed as an embodied system of accentual weighting that was never directly recorded, whereas the available data preserve only the surface of modern notation through barlines. As noted by Bozkurt, Ayangil, and Holzapfel (2014), symbolic notation captures only part of the information embodied in performance practice. This limitation becomes particularly acute in relation to *usul*: the relative weighting of *düm* and *tek* strokes, precisely the feature that transforms *usul* from a counting pattern into a temporal organization, is not encoded in any symbolic notation format. The methodology of the present study is therefore constructed with explicit recognition of this limitation.

Testing the Theory

The sample was drawn from a derivative of the SymbTr collection (Karaosmanoğlu, 2012), an open corpus of modern symbolic music transcriptions. The two instrumental forms examined, *peşrev* and *saz semâtsi*, constitute the core of the repertoire and provide a natural setting for tracing the *usul-melody* relationship through their *hâne-mülâzime* architecture. Records containing fewer than twenty valid notes, lacking composer attribution, or not assignable to one of the four attribution strata were excluded. During import, structurally corrupted rows, zero-duration symbols, and *usul* metadata rows containing no notes were also separated from musical events. The resulting sample comprised 149 works: 14 attributed to the pre-1650 period, 18 to 1650–1750, 21 to 1750–1850, and 96 to the post-1850 period, representing 77 composers. All 67 *saz semâts* in the corpus are in the *Aksak Semai usul*, a form-*usul* dependency addressed separately in the Method section. The selected-transmission character of the sample has been discussed above.

The analysis was conducted using the MakamKit computational engine. MakamKit reads symbolic notation files derived from SymbTr, extracts pitch, duration, metrical position, and section-boundary information, and generates quantitative indicators of *makam* and *usul* relationships. Five indicators (G1–G5) were defined. G1, the cross-barline connection ratio, is the proportion of connections between successive sound events that cross a barline. G2, final-pitch–bar-end alignment, is the proportion of events on the principal pitch that coincide with the end of a bar. G3, onset-phase entropy, is the Shannon entropy of note-onset phases within the bar, normalized to the maximum possible value. G4, phrase span, is the mean distance between successive phrase or section boundaries, measured in bars. G5, section-boundary closure ratio, is the proportion of closure behaviour at *hâne*, *mülâzime*, and *serbâne* boundaries. Because this final indicator could be calculated only for works containing section-boundary markers, its sample size was 132 rather than 149.

Five directional hypotheses were derived from the thesis, all referring to the expected change from the pre-1650 to the post-1850 period: H1, the cross-barline connection ratio increases; H2, final-pitch–bar-end alignment decreases; H3, onset-phase entropy increases; H4, phrase span expands; and H5, the section-boundary closure ratio decreases.

The predefined criterion for theoretical support required at least three of the five indicators to display stable period trends in the predicted direction after accounting for composer effects. The primary trend test was Spearman's rank correlation between the ordered four-period strata and each indicator (Spearman, 1904). Trend robustness was evaluated using the composer-clustered bootstrap procedure described above (2,000 iterations, fixed seed 20260720, percentile-based 95% confidence intervals using the 2.5th and 97.5th percentiles; Efron & Tibshirani, 1994). Because five ordered tests were conducted, *p*-values were adjusted using the Benjamini–Hochberg procedure (Benjamini & Hochberg, 1995). No single statistic was treated as decisive: direction, bootstrap slope, whether the confidence interval included zero, the bootstrap *p*-value, and the adjusted *q*-value were evaluated jointly.

Results

Table 10. Historical trend tests for the five *usul-melody* indicators (pooled, *n* = 149; *n* = 132 for section-boundary closure)

Indicator	<i>n</i>	Spearman's ρ	<i>p</i>	Bootstrap slope	95% confidence interval	Bootstrap <i>p</i>	Adjusted <i>q</i> (BH)
Cross-barline connection ratio	149	0.000	0.999	−0.019	[−0.053, 0.025]	0.352	0.999
Final-pitch–bar-end alignment	149	0.001	0.991	−0.000	[−0.021, 0.014]	0.997	0.999
Onset-phase entropy	149	0.130	0.114	0.004	[−0.001, 0.011]	0.123	0.570
Phrase span (bars)	149	0.055	0.504	0.888	[−0.832, 2.691]	0.287	0.840
Section-boundary closure ratio	132	−0.089	0.313	−0.023	[−0.104, 0.034]	0.510	0.782

The table 10 shows that none of the five indicators displays a stable directional trend across historical periods. The associations for the first two indicators are effectively zero. Although the strongest candidate, onset-phase entropy (ρ = 0.130), shows a slight movement in the predicted direction, its confidence interval includes zero, and no result survives correction for multiple testing (all *q*-values $\geq$ 0.570).

Table 11. Median values of the five indicators by period $\times$ form cell

Period	Form	<i>n</i>	Connection ratio	Final pitch–bar end	Phase entropy	Phrase span	Section closure
Pre-1650	<i>Peşrev</i>	9	0.829	0.152	0.989	2.243	0.400
Pre-1650	<i>Saz semâisi</i>	5	0.721	0.060	0.962	2.188	0.714
1650–1750	<i>Peşrev</i>	11	1.000	0.163	0.993	1.632	0.633
1650–1750	<i>Saz semâisi</i>	7	0.644	0.048	0.961	2.241	0.714
1750–1850	<i>Peşrev</i>	14	0.571	0.112	0.992	0.614	0.429
1750–1850	<i>Saz semâisi</i>	7	0.645	0.125	0.961	2.300	0.500
Post-1850	<i>Peşrev</i>	48	0.645	0.118	0.995	0.883	0.429
Post-1850	<i>Saz semâisi</i>	48	0.792	0.096	0.979	3.215	0.545

This descriptive analysis reveals fluctuations obscured by the pooled test. Among *peşrevs*, the median cross-barline connection ratio reaches 1.000 in the 1650–1750 period, declines markedly in 1750–1850, and rises again after 1850. Phrase span likewise reaches its lowest value in the same period. Because this trough is confined to a single period and forms a V-shaped pattern in which both ends return toward earlier levels, it cannot be interpreted as evidence of lasting historical transformation. It is more plausibly read as a fluctuation related to differences in the *usul* composition of the historical strata.

Table 11. Number of compositions per *usul* and median values of the five indicators⁵

<i>Usul</i>	<i>n</i>	Connection ratio	Final pitch–bar end	Phase entropy	Phrase span	Section closure
<i>Aksak Semai</i>	67	0.745	0.091	0.976	2.695	0.571
<i>Devr-i Kebir</i>	25	0.636	0.118	0.992	0.598	0.317
<i>Muhammes</i>	13	0.750	0.091	0.994	0.770	0.300
<i>Fabte</i>	8	0.540	0.136	0.995	0.818	0.429
<i>Hafif</i>	8	0.571	0.127	0.989	0.729	0.321
<i>Cenber</i>	4	0.911	0.193	0.996	1.938	0.514
<i>Çifte Düyek</i>	4	0.663	0.089	0.992	0.777	0.750
<i>Düyek</i>	4	0.818	0.221	0.993	4.950	0.429
<i>Ağır Düyek</i>	3	0.582	0.095	0.986	2.234	0.500
<i>Darb-ı Fetih</i>	3	0.667	0.136	0.992	0.277	0.250

Table 11 offers a clue as to what the indicators may actually measure. The connection ratio is higher in *Cenber* (0.911) than in *Aksak Semai* (0.745), and higher in *Aksak Semai* than in *Devr-i Kebir* (0.636) and *Fabte* (0.540). This ordering corresponds to differences in *usul* cycle length: longer cycles provide greater space for melodic expansion

⁵ *Usuls* represented by at least three works (*n* $\geq$ 3); *usuls* represented by a single work are listed separately in the article.

without crossing a barline. The possibility that this indicator reflects *usul* cycle length more strongly than historical period therefore constitutes an alternative explanation discussed below.

The principal finding is that none of the five indicators exhibits a stable directional trend across historical periods. Even G3, the strongest candidate ($\rho = 0.130$, $p = 0.114$, adjusted $q = 0.570$), remains within the bounds of uncertainty, and its confidence interval includes zero. No result survives correction for multiple testing.

Accordingly, despite the period- and *usul*-based fluctuations visible in the descriptive tables, the pooled trend tests reveal no statistically stable direction across the four historical strata.

Study 4. Can Feldman's Thesis of the Model–Nazire Relationship Generate Candidates for Expert Review through Computational Similarity Ranking?

Theory 4: Computational Prioritization of Model–Nazire Relationships

In *Music of the Ottoman Court*, Walter Feldman does not portray the transformation of the early Ottoman instrumental repertoire as a single linear progression (Feldman, 2024). In his account, short, repetitive, and relatively independent *terkibs*, self-contained melodic units spanning several bars, gradually become integrated into longer *makam seyir* trajectories. During this process, the relationship between melody and the *usul* cycle, the *hâne–mülâzime* structure, and model–*nazire* relationships among composers change together. A *nazire* is the practice of composing a new work in the style of an existing one, occupying a broad space between literal imitation and wholly original reinterpretation. A composer writing a *nazire* shares the model's *makam*, *usul*, and often its formal plan and conventions of *seyir*, while reconstructing the melodic material through new phrases. In Feldman's account, *nazire* is not merely a label of similarity but relational evidence of how repertorial models circulated among composers.

The origin of *nazire* introduces a significant tension. The practice emerged in literary history as an Ottoman poetic tradition tied to text and *aruz* metre: one poet composed a new poem in response to another poet's *gazel*, using the same metre, rhyme, and thematic imagery. By transferring the term to textless instrumental forms such as the *peşrev* and *saz semâîsi*, Feldman converts the verbal–metrical–rhyming structure of literary *nazire* into a *makam–usul–seyir* framework (Feldman, 2024). This transfer is not left entirely unquestioned in Feldman's text and should not pass without scrutiny here. Whereas a poetic *nazire* leaves a directly recognizable verbal and metrical echo for the listener, the trace of an instrumental *nazire* can be sought only in structural and *seyir*-based textures. Consequently, the weak detection of *nazire* through measured “structural fingerprint similarity” may reflect not a failure of the method but a change in the very criterion of validation when a literary concept is transferred into the instrumental domain.

This theory has scarcely been tested within computational musicology. Bozkurt, Ayangil, and Holzapfel (2014) observe that computational research on Turkish *makam* music using symbolic data has advanced in *makam* recognition, *usul* analysis, and surface-level similarity measurement, but that this body of work has not been developed into a prioritization tool for relationships such as *nazire*, which require the combined consideration of multidimensional formal design, *seyir* trajectory, and textural conventions. The research question of Study 4 therefore asks:

- Can potential *nazire* pairs be prioritized for expert review through similarities in *makam*, *usul*, form, and *seyir*?

This question is deliberately framed not as a hypothesis test but as a candidate-prioritization exercise; it does not claim that “new *nazires* have been discovered.” No modern corpus contains validated *nazire* labels. What can therefore be tested is not Feldman's historical thesis itself but the surface-level proxies expected to leave traces in notation: shared *makam–usul–form* configurations, similar structural-textural statistics, and comparable *seyir* trajectories.

Testing the Theory

The sample was restricted to *peşrevs* and *saz semâîsi* contained in the open symbolic corpus derived from SymbTr (Karaosmanoğlu, 2012); the unit of analysis was the musical composition rather than the individual note. Three exclusion criteria were applied: records containing fewer than twenty valid notes, records lacking composer attribution, and records that could not be assigned to one of the four attribution strata. The final sample comprised 149 compositions by 77 composers, divided into four historical layers according to composer attribution (14 pre-1650,

18 from 1650–1750, 21 from 1750–1850, and 96 post-1850).¹ These temporal boundaries (1650, 1750, and 1850), derived from composers' life spans, constitute a common convention throughout the present studies. They represent an approximate analytical simplification of Feldman's more detailed historical periodization and are not claimed to correspond exactly to the chronological boundaries proposed in his text. The sample is not a probabilistic historical sample; its selected-transmission character and the resulting restriction of the target population were discussed in detail in the Introduction.

MakamKit, the computational engine used in the analysis, reads symbolic notation derived from SymbTr, associates each work with *makam*, *usul*, and formal metadata, and extracts structural and *seyir*-related features at the composition level. In the present study it was employed to construct a two-layer similarity measure. The first layer was *structural fingerprint similarity*. Nine features, namely direct repetition, transposed repetition, *seyir* roughness, cadential convergence, section-edge similarity, directional continuity across section boundaries, cross-barline connection ratio, melodic density, and pitch range (in commas), were standardized using *z*-scores (with missing values imputed by the sample median) and compared using cosine similarity. The second layer was *trajectory similarity*. Each composition was divided into twelve equal segments, the duration-weighted pitch centroid was calculated for each segment, and the resulting standardized twelve-point curves were compared using unconstrained classical Dynamic Time Warping (DTW) distance (Sakoe & Chiba, 1978). The total alignment cost was divided by 24 and transformed to the 0–1 interval using the function $e^{-(\text{distance})}$. The composite similarity score combined the two components with weights of 65% for structural fingerprint similarity and 35% for trajectory similarity. Because no validated *nazire* labels exist, these weights were not optimized but explicitly declared as an exploratory design choice. Candidate pairs were searched only within identical *makam-usul-form* cells. Although this restriction improves comparability, it necessarily excludes genuine *nazires* that may have been transformed into different *usuls* or formal types. A single statistical validation test was planned to evaluate the ranking: a one-sided Mann–Whitney *U* test comparing the scores of same-composer pairs with those of different-composer pairs. To examine the influence of the weighting scheme on the ranking, the analysis was additionally repeated under four weighting profiles (25%/75%, 50%/50%, 65%/35%, and 75%/25%) and compared with the reference profile (65%/35%). All similarity values were rounded to three decimal places. The ranking procedure was entirely deterministic and involved neither random sampling nor model fitting.

Results

The search produced 34 comparable pairs within identical *makam-usul-form* cells. The highest composite similarity score was 0.885, whereas the lowest was 0.068.

Table 12. Top ten candidate *nazire* pairs ranked by composite similarity score (all 34 pairs are included in the data package)

Rank	Makam	Usul	Form	Composition A	Composition B	Structural fingerprint similarity	Trajectory similarity	Composite score
1	Segâh	Aksak Semai	Saz semâîsi	k1807 · Hızır Ağa	k1808 · Nâyi Osman Dede	0.896	0.863	0.885
2	Uşak	Aksak Semai	Saz semâîsi	k2123 · Dede Salih Efendi	k2125 · Neyzen Aziz Dede	0.907	0.718	0.841
3	Hüseynî	Aksak Semai	Saz semâîsi	k0793 · Lavtacı Andon	k0795 · Tatyos Efendi	0.942	0.641	0.837
4	Nihavent	Aksak Semai	Saz semâîsi	k1454 · Mesut Cemil	k1455 · Vecdi Seyhun	0.877	0.683	0.809
5	Hicaz Hümayun	Aksak Semai	Saz semâîsi	k0596 · Tanburi İsak	k0597 · Velî Dede	0.746	0.747	0.746
6	Beyatî	Aksak Semai	Saz semâîsi	k0190 · Kanuni Ömer Efendi	k0191 · Nâyi Osman Dede	0.770	0.695	0.744
7	Sabâ	Aksak Semai	Saz semâîsi	k1717 · Neyzen Aziz Dede	k1718 · Zaharya	0.741	0.739	0.740
8	Beyatî	Aksak Semai	Saz semâîsi	k0190 · Kanuni Ömer Efendi	k0192 · Mısırlı Udi İbrahim Efendi ("Arap")	0.777	0.660	0.736
9	Uşak	Aksak Semai	Saz semâîsi	k2124 · Kantemiroğlu	k2125 · Neyzen Aziz Dede	0.746	0.708	0.733
10	Segâh	Devr-i Kebir	Pesrev	k1764 · Nâyi Osman Dede	k1765 · Neyzen Yusuf Paşa	0.635	0.861	0.714

The composite-score formula can be verified directly by hand. For the first-ranked pair, for example: $0.65 \times 0.896172 + 0.35 \times 0.862931 = 0.884538 \approx 0.885$.

The most striking finding does not lie within the similarity columns themselves but in the dataset as a whole: none of the 34 candidate pairs is attributed to the same composer.

Table 13. Similarity diagnostics: results of the same-composer validation test

Indicator	Value
Number of comparable pairs	34
Number of same-composer pairs	0
Median score of same-composer pairs	— ^a
Median score of different-composer pairs	0.589
Interquartile range of different-composer pairs (Q1 = 0.246; Q3 = 0.728)	0.482
Mann-Whitney <i>U</i>	— ^a
One-sided <i>p</i> (same-composer pairs higher)	— ^a

^a Not computable because no same-composer control pairs were available.

The single statistical validation planned for evaluating the ranking could not be performed because the comparison pool contained no same-composer pairs. This outcome reflects not a methodological failure but rather the historical convention that composers of *saz semâîs* typically produced only a single work within the genre, leaving no internal control pairs for statistical comparison.

Table 14. Structure of the candidate pool based on 34 pairs

Indicator	Value
Number of <i>makam-usul-form</i> cells	21
Cells containing one pair	16
Cells containing three pairs	4
Cells containing six pairs	1 (<i>Mabur</i>)
Number of works represented in the pool	48 (32.2% of the corpus)
Works not included in any pair	101 (67.8% of the corpus)
<i>Saz semâîs</i> pairs in <i>Aksak Semai usul</i>	30 (88.2%)
<i>Peyrev</i> pairs in <i>Devr-i Kebir usul</i>	4 (11.8%)
Structural fingerprint similarity range	−0.344 to 0.942
Trajectory similarity range	0.634 to 0.863
Composite score range	0.068 to 0.885

The internal structure of the candidate pool is highly imbalanced. Two-thirds of the 149 works do not enter any pair, while nearly nine out of every ten candidate pairs belong to a single *usul-form* combination, namely *saz semâîs* in *Aksak Semai usul*. The two similarity components also differ substantially in discriminatory power. Structural fingerprint similarity ranges from near opposition (−0.344) to strong correspondence (0.942), whereas trajectory similarity is compressed within a considerably narrower interval (0.634–0.863). This suggests that the smoothed twelve-point trajectories are inherently similar to one another and that the trajectory component possesses limited discriminatory power when used independently.

Table 15. Sensitivity to weighting profiles: comparison with the reference profile (65% structural fingerprint / 35% trajectory)

Weighting profile (structural fingerprint/trajectory)	Rank correlation with reference	Top-ten overlap
25% / 75%	0.924	0.8
50% / 50%	0.992	1.0
65% / 35% (reference)	1.000	1.0
75% / 25%	0.997	1.0

The ranking is relatively robust to the choice of component weights. Even when the trajectory component is assigned an excessive weight of 75%, the rank correlation remains above 0.92 and at least eight of the top ten candidates remain unchanged. Once the structural fingerprint component receives a weight of 50% or more, the top-ten list becomes completely stable. This indicates that the highest-ranked candidates are not merely products of a single arbitrary weighting decision and that the two components generally point in the same direction. However, ranking stability does not establish ranking validity, and the empty control cells reported in Table 13 underscore this distinction.

In summary, the screening procedure generated 34 comparable pairs from the corpus of 149 works. None of these pairs involved the same composer, and the resulting ranking remained largely stable across four alternative weighting profiles⁶.

⁶ The period-wise distribution and the number of composers (77) could not be verified directly because these fields were not included as separate variables in the final results dataset; therefore, they could only be verified indirectly from the overall totals (149).

Study 5. Can Feldman's Four-Period Periodization Be Recovered through Blind Clustering of Musical Features with Dates and Composer Names Concealed?

Theory 5: Four-Period Formal Periodization

In *Music of the Ottoman Court: Makam, Composition and the Early Ottoman Instrumental Repertoire* (2024), Walter Feldman does not treat the *terkib-seyir* transformation summarized in the Introduction as a single linear process of “progress.” Rather, he conceptualizes it as the joint evolution of several simultaneously developing components, including the ordering of tonal centres, the relationship between melody and the *usul* cycle, *hâne-mülâzime* continuity, and model-*nazire* relationships. The aspect directly relevant to Study 5 is that this transformation ultimately implies that the repertoire can be divided into historically meaningful periods. Feldman's own periodization distinguishes as many as eight subperiods; however, the central theoretical claim is that these periods should leave readable traces not only in archival chronology but also in the formal texture of the works themselves.

This claim points to a recurrent epistemic problem in historical musicology. Stylistic periodization is commonly based on documentary evidence, biography, and chains of transmission, while the question of whether the internal organization of musical works independently corroborates that periodization often remains untested. If Feldman's theory is correct, then even when period labels are withheld, measurable features of the compositions, including motivic organization, *seyir*, sectional structure, and *usul-melody* articulation, should display patterns that differentiate the historical periods. Study 5 examines this expectation by testing not Feldman's theory in its entirety but a restricted and falsifiable sub-proposition derived from it: When dates and composer names are concealed, can musical features reconstruct the attributed historical periods? Two important limitations of this formulation must be stated at the outset. First, the data do not consist of original seventeenth- and eighteenth-century manuscripts but of an openly accessible symbolic corpus transcribed into modern notation. The independent variable is therefore not the exact composition date of each work but one of four broad historical strata inferred from composer attribution. Second, Feldman's thesis that the “Ottoman-Turkish style differentiated itself from the Persianate environment” is not directly tested here. Such a test would require a comparative corpus of Iranian and Transoxanian *pişrev* repertoires, which is unavailable in the present dataset. The study tests only a much narrower precondition of that broader thesis:

- Does the instrumental repertoire preserved in modern transcription exhibit a measurable formal structure associated with its attributed historical periods?

Testing the Theory

Sample

The analysis was conducted on a sample of 149 *peşrevs* and *saz semâîs*. The works were drawn from an openly accessible symbolic notation corpus transcribed into modern notation, of the type described by Karaosmanoğlu (2012), available for non-commercial research use with attribution. Records containing fewer than twenty valid notes, lacking composer attribution, or not assignable to one of the four attribution strata were excluded. The unit of analysis was the complete composition rather than the individual note.

The final sample was distributed across the four historical strata as follows: 14 works attributed to the pre-1650 period, 18 to 1650–1750, 21 to 1750–1850, and 96 to the post-1850 period. The sample represented 77 different composers. The pronounced imbalance in this distribution, particularly the small early-period cells of 14 and 18 works, is discussed below as a separate limitation. The source text does not demonstrate at the page level which specific transitions in Feldman's own historical narrative correspond exactly to these four boundaries. The four-period framework employed here is therefore a coarse approximation of Feldman's more fine-grained periodization, which extends to eight subperiods, and is not presented as a one-to-one mapping.

It should be emphasized that this sample retains the selected-transmission character discussed in detail in the Introduction and does not constitute a probabilistic sample of the historical universe addressed by Feldman. The target population of the results is therefore not “the entire Ottoman instrumental repertoire” but the body of works observable within the specified chain of transmission.

Feature Extraction with MakamKit

MakamKit is a multilayer computational analysis engine developed for Turkish *makam* music and incorporates both symbolic and audio-based modules. Only the symbolic layer was used in this study. The following feature families were extracted from each work:

Intervalllic motifs: differentiation between direct and transposed repetitions within windows of three and four intervals.

Seyir trajectory: each work was divided into twelve windows containing equal numbers of notes. For each window, the duration-weighted relative pitch centre was calculated, together with trajectory roughness between successive windows, the number of directional changes, and the degree of convergence toward the final pitch within the final fifth of the composition.

Section boundaries: histogram similarity and directional continuity were calculated within eight-note contexts surrounding transitions marked as *hâne*, *mülâzime/teslim*, and *serhâne*.

Usul-melody articulation: features included the metrical phase of note onsets, note events crossing barlines, coincidence of the final pitch with bar endings, and the span between sectional boundaries measured in bars.

None of these measurements constitutes a direct representation of historical *usul* accentual weights or original performance style. They represent only the measurable temporal and pitch organization visible on the surface of modern notation.

Blind Design

Composer names, period labels, composition titles, and dates were not supplied to the clustering model in any form. Only features related to motifs, *seyir*, section structure, *usul-melody* articulation, density, ambitus, and length were used. After feature standardization, KMeans clustering was run for values of k ranging from 2 to 8, with fifty different random initializations for each solution.

Within-cluster cohesion was measured using the silhouette coefficient (Rousseeuw, 1987). External agreement between the clusters and the four historical period labels was evaluated using the adjusted Rand index, ARI (Hubert & Arabie, 1985), and normalized mutual information, NMI. These measures address different questions. The silhouette coefficient indicates whether the data form internally coherent natural groupings, whereas ARI and NMI indicate whether any such groupings correspond to the attributed historical periods.

Classification

Independently of the clustering analysis, the extent to which musical features carry period information was examined through composer-grouped five-fold cross-validation. Works by the same composer were never permitted to occur simultaneously in both the training and test folds. This restriction was introduced to prevent the model from mistaking a composer-specific stylistic signature for a historical-period signal.

Balanced accuracy was reported for the four-period classification task. In addition, a random forest classifier (Breiman, 2001) was used to calculate ROC-AUC for a coarser binary distinction between the two early periods, pre-1650 and 1650–1750, and the two later periods, 1750–1850 and post-1850. The model used 500 trees and required a minimum of two observations per leaf. This binary division does not constitute a full test of Feldman's original eight-period periodization. It examines only whether a broad early-late stylistic gradient is present.

Uncertainty Assessment

For all diagnostics, composer-blocked permutation tests (1,000 iterations; fixed random seed = 20260720) and composer-bootstrap 95% confidence intervals were calculated (Efron & Tibshirani, 1994). In both procedures, works attributed to the same composer were resampled together, thereby preventing prolific composers from creating artificial precision. In each permutation iteration, only the composer-to-period assignment was shuffled; all works by a given composer received the same randomly assigned period label in that iteration. The clustering labels, cross-validated classification predictions, and predicted probabilities were held fixed. In other words, the KMeans and random forest models were not retrained in each iteration; only the observed, fixed model outputs were compared repeatedly with randomized period assignments. This design tests whether the fixed model output exhibits greater agreement with the observed period labels than would be expected under a random composer-to-period assignment. It does not test whether the model itself, when retrained on randomly permuted labels, could generate a comparable

level of agreement. This distinction is addressed separately in the Limitations section below in relation to the potentially optimistic nature of the reported p -values.

Results

Table 16. Cluster-number scan, $k = 2-8$

k	Silhouette	ARI (four periods)	NMI (four periods)
2	0.1349	0.0164	0.0338
3	0.1136	0.0522	0.0651
4	0.1149	0.0462	0.0746
5	0.1146	0.0150	0.0720
6	0.1149	0.0299	0.0702
7	0.1174	0.0289	0.0812
8	0.1110	-0.0036	0.0813

The highest internal cohesion was observed for the two-cluster solution. However, the silhouette value of 0.1349 remains low and indicates that the feature space does not contain clearly separated natural groupings. At no value of k does the ARI approach a strong correspondence with the four historical periods.

Table 17. Cross-validated confusion matrix for the four-period classification ($n = 149$)

Actual period	Predicted: 1	Predicted: 2	Predicted: 3	Predicted: 4
1 (pre-1650, $n = 14$)	0	1	0	13
2 (1650–1750, $n = 18$)	0	4	3	11
3 (1750–1850, $n = 21$)	1	3	0	17
4 (post-1850, $n = 96$)	2	2	3	89

The model does not reliably distinguish any of the four periods and assigns the great majority of works to the fourth class, post-1850, irrespective of their actual period labels. This tendency may result from the structural dominance of the largest class ($n = 96$) relative to the other three classes.

Table 18. Summary of periodization diagnostics ($n = 149$; number of composers = 77)

Criterion	Value
Four-period balanced accuracy	0.2873
Early/late binary ROC-AUC	0.6976
ARI ($k = 4$)	0.0462
NMI ($k = 4$)	0.0746
Silhouette ($k = 4$)	0.1149
Number of permutation iterations	1,000
Permutation seed	20260720

The balanced accuracy of the four-period classification (0.2873) is only slightly above the chance baseline of 0.25. By contrast, the ROC-AUC value for the early/late binary distinction (0.6976) indicates a more discernible signal. The uncertainty assessment reported in Table 19 shows that this distinction is also statistically consistent.

Table 19. Uncertainty and permutation assessment.

Diagnostic	Observed	95% composer-bootstrap CI	Permutation p
Four-cluster ARI	0.0462	[0.0076, 0.1194]	0.0020
Four-cluster NMI	0.0746	[0.0418, 0.1711]	0.0040
Four-period balanced accuracy	0.2873	[0.2392, 0.3844]	0.1159
Early/late ROC-AUC	0.6976	[0.5807, 0.8019]	0.0060

The permutation p -values indicate that the ARI and NMI differ statistically from those expected under random label assignments ($p = 0.0020$ and $p = 0.0040$, respectively), although the absolute magnitude of this agreement remains small. The permutation test for four-period balanced accuracy does not cross the conventional threshold of statistical significance ($p = 0.1159$), whereas the early/late ROC-AUC carries a statistically consistent signal ($p = 0.0060$). This result provides the numerical basis for the study's conclusion of "limited support for a coarse historical gradient." In summary, the findings do not reveal a fine-grained period-specific formal organization in modern notation, whereas traces of a broader early/late historical gradient remain detectable.

Study 6. Is Feldman's Thesis on the Relationship Between Formal Length and Historical Style in the First *Selâm* of the Mevlevi *Âyin* Ready to Be Tested through a Source-Anchored Validation Protocol?

Theory 6: The Form–Style Relationship in the First *Selâm*

In *From Rumi to the Whirling Dervishes: Music, Poetry, and Mysticism in the Ottoman Empire* (2022), Walter Feldman interprets the First *Selâm* section of the Mevlevi *âyin* as a domain in which the composer's technical accumulation and musical vision can be observed most intensively. According to Feldman, there is a directional relationship between the formal length of the First *Selâm* and the historical position of the composer to whom it belongs: whereas in ancient or early examples, such as the anonymous *Dügâh Âyini*, the verse and *terennüm* sections are compactly organized within a few *devr-i revân* cycles, later composers are expected to allow melodic phrases to extend beyond cycle boundaries and to expand the *makam* region. This interpretation is not a superficial observation reducible merely to an increase in the number of bars. What Feldman emphasizes is the reorganization of long-range melodic thought so that it encompasses the *usul* cycle, that is, the development of the text-*usul-seyir* relationship into a more complex architecture in the hands of the composer. The theory requires three elements to be monitored simultaneously: where the phrase boundary ends, how this boundary relates to the *devr-i revân* cycle, and how many times the *makam* region, or modulation, changes. Feldman's (2022) printed text supports these elements through partial formal descriptions of the *Dügâh* and *Beyati âyins*. For the *Saba Âyini*, however, he does not reproduce the notation itself; instead, he directs the reader to the web supplement and to İrsoy's *Mevlevi Ayinleri* (Vol. 11, pp. 565–568). Because neither the article nor the source documents provide a complete bibliographic reference for İrsoy's compilation, including the author's full name, publication year, and publisher, the present study refers to this source only through Feldman's (2022) citation, using the available volume and page information. Although Feldman's 2024 work *Music of the Ottoman Court* is also included in the bibliography of this study package, the thesis concerning the First *Selâm* is derived directly from the 2022 book.

On the basis of this framework, the research question was formulated as follows: when the source-verified First *Selâm* scores of the *Dügâh*, *Beyati*, and *Saba âyins* are coded according to the same measurement protocol, does their chronological ordering produce a descriptive pattern involving (a) phrases extending beyond *devr-i revân* cycles and (b) an increase in the number of independent *makam* regions? Three predefined hypotheses operationalize this question. H1 predicts that the expert-coded mean phrase length, measured in units of *devr-i revân*, increases in chronological order; H2 predicts that the proportion of phrases spanning at least one *usul* cycle boundary increases; and H3 predicts that the number of regions or modulations calculated from independent expert *makam*-region labels increases. The decision rule was fixed in advance: if all three indicators follow the expected direction, the result will be recorded as *descriptive_support*; if at least one moves in the opposite direction, it will be recorded as *descriptive_counterexample*. Because only three cases are involved, the production of a *p*-value or a population-level generalization was not planned. This positions the proposed relationship identified by Feldman as a selected-case replication attempt rather than as a claim based on probabilistic sampling.

Testing the Theory

The sample was restricted to three cases selected with chronological and theoretical contrast in mind: the anonymous *Dügâh Âyini*, regarded as ancient or early; the *Beyati Âyini* attributed to Köçek Mustafa Dede, representing the middle period; and İsmail Dede Efendi's *Saba Âyini* of 1823, representing the late period and having a precise date. This set of three cases was selected to represent the early-late contrast implied by Feldman's thesis both chronologically and theoretically; however, it is not a probabilistic sample of the Ottoman Mevlevi *âyin* repertoire. The article itself and the present section acknowledge this explicitly as a limitation. The analysis was designed as a protocol planned to be conducted through the symbolic layer of the five-layer MakamKit engine introduced in the Introduction, comprising symbolic melodic analysis, audio-based pitch analysis, *usul/aksak* analysis, lyric/emotion analysis, and similarity clustering. In this study, the role of the engine is limited to that of a validation instrument. This four-step validation operates as follows: passing source-traceable MusicXML files (W3C Music Notation Community Group, 2021) through schema validation; calculating phrase length by dividing expert-marked phrase boundaries by the relevant *devr-i revân* cycle length; determining cycle crossing by counting *usul* boundaries that remain within open intervals; and deriving the number of regions or modulations from independent expert *makam*-region labels. In cases where no expert boundary is available, the engine can produce a duration-based heuristic segmentation; however, this output is

used only for software diagnostics, is marked as *phrase_source: heuristic*, and cannot enter the confirmatory H1/H2 decision. The same principle applies to *makam* expansion: an algorithmic pitch histogram does not substitute for expert labels; a missing value is not converted into zero but is retained as *n_modulations: null*. This design choice is intended to prevent the error of treating an unmeasured variable as zero. The opening of the confirmatory run was made conditional upon the simultaneous satisfaction of four data-preparation gates: (1) the availability of source-traceable MusicXML files for all three cases, (2) the inclusion of an expert-generated *phrase_end_qL* sequence in each file, (3) the provision of independent *region_labels*, and (4) the double-checking of note, phrase, and region codes against the source page. Schema validation was kept strict enough to reject unknown annotation fields. Quality assurance additionally requires that the person responsible for the transcription and the second reviewer be different individuals, that disagreements be preserved before blinded consensus, and that reliability be reported separately. The consensus dataset and the sensitivity run excluding disputed boundaries will be presented separately; if the direction of the result changes, it will be classified as *coding sensitivity*. No formal multiple-comparison correction, such as Bonferroni correction, was applied beyond the combined three-indicator decision rule because no *p*-value is produced in any case owing to the presence of only three cases.

Results

This section is not a “positive result” report in the conventional sense: the findings presented below document not that the theory has been tested and supported or refuted, but that the conditions required for testing are not yet ready. The confirmatory sample size is zero; the values *confirmatory_ready=false* and *historical_test_status=not_run_missing_real_scores* are fixed in the machine-readable results file. H1, H2, and H3 all remain untested. This means not that the hypotheses have been refuted, but that they have not been tested. Tables 20 and 21, together with the status fields in the results file, are direct reproductions cross-checked against the results JSON.

Table 20. Data preparation by case

Case	Real score ready	Published anchors	Evidence coverage
<i>Dügâh</i>	No	6	Partial printed example and textual analysis
<i>Beyati</i>	No	7	Partial printed example and textual analysis
<i>Saba</i>	No	1	Identifies the local book source; the score is located in the web supplement or the İrsoy volume

Table 20 shows that no source-traceable real score is ready for any of the cases, although six, seven, and one verifiable published propositions, or anchors, respectively, could be extracted from Feldman’s (2022) printed text. An anchor refers to a concrete proposition attributed to a specific section in Feldman’s printed text and verifiable against the source page, for example, “a modulation to *Segâh* in the third *devr* of *Dügâh* Bend III.” These anchors can be used to supervise data acquisition, but they do not themselves constitute sufficient data for hypothesis testing.

Table 21. Confirmatory decision gates

Gate	Required	Current status
Three source-traceable scores	Yes	Not satisfied
Expert phrase boundaries	Yes	Not satisfied
Independent <i>makam</i> -region labels	Yes	Not satisfied
Double source verification	Yes	Not satisfied
Confirmatory analysis	Yes	Not run

Table 21 shows that none of the four data-preparation gates, nor the confirmatory analysis step dependent upon them, has yet been satisfied. This indicates not that the thesis has been refuted, but that the conditions required for testing have not yet been established.

Table 22. Protocol status flags

Indicator	Value
Number of real scores	0
Number of missing targets (<i>Dügâh</i> , <i>Beyati</i> , <i>Saba</i>)	3
<i>confirmatory_ready</i>	No
<i>historical_test_status</i>	<i>not_run_missing_real_scores</i>
<i>protocol_status</i>	Stage 1 registered report, real scores missing
Figures	None

Table 22 summarizes the protocol’s own status flags and confirms that none of the numerical or visual findings presented in the article carries a claim of historical results. The successful completion of software self-tests using

synthetic data does not alter this assessment, because these tests examine only parsing, schema rejection, cycle counting, and missing-value behavior; they do not generate historical evidence. In summary, none of the four data-preparation gates has yet been satisfied for any of the three cases, the confirmatory analysis has not been run, and H1–H3 have not been tested; the number of real scores is zero, and the number of missing targets is three.

Conclusion and Discussion

This article derived six separate, pre-bounded propositions from Walter Feldman’s broader account of transformation in Ottoman court instrumental music and tested them using a shared computational infrastructure (MakamKit), a shared data source (a SymbTr-derived symbolic score corpus), and transparent statistical protocols incorporating multiple-comparison correction. The six studies cannot be reduced collectively to a single sentence, but they reveal a highly internally consistent picture: five studies either found no support on the surface of modern notation or detected only a coarse and limited signal, whereas the sixth documented the data-preparation stage of a theory that has not yet become testable.

Table 23. Collective epistemic status of the six studies

Study	Theory	Status	Main finding
Study 1	Transition from <i>terkib</i> to <i>seyir</i> (motif repetition, phrase expansion, cadential convergence)	Not supported; sensitive to component and model	The combined support criterion ($\geq 3/5$ indicators) was not met; the only stable indicator, cadential convergence, moved in the opposite direction from that expected; phrase expansion in the <i>saz semâlsi</i> provided partial within-form support
Study 2	Holistic historical codification of <i>seyir</i> (five convergent indicators)	Not supported	No stable trend emerged in four indicators; cadential convergence moved in the opposite direction ($\rho = -0.325$, $p < 0.001$); all five hypotheses had point estimates with signs opposite to the direction predicted by the thesis
Study 3	Periodic change in the mode of articulation between the <i>usul</i> cycle and the melody	No trace found on the surface of modern notation	None of the five indicators produced a significant trend after correction (all $q \geq 0.570$)
Study 4	Computational prioritization of the model– <i>nazire</i> relationship	No validated discovery; candidate ranking only	Only 34 comparable pairs could be generated from the corpus of 149 works; zero same-composer control pairs
Study 5	Four-period formal periodization	Not recoverable as natural clusters; limited evidence for a coarse gradient	Balanced accuracy = 0.2873 (permutation $p = 0.1159$); early/late ROC-AUC = 0.6976 ($p = 0.0060$)
Study 6	Form–style relationship in the First <i>Selâm</i>	Not tested — Stage 1 registered protocol	Number of real scores = 0; confirmatory_ready = false

A first reading of this table may suggest that Feldman’s theoretical legacy has been “lost” in the modern symbolic transmission sample; such a reading would be premature. None of the six studies falsifies the theory itself. Each study tested a projection of the theory reduced to notation and measured through a narrow set of proxies. The rationale for this distinction must be established here once and forcefully, as the foundation for all six discussions below: the SymbTr-derived corpus used as MakamKit’s input (Karaosmanoğlu, 2012) is not a probabilistic sample of seventeenth- and eighteenth-century performance. It is a selected transmission sample that has passed through the filters of survival into the present, preservation within the *meşk* chain, transcription into notation, inclusion in an edition, and attribution to a composer. Every link in this filtering process constitutes a selection and reduces the flexibility of performance, ornamentation, pitch adjustment, temporal freedom, and the weight of *düm–tek* strokes to fixed signs. In addition, the 53-comma Arel–Ezgi–Uzdilek (AEU) system on which MakamKit’s pitch measurements are based is a framework theorized in the twentieth century. Applying this framework retrospectively to early works entails a distinctive risk of anachronism inherent in the conceptual foundations of measurement, separate from the uncertainty carried by the transmission chain. These two boundaries, the transmission filter and conceptual anachronism, are common to all six studies and will not be repeated for each study below; instead, the discussion will focus only on how each finding is positioned on this shared foundation. The fact that most of the results take the form of “not supported” or “limited evidence” therefore does not constitute a refutation of Feldman’s historical proposition. Rather, it maps the trace of that proposition on the modern transmission surface and the extent to which this trace can be captured by the selected proxies.

From *Terkib* to *Seyir*: An Attempt at Integration (Studies 1 and 2)

Feldman's (2024) central thesis is that, beginning in the second half of the seventeenth century, the *peşrev* and *saz semâisi* repertoire shifted from the modular *bâne-mülâzime* structure known as *terkib* toward a logic of *seyir* encompassing the work as a whole. Study 1 tested this transition using indicators such as motif repetition, phrase expansion, and cadential convergence, whereas Study 2 tested the same transition through five indicators based on tonal center of gravity: roughness, directional change, directional continuity at section boundaries, similarity of pitch distributions at section boundaries, and cadential convergence. The patterns produced by the two studies overlap strikingly: the number of indicators moving in the expected direction remained below the combined support threshold, whereas cadential convergence moved in both studies in a statistically significant direction directly opposite to that expected (in Study 2, $\rho = -0.325$, $p < 0.001$). In Study 2, all five hypotheses had point estimates with signs in the opposite direction; this indicates a consistent pattern rather than the accidental reversal of a single indicator.

When confronted with Feldman's text, this consistent reversal can be interpreted in two different ways. First, the cadential-convergence measure itself is a coarse proxy: it mechanically treats the final note as the finalis, whereas in some forms the concluding section (*mülâzime*) may close with a suspended cadence; the measure does not distinguish this subtlety. In that case, the reversal may indicate not that *seyir* historically weakened, but that the proxy misrepresents the cadential region. Second, as indicated in the limitations section of Study 2, the trajectory of cadential convergence across period medians is not unidirectional but instead peaks in the 1650–1750 stratum and subsequently declines. This suggests that the linear expectation framework constructed by H1–H5 may be more restrictive than Feldman's own nuanced and multidimensional narrative. Feldman already frames the transition from *terkib* to *seyir* not as a one-dimensional "progression," but as a multidimensional process in which the ordering of pitch centers, the articulation between *usul* and melody, the *bâne-mülâzime* organization, and the *nazire* relationship shift together (Feldman, 2024). The two studies in this article have disentangled only two tonal-structural threads from that broader fabric. The inability to retain one thread does not mean that the fabric itself does not exist. However, the convergence of two independent studies, using two different sets of indicators, on the same reverse pattern of cadential convergence also strengthens the possibility that this regularity is more than a proxy defect and that the transmission chain systematically distorts the signal in a particular direction. The general limitation identified by Bozkurt, Ayangil, and Holzapfel (2014) for computational *makam* research, namely that symbolic measures can capture only the shadow of historical-theoretical concepts, has here materialized in the same direction across two independent tests. Nor is the finding entirely empty: in Study 1, phrase expansion in the *saz semâisi* produced a form-specific partial signal consistent with the thesis within its own within-form correction family. This indicates that the result does not amount to a single flat judgment of "absence," but rather to a negative/mixed judgment sensitive to component and form.

Articulation Between the *Usul* Cycle and the Melody (Study 3)

The *usul* dimension of Feldman's (2024) thesis is based on the hierarchy of accentual weight among *düm-tek* strokes and on how the melody is articulated in relation to this hierarchy. Study 3 tested, through five indicators, whether the melody increasingly extended across the *usul* cycle for longer durations and with greater frequency, and found that none of the five indicators produced a statistically significant period trend after correction (all $q \geq 0.570$). Here, the negative result may be explained by a mechanism different from those operating in Studies 1 and 2: because the source data do not encode the intensity of *düm-tek* strokes, the study was compelled to use an event-phase proxy, and none of the indicators directly measures the accentual hierarchy of the *usul*. The barline itself is not a trace of the historical *usul* cycle, but of the editorial decision of the modern transcriber. The general warning offered by Bozkurt et al. (2014) takes an especially sharp form here: symbolic data provide timing, but not accentual weight. Whereas the *terkib-seyir* transformation in Feldman's text rests on knowledge transmitted from body to body through the *meşk* chain, what this study measures is the shadow of that knowledge cast onto the surface of modern notation.

A confounding factor identified in the limitations section of Study 3 requires an even more cautious interpretation of this negative result: all *saz semâisis* are in the *Aksak Semai usul*, and therefore the effects of form and *usul* cannot be disentangled in these data. The indicators are sensitive to *usul*-cycle length, and because the *usul* composition of the periods differs, the pooled test carries the burden of this confounding factor. Moreover, although a single-period

“trough” corresponding to 1750–1850, in which both the linkage ratio and phrase expansion decline, roughly coincides with the window predicted by Feldman, it constitutes a V-shaped fluctuation rather than a persistent opening; the transformation thesis, by contrast, predicts a unidirectional and lasting trend. Three possible mechanisms, the homogenization of the signal by the transmission chain, the confounding of *usul*-cycle length with period, and the possibility that the transformation proceeded as a nonlinear fluctuation, each find a counterpart in the data, and none can be disentangled from the others within the scope of this article. The trace left on notation by three centuries of transmission appears static; however, one cannot directly conclude from this that historical *usul* practice itself was static.

The Model–*Nazire* Relationship (Study 4)

Feldman’s concept of *nazire* presents an inter-composer relationship of artistic practice that cannot be reduced to the identity of a single motif (Feldman, 2024). Study 4 transformed this relationship into a computational candidate-prioritization problem and, because the corpus contains no historical *nazire* labels, produced only a ranking rather than a validation. The Hızır Ağa–Nâyi Osman Dede *Segâh saz semâîsi* pair at the top of the list (0.885) may reasonably be considered a genuine first point of investigation, because both composers belonged to the late seventeenth-century Ottoman court environment and because Nâyi Osman Dede occupied a distinctive position as a figure who compiled and notated the earlier repertoire of the period. Yet this high score is compatible with two different narratives: a genuine model–*nazire* affinity or a shared transmission practice associated with the same notational tradition. Resolving this distinction is not the task of the machine, but of the historian working through source investigation. The most instructive finding of the study is that the only planned direct sanity check, using a second work written by the same composer within the same cell as a control group, could not be conducted at all because of insufficient data. Although the corpus of 149 works mathematically permits 11,026 possible pairwise comparisons, only 34 of these remained within the same *makam–usul–form* cell, and zero same-composer control pairs were produced. This gap reflects not a weakness in Feldman’s theory, but the generic logic of the modern canon itself: because the *saz semâîsi* became standardized as a form in the nineteenth and twentieth centuries, composers of this period generally wrote only one *saz semâîsi* in almost every *makam*. The convention of “one *saz semâîsi* for each *makam*” structurally prevented the same composer from producing a second work within the same cell. This does not demonstrate that Feldman’s theory of *nazire* is incorrect. On the contrary, it shows how the selective logic of the modern canon may obscure the trace of a historical relationship. The sensitivity analysis across weighting profiles demonstrated that the ranking was not merely the product of a single arbitrary coefficient choice; however, internal consistency is not external validity. The gap generally identified by Bozkurt et al. (2014) has not been closed here, but only mapped: the extent to which the generic logic of the modern transmission canon obscures the historical model–*nazire* relationship remains an open question that cannot be answered with the present data but can be directly tested through future bibliographic investigation.

Formal Periodization (Study 5)

Study 5, which tested a four-part reduction of Feldman’s (2024) eight-period periodization, found that a four-cluster solution did not align with the period labels (ARI = 0.0462, NMI = 0.0746) and that four-period classification remained only slightly above the chance baseline (balanced accuracy = 0.2873; permutation $p = 0.1159$, not exceeding the threshold for statistical significance). By contrast, a coarser early/late distinction carried a statistically consistent weak-to-moderate signal (ROC-AUC = 0.6976; $p = 0.0060$). Although this dual result may initially appear contradictory, it is paradoxically consistent when read through Feldman’s own theoretical framework: Feldman’s periodization does not propose eight separate, sharply bounded blocks, but rather a gradual transition and an overlapping continuity from the brevity of the *terkib* form toward *seyir*-based integration. The discrete and mutually separated groups sought by clustering methods may therefore not constitute an appropriate test of this kind of gradual transformation; the fact that the early/late binary distinction produced a stronger signal than clustering is also consistent with this interpretation. The signal increases when the analysis searches for a broad direction rather than a sharp classification.

Nevertheless, an alternative interpretation of the low cluster alignment cannot be ignored: the modern transmission process itself may have homogenized historical stylistic signatures. Bozkurt et al. (2014) emphasize that symbolic transcription largely standardizes the microtonal and rhythmic details specific to performance practice. These two interpretations, the continuity inherent in the theory and the artefact of transmission, are not mutually exclusive and may jointly explain the finding. The partial success of the early/late distinction must also be interpreted cautiously: although an ROC-AUC of 0.6976 exceeds the chance level of 0.50, the lower bound of the confidence interval (0.5807) remains close to chance, preventing this result from being regarded as strong evidence of periodization. The contribution is not to “discover” periods through a machine, but to test the periodization claim without providing the model with historical labels and within a framework controlled against composer leakage. Because Feldman’s own mode of argumentation is already based largely on philological and comparative manuscript evidence (2024), this framework aims not to compete with that evidence but to provide a complementary and indirect computational test.

The First *Selâm*: Registered Protocol (Study 6)

Study 6 occupies an epistemic position fundamentally different from those of the other five studies, and this difference must be emphasized clearly here: Studies 1–5 worked with a modern symbolic transmission sample compiled through SymbTr and were therefore able to produce descriptive or inferential results; in Study 6, however, SymbTr is not the data source because the Mevlevi *âyin* repertoire falls outside the scope of the SymbTr corpus. No source-traceable authentic score for any of the *Dügâh*, *Beyati*, or *Saba âyins* is currently present in the confirmatory data package; the number of real scores is zero. Although six published anchors for *Dügâh*, seven for *Beyati*, and one for *Saba* have been verified, these are merely checkpoints used to monitor the data-acquisition process and do not possess the capacity to generate results. The purpose of this section is not to issue a judgment for or against the theory; the only honest assessment that can be offered is not that Feldman’s First *Selâm* thesis has been supported or refuted, but that it has not yet been made observable.

This situation does not indicate a weakness in the thesis. Feldman’s observations concerning the section lengths, numbers of *devr-i revân* cycles, and certain *makam* movements in the *Dügâh* and *Beyati âyins* are based on printed and traceable prose evidence. Yet the gap between prose evidence and machine-readable analytical input exposes a structural problem in computational Ottoman musicology. An observation by Feldman such as “a modulation to *Segâh* in the third cycle of *Dügâh* Bend III” can be used to verify a specific passage, but it does not transform the entire First *Selâm* of the work into an event sequence. The assessment of Bozkurt et al. (2014) remains applicable here: the computational infrastructure for *makam* music analysis is based largely on modern repertoire for which audio or notated records are available; the incorporation of historical and religious-ceremonial repertoire into this infrastructure requires a separate labour of transcription. The finding of this section is precisely that this labour has not yet been undertaken. The study’s principal contribution lies in refusing to simulate historical theory on incomplete data: the registered-report logic of Nosek et al. (2018) has been carried one step further here by publishing the protocol itself before data collection has been completed. The two layers distinguished by Munafò et al. (2017), software reproducibility and the validity of historical inference, are especially clear in this case: the fact that MakamKit’s self-tests on synthetic files are successful demonstrates only that the software operates as expected; it does not mean that Feldman’s historical claim has been validated.

General Evaluation

When we consider how Feldman’s theoretical legacy emerges from these six tests, the picture that appears negative at first glance actually indirectly confirms the strength of Feldman’s own method. Feldman’s argumentation rests largely on philological and comparative manuscript evidence (2024; 2022). The six studies in this article have attempted to offer not a competing form of evidence, but a complementary and indirect computational test. The fact that most of the results take the form of “not supported” or “limited evidence” demonstrates precisely the limits of this complementarity. The barline in modern notation does not carry the hierarchy of accentual weight within the historical *usul* cycle (Study 3). The final notated pitch is not identical to an expert-labelled cadence (Studies 1 and 2).

Shared *makam-usul-form* membership is neither necessary nor sufficient for *nazire* (Study 4). Even as a coarse reduction of eight periods, four periods do not produce a natural cluster structure (Study 5). This does not mean that computational methods are inapplicable to Feldman's theory. On the contrary, the studies clearly map which measurements carry no trace on this transmission surface, which carry a trace in the opposite direction, as independently observed in the case of cadential convergence in Studies 1 and 2, and which theoretical components remain untestable because a required structural precondition is absent, as in the absence of same-composer control pairs for *nazire* in Study 4 and the absence of real scores for the First *Selâm* in Study 6. This mapping constitutes a concrete contribution that identifies the field's next steps.

Several directions for future research are jointly indicated in the limitations sections of the six studies. First, indicators such as cadential convergence and phrase expansion should be redefined in a manner sensitive to concluding regions. The current measures mechanically treat the final note as the finalis, whereas in some forms the concluding section may end with a suspended cadence. Second, the feature-family ablation identified in Study 5, together with negative-control models, alternative clustering methods, and model-dependence assessments, will clarify the extent to which the weak-to-moderate signals found in this article depend on methodological choices. Third, as proposed in Study 4, expert cadence analysis and blinded expert evaluation will demonstrate how computational rankings can be complemented by historical-musicological judgment. Similarly, an expert-validation pilot addressing the shared limitation of Studies 1 and 2 would independently test the validity of measures such as cadential convergence. Fourth, and perhaps most concretely, satisfying the four data-preparation gates defined in Study 6, namely source-traceable scores for three Mevlevi *âyins*, expert phrase boundaries, independent *makam*-region labels, and double source verification, constitutes a specific agenda for moving Feldman's (2022) First *Selâm* thesis to Stage 2.

Two principles should guide the continuation of this agenda: the registered-report logic that fixes hypotheses and decision rules before the data are examined (Nosek et al., 2018), and a principle of transparency that distinguishes software reproducibility from the validity of historical inference (Munafò et al., 2017). When these two principles are maintained, future studies in the field will be able to evaluate similar claims without obscuring the distinction between "the tests passed" and "the theory is supported." The FAIR principles, which recommend the joint versioning of work identity, source page, transformation log, and derived measurements (Wilkinson et al., 2016), also provide a guiding framework for establishing this type of source-transmission infrastructure.

In conclusion, these six studies have neither falsified Feldman's theories of Ottoman court instrumental music nor directly confirmed them. The findings reflect the limited projection carried by a modern symbolic transmission sample, the final link in a chain extending from *meşk* to notation, from notation to edition, and from edition to a digital corpus, as measured through the selected proxies. At times, this projection even points in the direction opposite to that expected. This is not a negative outcome for the historiography of Ottoman court music. On the contrary, it records a chain of tests that honestly confronts the grand narratives of historical ethnomusicology with the measurement instruments of computational ethnomusicology and does not conceal the points at which those instruments remain silent. Feldman's legacy emerges from this picture as demonstrating that his philological method remains indispensable and that computational testing can advance only together with this method, not in its place but alongside it.

Limitations

Rather than being repeated individually, the limitations of the six studies should be consolidated here into a common structure arising from three principal sources.

Transmission and representation limitations. In all studies, the data consist not of original historical manuscripts but of the final link in a chain extending from *meşk* to notation, from notation to edition, and from edition to the digital representation provided by SymbTr. Every stage of this chain involves selection, and the findings can therefore speak only about the product of that transmission chain. In addition to this general limitation, the 53-comma Arel-Ezgi-Uzdilek (AEU) system underlying MakamKit's pitch measurements is a theoretical framework developed in the

twentieth century. The retrospective application of this framework to earlier works constitutes a distinct and recurring risk of anachronism across all six studies (Studies 1, 2, 3, 4, 5, and 6). Composer attribution is likewise not identical to the date of composition: period layers are based on composer attributions, anonymous works were excluded, and the later attribution practices transmitted through the *meşk* tradition could not be verified. This uncertainty may blur the chronological axis (Studies 1, 3, and 4).

Sample size and imbalance. The early-period cells are systematically small, comprising 5–11 works before 1650 and during 1650–1750 in Study 1, $n = 14$ and $n = 18$ in the early periods of Study 2, 14, 18, and 21 works in Study 3, and 14 and 18 works in Study 5. Power analyses indicate that these cells can reliably detect only medium-to-large effects (approximately $|\rho| \approx 0.69$ and $|\rho| \approx 0.62$ in Study 2, and roughly $|\rho| \gtrsim 0.23$ in Study 3). Consequently, results reported as “not significant” should not be interpreted as evidence of “no effect”; a small-to-moderate historical transformation may simply have escaped detection because of insufficient statistical power. The sample is also imbalanced with respect to period and *makam* composition. In Study 2, approximately two-thirds of the works belong to the post-1850 period, while in Study 4, 88.2% of the corpus consists of *saz semâîsis* in *Aksak Semai usul*. These imbalances introduce the risk of compositional confounding in pooled analyses.

Validity of measurement proxies. None of the indicators employed in the six studies directly measures the theoretical concepts intended by Feldman; all are indirect proxies derived from notation. The cadential-convergence measure (Studies 1 and 2) does not distinguish concluding regions and mechanically treats the final note as the finalis, implying that even the most robust finding requires remeasurement using cadence-sensitive definitions. Phrase expansion (Studies 1 and 3) is not normalized by *usul* and is constrained by sectional markings in the notation rather than by breathing or accentuation in performance. In Study 3, because the source data contain no encoding of *düm-tek* accent intensity, an event-phase proxy had to be used, and none of the indicators directly measures the accentual hierarchy of the *usul*. Moreover, because all *saz semâîsis* are in *Aksak Semai usul*, the effects of form and *usul* cannot be disentangled. In Study 4, shared *makam-usul-form* membership is neither a necessary nor a sufficient condition for *nazire*; the cell filter excludes genuine *nazires* that were transformed into different *usuls* or forms, similarity measures cannot distinguish the influence of a common copyist from the compositional style of the composer, and the 34 pairs derived from 48 works do not constitute independent observations. In Study 5, feature-family ablation, negative-control models, and model-dependence assessment through alternative clustering methods have not yet been conducted. Furthermore, because the composer-blocked permutation test kept the clustering labels and classification predictions fixed, the reported *p*-values may be optimistic relative to a stricter procedure in which the model is retrained during every permutation.

Lack of validation. None of the studies included a preplanned pilot investigation examining the extent to which the measurement proxies correspond to expert judgment. This constitutes the principal methodological limitation common to all six studies. In Study 1, the mixed-effects model failed to converge because of a software-layer error and therefore contributed neither supporting nor contradictory evidence to the three-pattern analysis (pooled, form-specific, and matched). In Study 4, because no historical *nazire* labels exist, neither sensitivity nor precision could be calculated, and the resulting ranking remained merely an inspection priority rather than a validation metric. The limitation of Study 6 is qualitatively different: the missing element is neither statistical power nor proxy validity, but the raw data required for analysis itself. Source-traceable authentic scores, expert phrase boundaries, independent *makam*-region labels, and double source verification all remain unmet prerequisites.

References

- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- 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>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Efron, B., & Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. Chapman & Hall/CRC. <https://doi.org/10.1201/9780429246593>

- Feldman, W. (2022). *From Rumi to the Whirling Dervishes: Music, Poetry, and Mysticism in the Ottoman Empire*. Edinburgh University Press. <https://doi.org/10.3366/edinburgh/9781474491853.001.0001>
- Feldman, W. (2024). *Music of the Ottoman Court: Makam, Composition and the Early Ottoman Instrumental Repertoire*. Brill. <https://doi.org/10.1163/9789004531260>
- Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of Classification*, 2(1), 193–218. <https://doi.org/10.1007/BF01908075>
- 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*, 223–228. <https://dblp.org/rec/conf/ismir/Karaosmanoglu12>
- Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 38(4), 963–974. <https://doi.org/10.2307/2529876>
- Munafò, M. R. ve diğerleri. (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. *PNAS*, 115(11), 2600–2606. <https://doi.org/10.1073/pnas.1708274114>
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1), 43–49. <https://doi.org/10.1109/TASSP.1978.1163055>
- Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72–101. <https://doi.org/10.2307/1412159>
- W3C Music Notation Community Group. (2021). *MusicXML 4.0*. <https://www.w3.org/2021/06/musicxml40/>
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), 80–83. <https://doi.org/10.2307/3001968>
- Wilkinson, M. D. ve diğerleri. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>