

Auditory Streaming Complexity and Renaissance Mass Cycles

FINN UPHAM[1]
McGill University

JULIE CUMMING[2]
McGill University

ABSTRACT: How did Renaissance listeners experience the polyphonic mass ordinary cycle in the soundscape of the church? We hypothesize that the textural differences in complexity between mass movements allowed listeners to track the progress of the service, regardless of intelligibility of the text or sophisticated musical knowledge. Building on the principles of auditory scene analysis, this article introduces the Auditory Streaming Complexity Estimate, a measure to evaluate the blending or separation of each part in polyphony, resulting in a moment-by-moment tally of how many independent streams or sound objects might be heard. When applied to symbolic scores for a corpus of 216 polyphonic mass ordinary cycles composed between c. 1450 and 1600, we show that the Streaming Complexity Estimate captures information distinct from the number of parts in the score or the distribution of voices active through the piece. While composers did not all follow the same relative complexity strategy for mass ordinary movements, there is a robust hierarchy emergent from the corpus as a whole: a shallow V shape with the Credo as the least complex and the Agnus Dei as the most. The streaming complexity of masses also significantly increased over the years represented in this corpus.

Submitted 2020 October 29; accepted 2021 February 17.

Published 2021 June 28; <https://doi.org/10.18061/emr.v15i3-4.7980>

KEYWORDS: *corpus research, Renaissance, polyphony, auditory scene analysis, mass ordinary cycle, texture, counterpoint, Josquin des Prez, Palestrina*

INTRODUCTION

COMPLEX polyphony bloomed in Renaissance church music, as the musical genre of the polyphonic mass ordinary cycle emerged in the early 15th century. While some of these pieces continue to be performed today in concerts and churches, the experience of modern audiences is probably quite different from that of people hearing these works in their original context. How might the mass cycle have been followed and appreciated by those attending services in the 15th and 16th centuries?

Imagine that you are a woman in Bruges in 1510. The annual low mass for your deceased father is to be said by a single priest in a side chapel of St. Donatian's Cathedral. You enter the church, where people are walking around, talking, and saying prayers in front of statues and paintings of local saints. You find the chapel associated with your father's guild (with a painting of the guild's patron saint), where you kneel with some other family members, and hear the priest begin the mass by reciting the text of the Introit in Latin. You are "present with the body and intent with the mind," thinking of your father, with no expectation of understanding the priest's words. But you also hear chant (the Introit) being sung at the high altar, behind the rood screen, and your ear is drawn to the music, as it cuts through the other noises in the church. After a minute or two the chant gives way to polyphony – multiple voices singing unintelligible text, largely on the "ee" vowel (the Kyrie). You let your mind wander and look around at the architecture. Then one voice, singing "Gloria in excelsis Dei" and



more polyphony (the Gloria)– with syllabic Latin text and somehow less complex. Then back to melismatic chant and Latin readings – you turn back and look at your relatives and think about your father. Polyphony is heard again coming from the high altar, not that different from the Gloria, but if you catch the opening “Credo in unum Deum”, you know it is the creed. After it ends, there is no more polyphony for some time; but then you hear “Sanctus,” melismatic polyphony that soars through the church, sounding holy and uplifting, while also very solemn, sometimes with just a few voices, and sometimes with a joyous Hosanna, and you know that the host is being consecrated, and that God is present in the church. Finally, after mumbling along with the Lord’s prayer, you hear Agnus Dei – which somehow is even more complicated, with even more voices praying for mercy (miserere nobis). By this time the mass for your father has ended, and you leave the church, happy to have experienced the mass through the medium of the polyphonic music of the mass cycle.

The service of mass, when sung, was structured by the alternation of chant, spoken texts, and choral polyphony. Given the acoustic conditions of these chapels, without amplification or much effort to broadcast the priest’s actions, the loudest cues to mark the complicated progression through the mass would have been the choral singing. So long as a parishioner could tell what movement of the mass ordinary was being sung, they could have some sense of where in the service they might be. While the movements of the mass ordinary are easily distinguished on a score by the text, it is not clear how well parishioners might hear the words. Might there be other features a non-musician listener might use to recognize a Credo or Agnus Dei?

Composers in the Renaissance were expanding the palette of polyphonic textures. The polyphonic mass ordinary settings of the 15th and 16th centuries included anywhere from three to twelve voices (and occasionally more) but the prominence of each line as a distinct sound source shifts with the notes in and around it. Renaissance polyphony explored contrapuntal textures within and between sections, from a single melody that enters alone, to contrapuntal duos, to five-voice homorhythmic passages, to impenetrable sound masses of eight voices moving independently. This range of textures can be perceived as a form of complexity in the auditory scene: how many independent lines are active, how many sound sources are distinguishable, and how much effort would it require to follow their paths and predict what comes next? Experience with polyphonic mass cycles suggests that this dimension of complexity was a feature explored by composers at the time, either directly or as a consequence of other compositional objectives. This paper articulates the concept of auditory streaming complexity and tests whether it varies meaningfully in relationship to the form, functions, and changing aesthetics of mass music in the church.

Without sound recordings and personal accounts by contemporary listeners, we have chosen to estimate streaming complexity from symbolic representations of these scores. The auditory streaming complexity estimate is calculated as a continuous feature reporting a moment-by-moment tally of how many independent streams or objects might be heard, given the information available. While it is easy to count the number of voices sounding at any given moment, the relationships among the voices influence how many independent streams are heard by a listener, right down to the single-stream unison.

To capture this heard numerosity or complexity, this model evaluates how individual notes contribute to the blending or separation of each part from the rest. While some cues for sound stream merging and segregation have very robust effects on the attentive listener (Bregman, 1994), attention also plays an important role in how a listener interprets the auditory scene (Cusack et al., 2004). The perceptual outcomes of auditory streaming involve interactions of bottom-up processing of sensory data with top-down expectations tuned to the signals detected and to the orientation of the listener’s auditory focus (Shinn-Cunningham et al., 2017). While the exact perceptions of a 15th century European churchgoer cannot be fully modeled today, it is reasonable to expect that much of their low-level sensory processing is similar to those of participants in 20th- and 21st-century psychoacoustic experiments with moderate attentiveness. Rather than representing the absolute number of voices a listener might consciously assess and report, this calculation approximates an average impression of the number of sound sources in the music according to non-musician congregants in a complex and noisy overall soundscape.

The principles of auditory scene analysis have been applied to counterpoint and polyphony before. David Huron’s work on voice leading identified many connections between the cues influencing sound stream integrity and the rules of species counterpoint (Huron, 2001). The relationship between streaming cues and the merging of parts in polyphonic works has also been looked at experimentally with perceptual judgments. Ben Duane asked participants to judge the audible merging of parts in string quartet excerpts

(played in midi) and compared these perceptual accounts to quantifications of stream separation cues discerned from the scores (2013). Duane's results suggested asynchronous note onsets and harmonic dissonance were important factors that decreased the merging of two string parts into a blended stream. Emiliios Cambouropoulos combined similar features to identify voices within polyphonic piano music, combining principles of onset synchrony, pitch comodulation, and stream continuity to calculate which notes would be integrated vertically and horizontally (2008).

This project is focused on capturing the dynamics of perceived numerosity in polyphonic vocal works of the Renaissance. By evaluating the continuous interplay of merging and segregating cues in a corpus of mass ordinary cycles, we computationally estimate the apparent complexity of the music as a scene of interacting auditory objects and consider whether this feature of polyphonic writing changes substantively over time and in ways that might have been relevant to the composers and audiences of the day.

THE RENAISSANCE MASS

The Renaissance mass is a polyphonic setting of five texts (known collectively as the mass ordinary) that were said or sung at every Catholic mass.[3] The liturgy of the mass evolved over the course of the middle ages. Each of the mass ordinary texts came into the mass at a different time and served a different function. Some had one main melody to which the text was sung; others had hundreds of melodies (see Table 1). For most of the middle ages, the five sections of the mass ordinary were not considered as a unit.

In the early 15th century, composers began to write mass ordinary cycles: polyphonic settings of the five texts that had musical features in common. The shared musical features were often derived from another piece of music (chant or polyphony) associated with specific events or liturgical functions. This practice became very popular for a variety of reasons enumerated by Andrew Kirkman.

First and foremost, it could be linked to a specific occasion . . . ; second, the specific message embodied in the cycle as a whole could be brought into contact, in the Sanctus, with the moment of grace ... ; and third, . . . it was extended across the rite, bringing the entire ceremony under the aegis of an all-encompassing plea. (Kirkman, 2010, p. 203).

From c. 1400 to the present, thousands of composers set this set of texts according to a rather strict format, as shown in Table 1 (Fallows, 2012).

The polyphonic mass cycle was at the top of the musical genre hierarchy in the Renaissance. Johannes Tinctoris, author of the first musical dictionary (1476), defines the mass in relation to the two other principal genres according to size:

A cantilena (song) is a small piece which is set to a text on any kind of subject... A motet is a composition of moderate length... The mass is a large composition for which the texts Kyrie, Et in terra (Gloria), Patrem (Credo), Sanctus, and Agnus ... are set for singing by several voices, (Tinctoris, 1963, pp. 40-41).

In his treatise on counterpoint (1477) Tinctoris ranks the genres according to the amount of “variety” – a highly valued feature of many types of art that could be understood as a type of complexity (Luko, 2008).

There is not as much variety in a chanson (song) as in a motet, nor is there as much variety in a motet as in a mass. (Cumming, 1999, p. 42).

In 1510 Paolo Cortese published a handbook on how to be a good Cardinal, in which he discussed the role of music in promoting pleasure, knowledge, and morals. He called the mass “the propitiatory mode” and said that “no one should be included in the number of the most eminent musicians, who is not very conversant with the making of the propitiatory mode.” He also identifies Josquin des Prez as an excellent composer of masses (Pirrota, 1966, pp. 54-5).

The polyphonic mass ordinary cycle in the Renaissance is therefore an ideal genre for a computational study of streaming complexity. It is the largest and most complex musical genre in the Renaissance; there are many surviving mass cycles; and every mass sets the same text, unlike other genres, such as the motet and most secular music. It also had a clear place within the liturgy.

To understand the contrasting features of the individual mass ordinary movements, we must look at the text and the liturgical role of each movement. As you can see in Table 1, the mass ordinary movements are not sung one after another (as in a symphony). Only the Kyrie and Gloria are back-to-back; other elements of the mass (shown in grey: mass proper chants, prayers, readings, etc.) are spoken or sung before, in between, and after the polyphonic mass ordinary movements. Most of the mass ordinary texts are very different from each other in terms of length, structure, and function. Therefore, the Renaissance mass cycle created a unified musical structure out of disparate texts over the course of an extended and complex liturgical ceremony; the musical similarities among the mass ordinary movements created a kind of giant rondo form out of the mass as a whole.

Table 1. *Mass ordinary movements in the context of the mass as a whole (Hiley, 1993; Kirkman, 2010)*

Movement (italic text sung in chant)	Translation	Function in mass	Chant melodies	Text setting	Sections
Part I: The Liturgy of the Word					
Introit mass proper chant					
Kyrie eleison, Christe eleison, Kyrie eleison	Lord have mercy, Christ have mercy. Each phrase sung 3 times.	Litany (prayer for mercy) in Greek	More than 100 melodies	Melismatic, 18 words (counting repeats)	3: Kyrie; Christe; Kyrie
Gloria in excelsis Deo /Et in terra pax	<i>Glory be to God on high</i> /And on earth peace	Hymn of praise to God	Over 50 melodies	Mostly syllabic, 84 words	Normally 2: Et in terra; Qui tollis
Prayer, Epistle and Gospel readings, Gradual and Alleluia mass proper chants; optional sermon					
Credo in unum deum /Patrem omnipotentem	<i>I believe in one God</i> /Father omnipotent	The Nicene Creed: basic statement of Christian belief; in the first person	One main melody	Mostly syllabic, 163 words	2 or 3
Part II: Liturgy of the Eucharist					
Collection of offering; Offertory mass proper chant; prayer over offerings; Eucharistic prayer					
Sanctus , sanctus, sanctus, Dominus Deus sabaoth	Holy ... is the Lord of Hosts	Acclamation of praise; end of the Eucharistic prayer (said, then sung during the process of consecrating of the Host)	Over 270 melodies	Melismatic, 25 words	5: Sanctus; Pleni (fewer voices); Hosanna; Benedictus (fewer voices, sometimes sung during transubstantiation) Hosanna
Lord's prayer					
Agnus Dei , qui tollis peccata mundi, miserere nobis (2x); Agnus Dei ... dona nobis pacem	Lamb of God, who takes away the sins of the world, have mercy... give us peace	Litany; sometimes sung when priest or congregation take communion	100s of melodies	Melismatic, 25 words	3: Agnus Dei ... miserere; Agnus Dei ... miserere (fewer voices) Agnus Dei; dona nobis pacem (more voices).
Communion mass proper chant, prayer, dismissal					

The nature and topic of the different texts of the mass ordinary movements responded to the structure of the mass as a whole. Part I of the mass is known as the liturgy of the word because the primary focus is on central Christian texts: the Epistles and Gospels from the New Testament. The mass ordinary movements exemplify the basic devotional acts of a member of the Christian community: the Kyrie is a prayer to God for mercy, the Gloria is an outpouring of praise to God, and the Credo is the articulation of the basic beliefs of a Christian. While the congregation did not sing during mass in the Renaissance, the church choir represented the voice of the congregation.

Part II, known as the liturgy of the Eucharist, is focused on the supreme moment in the mass: transubstantiation, where the bread and wine become the body and blood of Christ (the eucharist), as prescribed by Jesus at the last supper. During the Offertory chant of the mass proper the bread and wine were prepared and placed on the altar. The priest said a prayer over the bread and wine, and he then started the Eucharistic prayer, a multi-part prayer that begins with a dialogue with the congregation and ends with the Sanctus text of the mass ordinary. It is during the Sanctus (and sometimes during the Benedictus section of the Sanctus) that transubstantiation happens. Then the Lord's prayer, "Pater noster," is said by all. During the Agnus Dei movement of the mass ordinary the priest (and sometimes the congregation, who normally only took communion at Easter) took communion. The "Lamb of God" addressed in the Agnus Dei is the communion: the body of Christ. This final movement of the mass ordinary is also a prayer for mercy, and it echoes the structure and message of the Kyrie: one more proper chant, a prayer, and a dismissal round out the ceremony.

Laypeople's Experience of the Mass in the Renaissance

Daniele Filippi has recently investigated the practice of going to mass in the Renaissance. Lay Christians were required by Church law to go "hear the entire mass on Sundays." It is important to remember that the entire mass text was in Latin, the church was crowded, and there were no written supports, such as an order of service, hymnal, or prayerbook. There were multiple "methods of hearing the mass" that depended on the listener's education and inclination (Filippi, 2017, p. 13). Some modern churchgoers assume that the norm was "informed participation," which presumes that the liturgy is a form of communication. But this method was probably limited to people who knew Latin or had been taught in detail about the various sections of the mass. "Devotional participation," which assumes that there was no attempt to understand what was happening during the mass, was probably the dominant mode (Filippi, 2017, p. 17). Here the liturgy is understood as a set of actions; being present in a devout frame of mind is enough. Georg Gobat, a 17th-century Jesuit who summarized many earlier writers, said that "hearing the mass does not mean to understand the words of the mass, which ordinary people cannot do ... but rather "to be present with the body and intent with the mind" (Filippi, 2017, p. 14). Being in the same space with the real presence of Christ in the Eucharist was understood to be beneficial to the body and soul (Filippi, 2017).

Polyphonic mass ordinary settings were only sung at "high" masses, normally sung at the high altar behind the choir screen (and thus invisible to the congregation). The music could be heard throughout the whole church. High masses would have been sung on Sundays and other important feast days (saints' days, etc.). "Low masses" were spoken (with no music) or chanted by a single priest inside chapels. Robert Nosow has documented the schedule of high and low masses at St. Donation of Bruges in the 15th century (the practices in Bruges also applied to many other important churches) (2019). He shows that at any one time there could have been three or more masses being performed simultaneously in the church: one with polyphonic music, two others without music. Therefore, the polyphonic music of the high mass could also be part of the soundscape of the low mass.

Historical Perception and Cognition of the Polyphonic Mass Ordinary

What might a 15th-century churchgoer who practiced devotional participation perceive or experience when attending mass in a Renaissance church? Polyphonic music sounded elaborate, festive, and special. The alternation between the complex music of the mass ordinary cycle (where each movement begins with similar music) and the chant and spoken text of the rest of the mass could have provided an overall unity to the ceremony. The standardized sequence of the five movements, each with a different approach to text-setting and structure, might have allowed listeners to locate themselves in the ceremonial time of the mass, even if they had no idea what was being said. The melismatic Kyrie and Agnus Dei, both of which have three

sections, serve as a kind of internal frame for the mass as a whole. The long syllabic Gloria and Credo movements articulate the liturgy of the word. The melismatic Sanctus and Agnus frame the mystery of transubstantiation. In a complex soundscape like the one described in the introductory vignette, the music of the polyphonic mass ordinary provided an aural focus for being “present with the body and intent with the mind.” In this context, different streaming complexity patterns could have helped locate the listener in time and in the spiritual world of the liturgy.

Auditory Streaming Complexity Estimate

The organization of streams in an auditory scene depends on the characteristics of the active sound sources and the timing and quality of the sounds created. In polyphonic masses of the Renaissance, we know some things about the sound sources. These pieces were performed by groups of trained singers, often one to a part, and usually grouped close together in the worship space and at some distance from the parishioners. Renaissance composers of masses were always also professional singers and might have been among the choir singing the mass cycle. As described above, this music would also happen at the same time as other audible worship activities, rather than in silence as is common today. Spatial positioning and competition with non-musical sounds would encourage harmonious singing to be heard as a single source or blended stream, while the timbral distinctness of individual voices might help distinguish parts into separable sources. Church music at this time also included unison chant, in which the coordination of singers would favor a single-stream hearing. Polyphonic music emerged in contrast to this aural unity, with an emphasis on counterpoint resulting in the independence of voices in pitch, rhythm, and text. This independence comes from the cues in each part that directly contradict the default unity of monophonic chant: onsets against held notes, contrary motion, and independent entries. These cues make up the evidence for separate sound sources.

Polyphonic music of the Renaissance seems to play with both merging and separation of voices, rather than maximizing the independence of parts. Within a single mass movement, the number of voices active (i.e., sounding) at one time will range from one to four, or six, or even twelve, and the concentration of separation cues will also change over time. The independence of sources in actively contrapuntal sections can create a complex texture that contrasts with the smooth consistency of homorhythmic passages using just as many or even more voices. Some contributing factors to the separation of sound sources in Renaissance mass cycles are determined in performance. Still, much of the information contributing to the perceived interactions of parts is in the score. The Auditory Streaming Complexity Estimate is a descriptor calculated directly from the notes in python using code libraries music21 (Cuthbert, 2019) and pandas (McKinney, 2010). This section explains how the Auditory Streaming Complexity Estimate, or streaming complexity for short, is calculated with an example from Josquin's *Missa Fortuna desperata*, shown in modern music notation in Figure 1 and as pitch over time in Figure 2.

The figure displays a musical score for four voices: Superius, Altus, Tenor, and Bassus. The Superius part is marked 'Section: Christe' and 'MenCutC'. The Altus part is marked 'MenCutC'. The Tenor part is marked 'MenC'. The Bassus part is marked 'MenCutC'. The score shows a complex polyphonic texture with various rhythmic and melodic patterns across the four parts.

Figure 1. An excerpt from the *Kyrie* movement of Josquin's *Missa Fortuna desperata*. The excerpt is taken from the beginning of the *Christe* section (the middle section of the *Kyrie*).

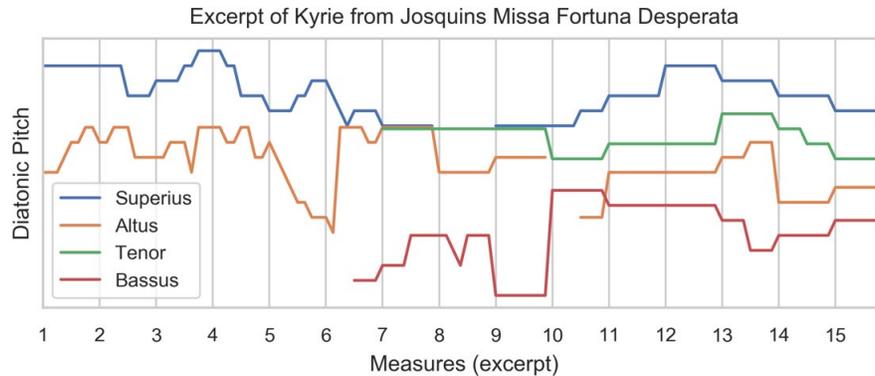


Figure 2. Same excerpt as in Figure 1, plotted as diatonic pitch over time.

Separation Cues

The first task for calculating this Streaming Complexity Estimate is to isolate cues that could contribute to the separation of sources. From a symbolic music file, we identified three types of events:

1. Voice Entries: when a part enters for the first time or after a rest.
2. Note Onsets: the beginning of each articulated note in the score.
3. Pitch Changes: when a part changes pitch and in which direction.

These events were chosen as reliable cues for stream formation in this genre of music. For example, note offsets were specifically not evaluated because the sustained singing style typical of this music renders these mostly redundant to note onsets. And while Duane found that inharmonicity in vertical intervals discourages blend between parts in classical string quartets (2013), such intervals are less common in our corpus of Renaissance masses, and the perception of dissonance is too culture-dependent for us to confidently extrapolate contemporary impressions from measurements on modern listeners. The three event types we use may not capture all the musical information influencing the perception of streaming complexity. Still, they are important, relevant to compositional strategies of the day, and can be reliably extracted from the symbolic scores.

The effect of these three event types on stream separation depends on how they co-occur. Two voices that enter on the same beat are easily heard as an integrated source until one moves against the other. Shared events encourage merging and must be filtered out to isolate the cues that encourage separation. Figure 3 shows the incidence of the three event types in each part of the Josquin excerpt found in Figures 1 and 2. Voice entries are sparse, occurring in only a few measures of this excerpt, while note onsets are numerous, and pitch changes are nearly as frequent and can go up or down. In order to filter out merging cues, we consider when these events co-occur with the same event type in another part. Simultaneous entries, onsets, and pitch changes subsumed by a higher part are plotted in grey (pitch change direction is shown by triangles pointing up or down). For example, at measure 11, all four parts articulate a new note, sharing a single onset as three of the four change pitch in the same direction. When calculating the reduction of events to separation cues alone, we attribute the "source" event to the highest voice (for many people, higher voices are more salient), but the same reduction could be performed prioritizing other lines without changing the final estimate values.

Cue Integration per Part

Next, these different types of separation cues are merged per voice to describe each line's distinction against the whole. This is complicated as the three types of event are related and not interchangeable: entries and pitch change cues always involve note onsets but not vice versa. It is also necessary to consider the relative power of these cues. A voice entry is more distinguishing than the re-articulation of a note on the same pitch, but we do not have precise empirically-grounded weightings with which to aggregate their effects. As an initial foray, we gave different relative weights to each cue type, 1 to voice entries, 0.5 to pitch changes, and

0.25 for onsets. Furthermore, when different cue types co-occur in the same voice, the larger cue weight is retained rather than the sum.

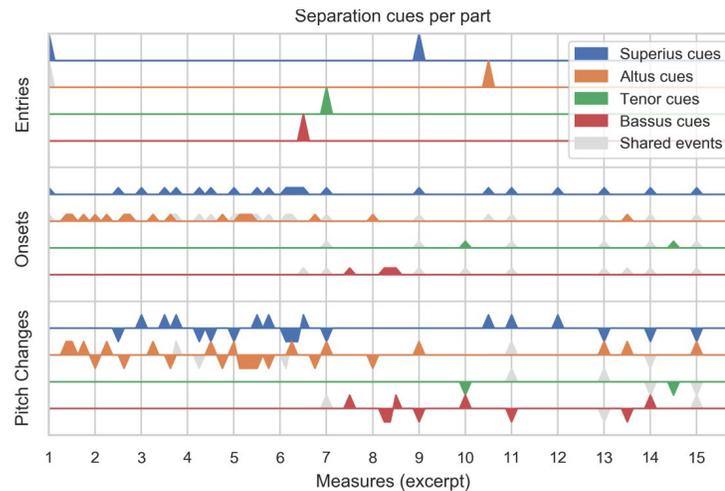


Figure 3. Individual events and separation cues from each part of the excerpt in Figure 1: Entries, Onsets, and Pitch Changes. Events are grey when the same type of event occurs simultaneously in a higher voice, encouraging the merging of streams rather than their separations.

Auditory streaming is a little bit sticky: once sounds in the auditory field have been attributed to separate sources, our ears will try to hold on to that organization until contradictory cues overwhelm it (Bregman, 1994). In this music, we assume there is encouragement in the voice's spectral qualities and performance context to hear a choir as a single source, at least for contemporary listeners. However, separation cues will have some duration of effect against that inclination. Given how rapidly the ear can reassess soundscapes (Shinn-Cunningham et al., 2017), we define a short linear decay for these cues, reducing at a rate of 0.2 per semi-minim or quarter note. In practice, a voice entry would give a part distinction as a separate stream that would dissipate within 1-1.5 seconds, depending on performance tempo.

In Figure 4, the separation cues in each part have been integrated across types and extended in time with the linear decay. This allows runs of notes to accumulate distinction, as seen in the Altus line through the first six measures, and for slow passages like measures (mm) 9-12 to drift towards unity with every held breve.

The last essential step of the cue integration process is to cap the streaming separation of each voice. Outside of compound melody techniques (not present in this repertoire), the maximum number of sources represented by a single voice is one, and this hard limit is enforced in our calculations.

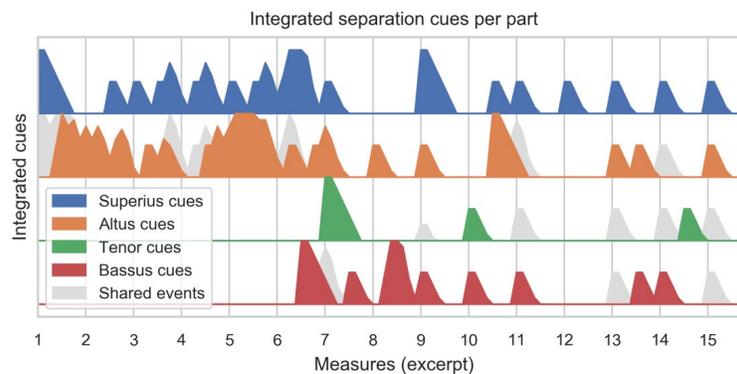


Figure 4. Separation cues integrated per part for the same excerpt as Figure 1. In color, the accumulations of cues separating each line, given its relationships to the other parts. In grey are the potential contributions of events shared with higher voices, cues that could add to separation in a different polyphonic context.

Total Streaming Complexity Estimate

The last step to having a continuous estimate of streaming complexity for a piece of polyphony is to sum the separation cues across voices, with a few more caveats.

The first is the issue of held notes. Renaissance polyphony often involves one or more parts holding a single pitch against which other voices move. These held notes are not terribly salient, but they contribute to the fullness of the texture. To include these in the total Streaming Complexity Estimate, a value of 0.25 is added to the highest sounding voice without separation cues.

To combine these streaming separation cues into a meaningful complexity estimate, a last adjustment must be made: the identification and amplification of the foreground source. Separation cues are only meaningful against an existing stream, and so long as voices are sounding, the Streaming Complexity Estimate must be no less than 1. Before summing for the total, the part with the highest streaming separation value, beat by beat, is raised to a value of 1.

The total auditory Streaming Complexity Estimate for the Josquin excerpt is shown in Figure 5, along with the count of active voices in red. At the beginning of this excerpt, the duo texture wobbles between 1.25 and 1.5 as the Superius and Altus trade prominence. The entry of the other parts in mm 6-7 raises the estimate above 2, but the impacts are fleeting. The slow homorhythmic texture after measure 11 returns to the range of not-quite-2, despite there being now four voices singing at once.

Streaming Complexity vs. Voices Active Count

The Streaming Complexity Estimate is a feature designed to prioritize the unskilled listener's perspective on what is in the music, in a noisy context. We contend that church attendees such as the mourner described in the introduction would have had a limited view of the music heard, interpreting what they heard as more or less complex without the ability to assess the real numbers of singers involved. Other people present, particularly the musicians performing, might have a different impression of the number of auditory streams in the musical soundscape.

Imagine you are a singer at St. Donation's Cathedral in Bruges. Your job is to sing chant and polyphony in all the high masses, and also to sing both chant and polyphony in some of the offices, especially Vespers and Compline. You sing with a small choir of professional singers (1-3 on a part). You can improvise polyphony against a plainsong, and you do some composing as well. You have been singing most of your life; you began as a choirboy and moved around until you settled in this church, where you will probably sing for most of your life; you have sung with this particular group of singers for years and know them very well. When you sing at mass, you all stand close together around a large choirbook (perhaps 50 to 70 centimeters high, 20 to 30 centimeters wide when closed), where you can see all the parts laid out – not in score, but with each voice copied on a different part of the page (you are a bass, so your part is the bottom right corner, and you stand next to the high lectern that holds the book, and are in charge of turning the pages).

For each movement of the mass, you can see exactly how many parts are involved in any one section; you can also hear the voices of your fellow singers, and you recognize the exact timbre and mannerisms of each one. You tune to the other singer on your part and to the singers on other parts – so you are aware of how many voices are active at any moment (and who is not singing). You are also aware of musical interactions between the parts – you can see and hear the long-note tenor cantus firmus, hold fast to your part when there are conflicting rhythms, and when you have homorhythmic parallel thirds with another part, you work to align the rhythms and articulation. (Cumming and Tribble, 2019).

The real number of sound sources in these masses is easy to evaluate from the perspective of the performer: each score includes a specific number of parts. A more granular measurement is the number of parts sounding at any given moment, what we refer to here as the number of voices active. In this genre of music, the number of voices active is an absolute upper limit on the Streaming Complexity Estimate values, and more parts allow for higher possible streaming complexity values. To allow for the lingering of streams, our calculation of voices active has been smoothed with a fast decay, like the streaming complexity measure.

The number of voices active and Streaming Complexity Estimate are not interchangeable measures. Figure 5 reports streaming complexity values in a two-voice section (mm. 5-6) that exceed many of those in the following four-voice texture (mm. 11.5-13). To understand the relationship between these two perspectives (that of the performer in active voices, and that of the unskilled listener in streaming complexity), we use both to evaluate patterns in the corpus of masses. Firstly, to determine if streaming complexity is capturing any information beyond that described by the voices-active measure, and secondly, to see whether composers were directly or indirectly choosing to write different levels of streaming complexity that are distinct from the number of parts heard in each moment.

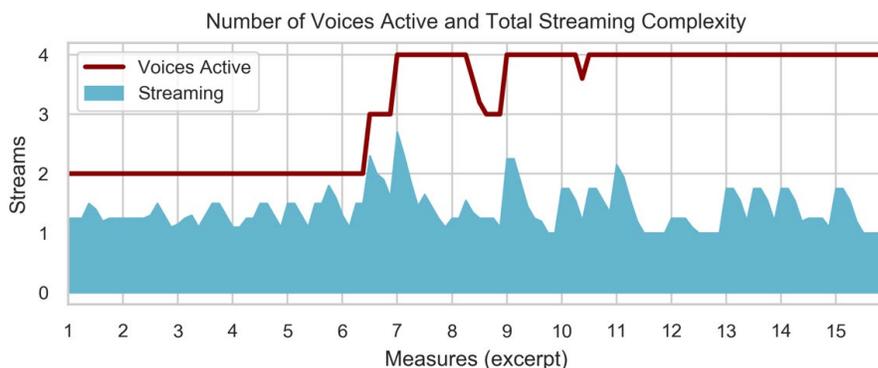


Figure 5. Estimated continuous streaming complexity for the excerpt in Figure 1 (light blue), with the number of voices active by beat (dark red line).

THE EXPERIMENT

To assess the relevance of streaming complexity as a feature of polyphonic music in the Renaissance, and to explore complexity in relation to the liturgical role of the mass ordinary, we aggregated symbolic encodings of Renaissance masses and tested for patterns in their statistics.

The Masses

The symbolic music files came from two public collections of polyphonic Renaissance music: The Josquin Research Project (JRP), an open-access database of Renaissance scores in multiple symbolic formats curated by the strength of composer attribution (Rodin & Sapp, 2019), and RenComp7, a repository of MIDI scores with additional polyphonic vocal works of the Renaissance (McKay et al., 2017). From the RenComp7 dataset, we took 18 mass ordinary cycles by Tomás Luis de Victoria and the 101 by Giovanni Pierluigi da Palestrina. From the JRP we gathered the remaining 84 mass ordinary cycles in kern format by 13 identified composers and a few without clear attribution. In total, our corpus includes 216 masses by at least 15 different composers (Table 2).

Dates of composers and especially of individual works is very tricky in the Renaissance. We rarely know the birth dates of composers; they were not famous when they were born, and people's ages are rarely mentioned. Before 1502, masses were found only in manuscripts, for which we almost never have precise dates. Starting in 1502 (with the first printed book of masses, Petrucci's *Missae Josquin*), we do get publication dates for masses. But the copying date and the publication date only tell us the last possible date for the composition of a mass since it could have been composed decades earlier. We therefore divided the repertoire into quarter-century chunks (eras). Cumming assigned each mass to a quartile based on what is known about the composers' lives and the dates of the prints and manuscripts in which they are found, using composer articles from Grove Online and publications on Josquin (Fallows, 2009) and Palestrina (Marvin, 2002). In the case of Josquin, many works were attributed to him in later sources, so we divided his masses into two categories: strong attribution and weak attribution, as determined by Jesse Rodin in the JRP (2019).

Table 2. *Composers and dates of masses in corpus*

Composer name	Composer dates	No. of masses	Quartile
Du Fay, Guillaume	1397–1474	4	1450-1475
Ockeghem, Jean de	c. 1410–1497	11	1450-1475
Busnoys, Antoine	c. 1430–1492	1	1450-1475
Anon. Missa <i>Gross Senen</i>	c. 1450–1475	1	1450-1475
Martini, Johannes	c. 1430–1497	8	1476-1500
Tinctoris, Johannes	c. 1430–1511	2	1476-1500
Anon. Missa <i>L'ardent desir</i>	c. 1475-1500	1	1476-1500
Agricola, Alexander	c. 1445–1506	2	1476-1500
Obrecht, Jacob	c. 1457–1505	3	1476-1500
Pipelare, Mattheaus	c. 1450–c.1515	1	1476-1500
Orto, Marbrianus de	c. 1460–1529	5	1476-1500
Févin, Antoine de	c. 1470–c.1512	1	1476-1500
Josquin des Prez (strong attribution, JRP)	c. 1450–c.1521	4	1476-1500
Josquin des Prez (strong attribution, JRP)	c. 1450–c.1521	8	1501-1525
Josquin des Prez (weak attribution, JRP)	c. 1450–c.1525	1	1450-1475
Josquin des Prez (weak attribution, JRP)	c. 1450–c.1525	1	1476-1500
Josquin des Prez (weak attribution, JRP)	c. 1450–c.1525	15	1501-1525
La Rue, Pierre de	c. 1452–1518	26	1501-1525
Anon. Missa <i>Bergerette savoysienne</i>	c. 1500–1525	1	1501-1525
Daser, Ludwig	c.1525–1589	1	1551-1575
Palestrina, Giovanni Pierluigi da	1525–1594	21	1551-1575
Palestrina, Giovanni Pierluigi da	1525–1594	80	1576-1600
Victoria, Tomás Luis de	1548–1611	18	1576-1600
Total		216	

Most of the composers seem to have written most of their masses in one quartile; others (Josquin, both strong and weak attribution, and Palestrina) composed over longer periods, so we have divided their masses into the different quartiles in Table 2. There are no masses in this corpus from between c.1525 and Palestrina's first masses (published in the 1550s), because we found no symbolic files for masses from this period. All these masses were conceived as cycles by their composers and presented in their sources as mass cycles. The specific techniques of unification are not relevant to this study.

Data Cleaning

Nearly all of these masses included the typical five movements of the mass ordinary: Kyrie, Gloria, Credo, Sanctus, and Agnus Dei. While files of mass proper movements such as the Introit or Communion were sometimes available for specific cycles, these were not included in the dataset. There were also a few masses that were missing symbolic files of some movements; these cycles were retained where statistical tests allowed.

Masses from later periods were often further broken down into six or seven or eight movements, with first and second parts of the Agnus, for example, and the Benedictus separated from the first part of the Sanctus. For the sake of analysis, these components were concatenated to match the standard five movements, all sharing the same text and liturgical role. A few more movements were excluded because the files available appeared to only contain one or two parts, too few for this style and these calculations. Across all 216 masses, 1042 movements were used for the analysis.

The Statistics

The experiments were run on statistics describing individual movements. The voices active and streaming complexity time series were reduced to their respective average values for each of the 1042 movements. These statistics, as shown in Figure 6, are better interpreted as the center of mass of the distribution than representative of the most common values in these time series.

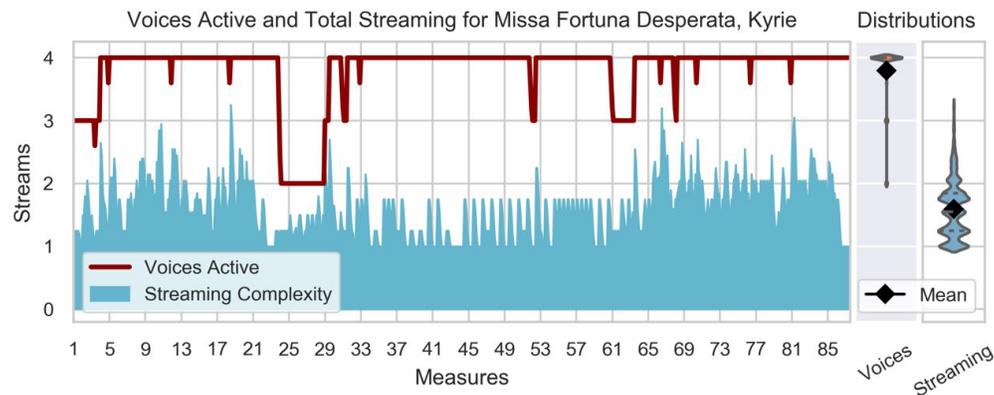


Figure 6. To the left are the streaming complexity and active voices time series for the complete *Kyrie* movement of Josquin's *Missa Fortuna desperata*. To the right are the distributions of these two measures in violin plots and their average values (black diamonds): first voices active in red, then streaming complexity in blue.

A secondary statistic was calculated from these averages, normalizing to the average value per mass cycle, to facilitate comparison between movements: the relative average number of voices active (relative voices active), and the relative average streaming complexity estimate (relative streaming complexity). Independent variables included the movement category (*Kyrie*, *Gloria*, etc.), the composer if known, and the quarter-century in which the mass was composed.

Hypotheses

With the information in this corpus of Renaissance masses, we test the following hypotheses:

- The Streaming Complexity Estimate captures information distinct from the number of parts in the score or the distribution of voices active through the piece.
- Streaming complexity of masses increased over the years represented in this corpus.
- There is a robust hierarchy across mass movements in terms of which are most complex.
- Composers did not all follow the same relative complexity hierarchy for mass ordinary movements, according to this measure.

RESULTS

Parts, Active Voices, and Streaming Complexity

As expected, there is a strong correlation between the number of parts in a Renaissance mass and the average streaming complexity. Spearman's Rho correlation between the number of parts in the movement and the average streaming complexity reached $r_s(1040) = .73$, for $p < .001$ ($p < 10^{-175}$). As shown in Figure 7, the number of parts in a mass movement does affect the range of mean streaming complexity, but these ranges overlap a great deal.

The average number of voices active across each movement correlates more strongly with the mean streaming complexity measure, up to $r_s(1040) = .81$. Controlling for the number of parts per movement lowers this correlation substantially: in four parts $r_s(544) = .57$, for $p < .001$ ($p < 10^{-47}$), in five parts $r_s(272) = .43$, for $p < .001$ ($p < 10^{-14}$).

These descriptors of numerosity in polyphony evidently share a great deal of information, but there is some distinction between the number of parts, the average number of voices active, and the estimated streaming complexity. This allows us to consider other explanations for the variation in streaming complexity values.

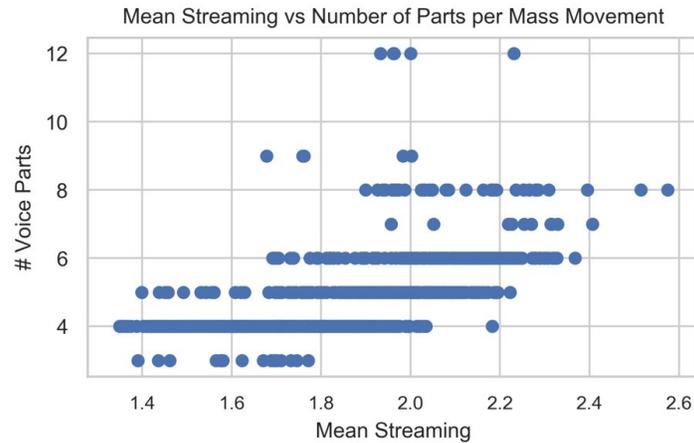


Figure 7. A scatter plot of the number of parts (maximum active voices) and the average (mean) streaming complexity for all 1042 movements.

Chronology

The masses in this corpus were written within 150 years and can be placed in a rough chronological order (as shown in the quartiles listed in Table 2). Figure 8 shows the distributions of average voices active and Streaming Complexity Estimates for movements in each quarter-century, showing a change in range over time. The early years mostly featured two and three voice textures within movements of four or more parts, while after 1550, more works featured four or more voices active at once. A similar trend can be seen in the mean streaming complexity measures, with textures tending towards two or more differentiable streams after 1550.

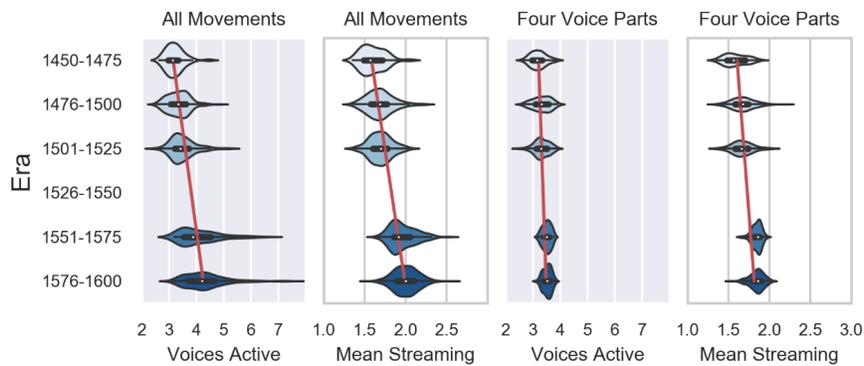


Figure 8. Violin plots of the distribution of average Voices Active (grey background) and Streaming Complexity (white background) in each identified Era. Red lines report linear regression trend over time on a random subset of 50 movements from each era. The first pair of plots report the distributions and trends across all movements in the corpus, and the last two describe only those movements with four voice parts.

The linear regression between era and stream values were applied to a subset of movements in this corpus, 50 from each era. The first two plots of Figure 8 show the distributions of values per era in violin plots and the regression line on one of these random subsets, as calculated with scipy (Jones et al., 2001). The trend to have more voices active was strong: from 100 random samplings, the regression on voices active

show significant change $F(4,246) = 36.200, p < .001, R^2 = .358$. However, the linear trend was stronger for the mean Streaming Complexity Estimate, $F(4,246) = 70.494, p < .001, R^2 = .51$.

To check on whether this trend was an artifact of the trend to include greater numbers of parts over time, we ran this same analysis on mass movements with only four parts. The last two plots of Figure 8 show the more restricted distributions of streaming values on these works, with regression lines calculated on random sample subsets of 50 per quarter-century, as before. The increase in voices active is more restricted here, but still significant: $F(4,246) = 14.387, p < .001, R^2 = .191$. But again, the linear increase in streaming complexity is more pronounced: $F(4,246) = 47.024, p < .001, R^2 = .419$. This pattern of increasing streaming complexity is also visible when comparing the outputs of individual composers. Figure 9 shows the distributions of mean streaming complexity for each of the identified composers in this corpus, roughly ordered chronologically (Josquin works only include those masses with strong attribution). In addition to the increase in complexity, these composers varied in the range of textures they produced, within and between mass cycles.

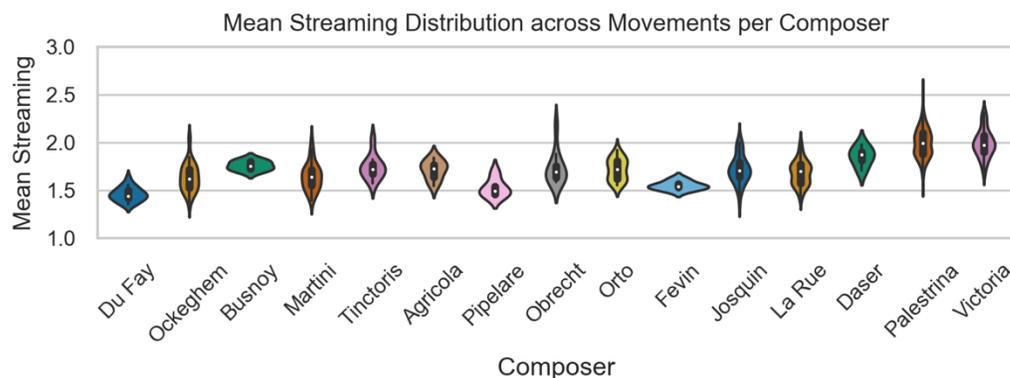


Figure 9. Violin distribution plots of the mean Streaming Complexity Estimates for movements by each composer in the corpus. Composers ordered roughly chronologically by mass composition dates.

Some of the increase in complexity over time may most easily be attributed to the composers represented in each quarter-century set in this corpus. As reported in Table 1, the last two quarter-centuries are dominated by Palestrina and Victoria. But the trends are just as clear in the earlier epochs that include a wider variety of contributors. Together this information positions Victoria and Palestrina as composers who continued the trend of increasing complexity established in the preceding century.

Movements

Our initial motivation for quantifying streaming complexity was to consider its relationship to function. Are there discernable and reliable contrasts in the streaming complexity of the different movements in the mass ordinary cycles? As the absolute values of these stream features changed with time (and composer), we normalized values per mass cycle, adjusting the average voices active and streaming complexity per movement by subtracting the mean feature value of its mass cycle. These relative statistics expose which movements tend to be high (above zero) or low (below zero) given the rest of the movements in the mass.

Figure 10 shows the distribution of these relative streaming values per movement. Two-factor ANOVAs, movement and composer, on these relative values across the whole corpus showed a significant effect of movement and a significant interaction between movement and composer for both statistics. These and subsequent ANOVAs were calculated using the python code library *pinguin* (Vallat, 2018).

For streaming complexity, the main effect of movement yielded an F ratio of $F(4,788) = 88.1, p < .001$ (uncorrected $p < 10^{-40}$), indicating a significant difference in relative mean streaming complexity between movements. Significant pairwise differences between most movements were strong ($p < .001$), except for Credo-Gloria and Gloria-Sanctus ($p < .05$) and no significant difference between the Credo and Sanctus (Games-Howell post hoc test, allowing for unequal variance). The significant interaction between composer and movement, $F(64,788) = 3.4, p < .001$ (uncorrected $p < 10^{-15}$), suggests analysis of individual composers' patterns is warranted.

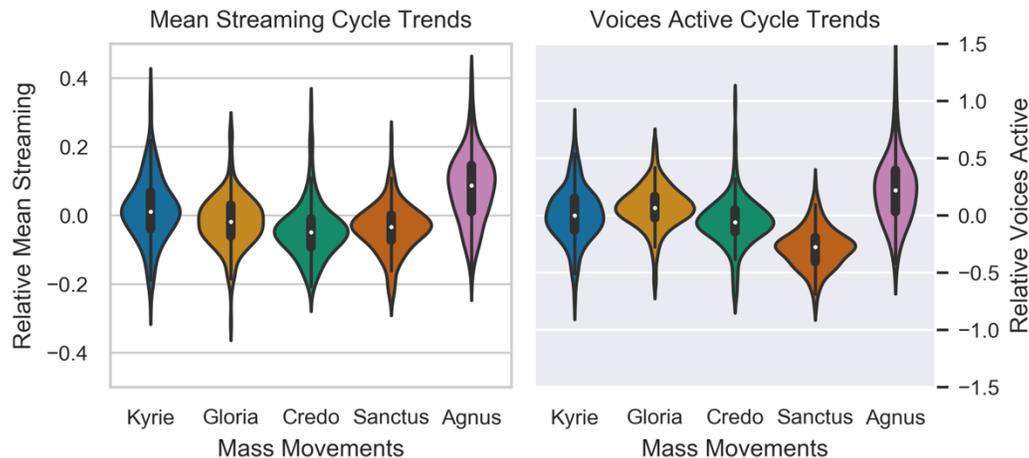


Figure 10. Relative streaming statistics per mass movement. Relative average streaming complexity values in violin plots to the left, with median, quartiles, and distribution. Relative average voices active to the right.

For Voices Active, the main effect of movement yielded an F ratio of $F(4,788) = 188.0$, $p < .001$ (uncorrected $p < 10^{-42}$), indicating a significant difference in relative mean streaming complexity between movements, with all pairwise differences ($p < .001$) except between the Kyrie-Credo and Kyrie-Gloria movements ($p < .05$) (Games-Howell). The interaction between composer and movement was smaller but still significant, $F(64,788) = 8.1$, $p < .001$ (uncorrected $p < 10^{-52}$).

While the relative streaming statistics across mass movements vary widely, with examples of each movement both above and below cycle averages across these 216 masses, there is a substantial tendency for some movements to be more complex and to involve more sounding parts. The Agnus Dei movements tend to be highest according to both measures, however these two measures do not agree on which movement has lowest values. The Sanctus consistently features the lowest average relative to number of voices active, and the Credo falls to the bottom according to the Streaming Complexity Estimate, suggesting a more balanced U shape over the course of the mass cycle.

Composers

The interaction of composer and mass movement for average streaming complexity and voices active values in this corpus suggests that it is worth looking for patterns in individual composers' mass cycles. Palestrina contributed the largest subset of the corpus with 101 masses. The range in relative streaming complexity is very pronounced in Palestrina's works, with all pairwise comparisons significant, save between the Gloria and Sanctus (shown in Figure 11). The V-shape trend in the streaming complexity of the movements contrasts with the relative mean Voices Active distributions per movement, for which the only pair without significant difference was the Kyrie and Credo. Figure 13 shows the distributions of these measures per movement for a single mass, showing the same hierarchy as described in the average trends in Figure 11. These movements overlap a great deal in the ranges of streaming complexity values and voices active: these features are not likely to help a listener identify the movement from hearing a single phrase. The difference is instead in the balance of relatively high and low values over time. If a listener perceives an increase in accumulated complexity, this suggests the mass is nearing the end. While not all masses by Palestrina strictly follow the hierarchies visible across the full set, it is a strong trend. The hierarchy of movements in both measures are very similar to those of the masses by Pierre de la Rue and Victoria, ranking the outer movements as relatively high, Credo as distinctly low for streaming complexity, but with the Sanctus as lowest in relative Voices Active.

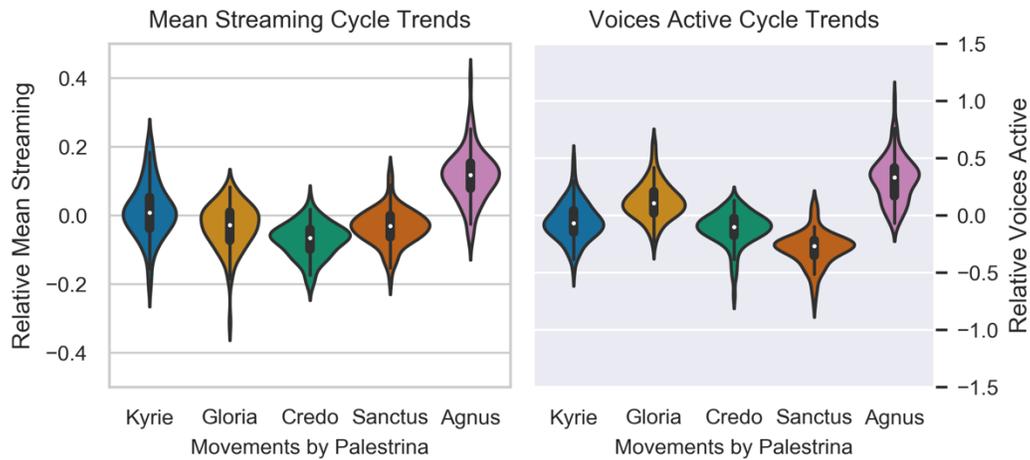


Figure 11. Cycle trends in Relative Mean Streaming Complexity and Voices Active for 101 masses by Palestrina.

A set of masses with a clear contrast to this dominant pattern are those strongly attributed to Josquin des Prez (Figure 12, upper plots). In these mass cycles, it was more common to rise to the Credo from a less complex Kyrie, fall in the Sanctus and then rise again in the Agnus (as is common for all masses). From the 12 masses in the set, significant pairwise contrasts between these movements are found between the Kyrie and Credo ($p < .05$), the Credo and Sanctus ($p < .01$), and the Agnus Dei and Sanctus ($p < .01$). In these masses, the hierarchy in streaming complexity follows that of the voices active.

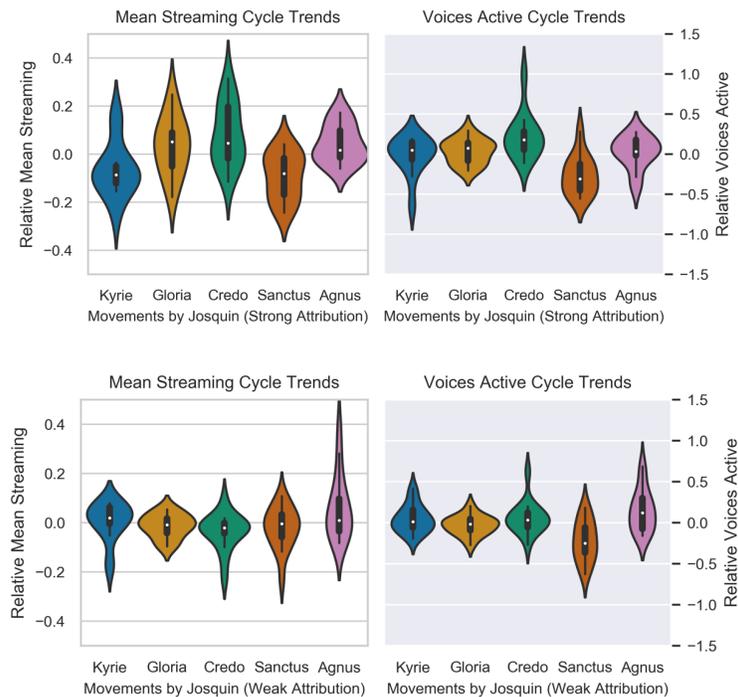


Figure 12. Cycle trends in Relative Mean Streaming Complexity (left) and Relative Voices Active (right) for the 12 masses with secure attributions to Josquin (above) and the 17 masses where the attribution is not secure (below).

The 17 masses where attribution to Josquin de Prez is considered weak show a shape more consistent with the overall trend, with streaming complexity for the Sanctus movement contrasting with the voices-active measurements (Figure 12, lower plots). Testing evaluating attribution security as a factor against

movements, there is a significant interaction with movements for average Streaming Complexity $F(4,104)=3.694$ ($p<.01$) but not for Voices Active $F(4,104)=1.549$ ($p=.19$). This suggests an inconsistency in how streaming complexity is used within the masses of weak attribution, as if these masses were composed by multiple people, rather than all by Josquin des Prez.

DISCUSSION

The Auditory Streaming Complexity Estimate was constructed out of principles of auditory perception, with parameters such as relative cue weightings and decay rate set to reasonable values without optimization. While we hoped to find differences in values over time and between movements, there was no empirical standard to which we could tune the calculation. But even without the opportunity to sharpen the model, this measure appears to pick up distinct and meaningful patterns in this corpus of Renaissance masses. Average values per movement strongly correlate with the number of parts in the score and the average number of voices active over time. However, this streaming complexity measure increased more robustly with composition date than with the other descriptors, at least within the 150 years represented in this corpus. The different hierarchies for voices active and streaming complexity averages across the mass cycle also underline how these measures capture distinguishable aspects of the musical soundscape and mark a difference from the streaming perceivable to the performer and experienced by unskilled listeners in the church. While this Streaming Complexity Estimate can surely be improved and refined, the present calculations already expose promising information in the corpus.

Changes over Time

Masses composed between 1450 and 1600 show a trend of increasing complexity. Most obvious is the increasing popularity of movements with more than four voices (five, six, eight, or 12 parts), reflected indirectly in the increase in average active voices shown in Figure 8. More delicate is the question of whether the change in the number of parts was accompanied by additional changes in polyphonic textures. The Streaming Complexity Estimate averages fit a stronger regression to the five quarter-century sets of works in the corpus, showing a more robust effect of time than did the voices active calculation. And the change was not only a consequence of more parts: a parallel analysis of movements with exactly four parts shows a similarly strong increase in streaming complexity. These results reassure us that polyphony writing was getting more complex in the mass ordinary over time, regardless of the number of parts, and that the trend of increase in streaming complexity is even more consistent than the increase in voices active.

These findings correspond to what musicologists and performers already know about style change in the Renaissance and in the mass in particular. From 1450-1540 most masses were composed for four voices, and around 1540 five-voice writing became much more common (though four-voice writing never stopped). In the late 16th century, there is more experimentation with large numbers of voices and a new interest in polychoral music, where choirs sing separately and together, especially in Victoria's music. Textures also changed over time with an increase in the number of voices active. Between 1480 and 1520, reduced textures were common; movements often included long duo sections and widely spaced imitative entries. After 1520, extended duo sections became rarer, and imitative textures with more densely timed entries started to predominate. When there are fewer rests in all parts, the voices are active (or sounding) for more of the time.

But our findings also provide us with new information that is impossible to detect when looking at individual masses. Not only do we have more voices in the score, and more voices active over time; we now know that the increase in streaming complexity means masses were composed in a way that individuated voices from each other more of the time: less merging, more separation.

Movements of the Mass Ordinary

The relative complexity of the mass ordinary movements within individual masses showed significant trends in terms of the average streaming estimate and the number of voices active. A reasonable question is: how might such differences be audible in individual works? Figure 13 shows the distribution of all values of these measures in each movement in Palestrina's *Missa Iste confessor* with averages marked with black diamonds.

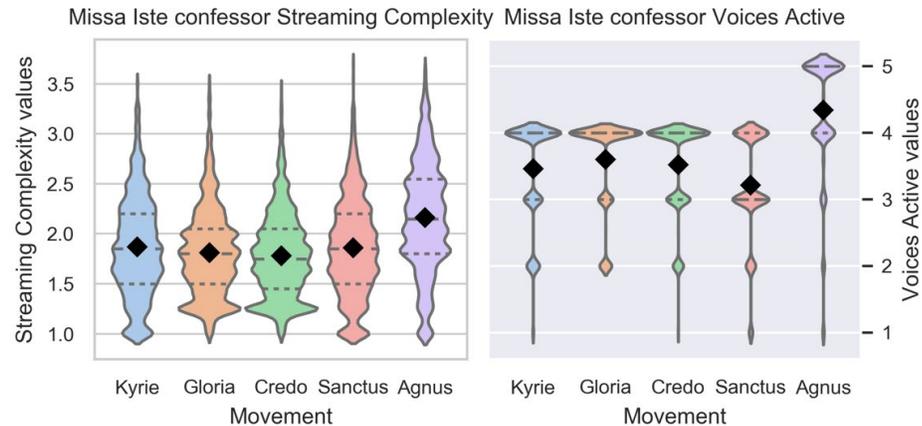


Figure 13. Distributions of Streaming Complexity and Voices Active values for each movement of Palestrina's *Missa Iste confessor*, with movement averages marked by black diamonds, as in the histograms on the right of Figure 6.

The beaded bulges of the Voices Active plot show plainly that for this mass, the Agnus Dei includes five voices a lot of the time, and the Sanctus has a more varied texture, with more three-voice writing than the other movements. In terms of streaming complexity, the distribution bulges follow the suggested V with a great deal of overlap. Here the Credo has the lowest median and more time with values below 1.5 than the other movements, although it is not very different from the distribution in the Gloria. The Sanctus, however, spends substantially more time above values of 2 than the Credo does; it is more similar to the Kyrie, despite the different proportions in voices active.

This mass by Palestrina, *Missa Iste confessor*, shows similar hierarchies to those of the corpus averages for this composer, but not all masses follow the pattern so closely. In practical terms, there were no explicit rules about which movement should feature more or less complex textures, but there was enough consistency across masses for auditory streaming complexity to be a useful cue for the listener, particularly for identifying the last three movements: Credo, Sanctus, and Agnus Dei.

The relative streaming complexity of the different movements (as shown in Figures 10 and 13) corresponds to musicologists' understanding of the conventions associated with each movement (Fallows, 2012). The subtle "V" shape indicates that streaming complexity decreases from the Kyrie to the Credo and then ascends through the Sanctus to the highest complexity in the Agnus Dei. The Credo is normally the lowest in streaming complexity; it has the longest text (163 words) and tends to use syllabic text setting and homorhythm, possibly to ensure that this fundamental statement of Christian belief was intelligible to educated listeners who knew Latin. The Sanctus, in contrast, is very melismatic; it has five sections (Sanctus-Pleni-Hosanna-Benedictus-Hosanna); the Pleni and the Benedictus are normally written for a smaller number (2-3) of voices. This movement shows the largest contrast between the voices active (relatively low) and the streaming complexity (relatively high) in this corpus. It is sung during the liturgical high point of the mass, transubstantiation. We might imagine that the smaller number of parts results in a muted and solemn sound, while the relatively high streaming complexity evokes the mystery of the real presence of God in the transubstantiation. The Agnus normally adds voices in the last of its three sections. Like the Kyrie and the Sanctus, it is very melismatic, permitting greater independence of voices. In addition to more voices, the last Agnus often has special effects: canonic structures, a cantus firmus in very long notes, or wider ranges (Fallows, 2012). The Agnus is designed to go out with a bang.

Overall, the consistency of the "V" shape in our corpus of mass cycles reinforces our initial claim that the sung mass ordinary cycle served as a way to mark experiential time in the context of the larger mass liturgy, perceptible without needing to understand the words. Fallows calls it "one coherent structure with a clear beginning and a clear end, whatever else is going on in the course of the celebration" (Fallows, 2012).

Composers

As in any musical genre, composers could make different choices about which movements should be most complex. Within this corpus, only a few composers were represented by enough works to test

quantitatively, but the one whose masses stood out most clearly was Josquin des Prez. Unlike the general average hierarchy in the masses of Palestrina, Victoria, and de la Rue, the seventeen masses that are reliably attributed to Josquin have streaming complexity averages that put the Credo movement high in complexity and the Sanctus relatively low. Figure 14 presents the distributions of complexity measure values per movement for one such mass, *Missa Sine nomine*.

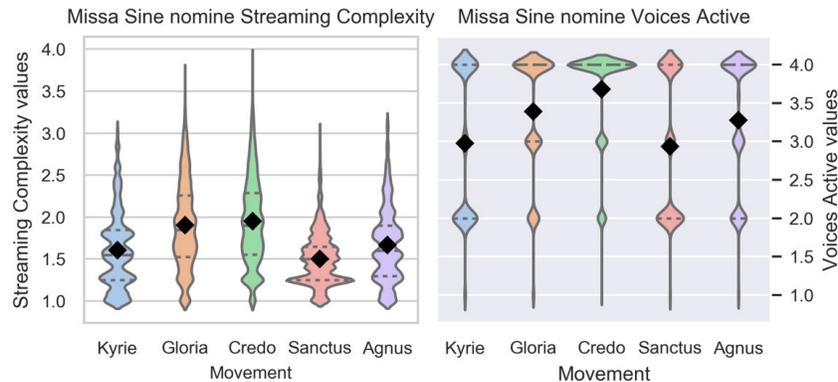


Figure 14. Distributions of Streaming Complexity and Voices Active values per movement of Josquin's *Missa Sine nomine*, showing typical relative values per movement for this composer.

Here the Credo spends the most time with four voices active and reaches the highest streaming complexity values. It is followed by a squat Sanctus distribution that spends a lot of time in two voices. These movements would undoubtedly contrast audibly in the complexity of the polyphonic texture. Of the masses weakly attributed to Josquin in this corpus, most of these followed distributions closer to that of the Palestrina mass in Figure 13 than to the Josquin example in Figure 14, confirming current judgements about attribution. The difference between the Josquin masses with strong attributions and the rest of the corpus is one of the most striking findings of our study. Unlike the other composers, Josquin chose a unique strategy: make the Credo (the longest, middle movement of the mass) the most complex, inverting the “V” pattern to a “mountain” pattern, and tone down the streaming complexity of the Sanctus. Like all the masses, Josquin's Agnus movements normally have a relatively high streaming complexity (see Figure 12), but not higher than his Credos.

CONCLUSIONS

This project arose out of an interdisciplinary collaboration combining insight and techniques from musicology, music technology, and music cognition. It was also informed by great familiarity with this genre of music. Julie Cumming has been playing, singing, teaching, and studying Renaissance music for 50 years. Finn Upham has been an amateur performer of early music since teenagehood and spent a year performing Renaissance masses weekly for an Anglican Cathedral. We leveraged our experiential appreciation for this genre of music to find a way to describe texture in polyphonic music, a feature that has rarely received the study it deserves. In our corpus of 216 polyphonic mass ordinary cycles, we found a number of musicologically interesting patterns.

Over 150 years, this liturgical genre increased in textural complexity, in the number of parts, the average number of voices active at a time, and in the average streaming complexity, as calculated from the interactions among the voices. Streaming complexity increased independent of the change in numerosity attributable to more parts. The five movements of the mass, with their own liturgical functions and formal contributions to the mass service, tended to differ in their streaming complexity and average number of voices active, creating a hierarchy where some movements tended to be more complex than others across the repertory, according to either feature. These differences were not absolute, but they were still common enough for multiple composers to show the same tendency, a soft V shape with the central text-heavy Credo maintaining lower streaming complexity and the closing glorifying Agnus Dei reaching the greatest heights. This persistent form allows us to imagine how the cycle could have functioned to mark time for the Renaissance listener. The exception, Josquin des Prez, appears to have had his own rules for which

movements were most complex in terms of auditory streams. The masses strongly attributed to this composer were measurably distinct from those works with weak attribution, providing new computational arguments for debates over who composed these masses.

The analysis could also be adapted to include more qualities informing auditory stream formation. The most pressing addition would be to include motivic repetition, since familiar and salient melodic segments can call attention and give identity to individual parts. This corpus covered the period in which imitative texture became dominant for sacred music, and it may well play a role in the growth in complexity already shown. Another research question opened by this analysis is whether composer-specific trends carried forward: were there composers after Josquin who wrote relatively complex Credo movements?

It would be immensely useful to tune the Streaming Complexity Estimate directly to reported or measured experiences of listeners. But even without this stage of validation, the strength of the results in this corpus analysis suggest that it establishes the importance of texture as a musical feature that changes over individual mass cycles, and over time, in significant ways. Our calculations were tailored to this genre of music, to specific types of listeners (non-expert), and to a particular listening context. Similar evaluations could be performed on other genres of polyphonic music, such as motets and madrigals, and with some additional changes, to classical genres such as the string quartet.

ACKNOWLEDGEMENTS

This article was copyedited by Annaliese Micallef Grimaud and layout edited by Diana Kayser.

NOTES

[1] Correspondence can be addressed to: Finn Upham, finn.upham@mail.mcgill.ca; Julie Cumming, McGill University, Schulich School of Music, 555 Sherbrooke St. West, Montréal, Québec, H3A 1E3, Canada

[2] julie.cumming@mcgill.ca.

[3] Most of the other texts and melodies said or sung in the mass changed according to the time of year or the feast day; the changing sections with music are known as mass propers.

REFERENCES

Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*. MIT press. <https://doi.org/10.7551/mitpress/1486.001.0001>

Cambouropoulos, E. (2008). Voice and stream: Perceptual and computational modeling of voice separation. *Music Perception*, 26(1), 75–94. <https://doi.org/10.1525/mp.2008.26.1.75>

Cumming, J. (1999). *The motet in the age of Du Fay*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511481789>

Cumming, J., & Tribble, E. (2019). Distributed cognition, improvisation and the performing arts in early modern Europe. In M. Anderson & M. Wheeler (Eds.), *Distributed Cognition in Medieval and Renaissance Culture* (Vol. 2, pp. 205–228). Edinburgh University Press.

Cusack, R., Decks, J., Aikman, G., & Carlyon, R. P. (2004). Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *Journal of Experimental Psychology: Human Perception and Performance*, 30(4), 643. <https://doi.org/10.1037/0096-1523.30.4.643>

Cuthbert, M. S. (2019). *Music21 Toolkit* (Version 5) [Python]. cuthbertLab. <https://web.mit.edu/music21/>

Duane, B. (2013). Auditory streaming cues in eighteenth- and early nineteenth-century string quartets: A corpus-based study. *Music Perception: An Interdisciplinary Journal*, 31(1), 46–58. <https://doi.org/10.1525/mp.2013.31.1.46>

- Fallows, D. (2012). The last Agnus Dei: Or: The cyclic Mass, 1450-1600, as forme fixe. In A. Ammendola, D. Glowotz, & J. Heidrich (Eds.), *Polyphone Messen im 15. Und 16. Jahrhundert: Funktion, Kontext, Symbol* (pp. 53–64). V & R unipress. <http://d-nb.info/1009693018>
- Fallows, D. (2009). *Josquin*. Brepols.
- Filippi, D. V. (2017). ‘Audire missam non est verba missae intelligere...’: The Low Mass and the Motetti missales in Sforza Milan. *Journal of the Alamire Foundation*, 9(1), 11–32. <https://doi.org/10.1484/JJAF.5.114048>
- Hiley, D. (1993). *Western plainchant: A handbook*. Clarendon Press; Oxford University Press.
- Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception*, 19(1), 1–64. <https://doi.org/10.1525/mp.2001.19.1.1>
- Jones, E., Oliphant, T., Peterson, P., & et al. (2001). *SciPy: Open Source Scientific Tools for Python*. <https://www.scipy.org>
- Kirkman, A. (2010). *The cultural life of the early polyphonic Mass: Medieval context to modern revival*. Cambridge University Press.
- Luko, A. (2008). Tinctoris on Varietas. *Early Music History*, 27, 99–136. <https://doi.org/10.1017/S0261127908000296>
- Marvin, C. (2002). *Giovanni Pierluigi da Palestrina: A Research Guide*. Taylor & Francis Group. <https://doi.org/10.4324/9781315050928>
- McKay, C., Cumming, J., & Fujinaga, I. (2017). Characterizing composers using jSymbolic2 features. *Extended Abstracts for the Late-Breaking Demo Session of the 18th International Society for Music Information Retrieval Conference, Suzhou, China, 2*.
- McKinney, W. (2010). Data structures for statistical computing in Python. In S. van der Walt & J. Millman (Eds.), *Proceedings of the 9th Python in Science Conference* (pp. 51–56). <https://doi.org/10.25080/Majora-92bf1922-00a>
- Nosow, R. (2019). Simultaneous ceremonies at the collegiate church of St. Donation in Bruges. In A. Pavanello & D. V. Filippi (Eds.), *Motet cycles between devotion and liturgy* (pp. 37–52). Schwabe Verlag.
- Pirrota, N. (1966). Music and cultural tendencies in 15th-century Italy. *Journal of the American Musicological Society*, 19(2), 127–161. JSTOR. <https://doi.org/10.2307/830579>
- Rodin, J., & Sapp, C. (2019). *The Josquin Research Project* [Music Database]. The Josquin Research Project. <https://josquin.stanford.edu/>
- Rodin, J. (2019). "Attribution levels." The Josquin Research Project. <https://josquin.stanford.edu/about/attribution/>
- Shinn-Cunningham, B., Best, V., & Lee, A. K. (2017). Auditory object formation and selection. In *The auditory system at the cocktail party* (pp. 7–40). Springer. https://doi.org/10.1007/978-3-319-51662-2_2
- Tinctoris, J. (1963). *Dictionary of musical terms: An English translation of Terminorum musicae diffinitorium: together with the Latin text* (C. Parrish, Trans.; pp. xi, 108 pages). Free Press of Glencoe.
- Vallat, R. (2018). Pingouin: Statistics in Python. *The Journal of Open Source Software*, 3(31), 1026. <https://doi.org/10.21105/joss.01026>