Intelligibility Redux: Motets and the Modern Medieval Sound *

Anna Zayaruznaya

NOTE: The examples for the (text-only) PDF version of this item are available online at:
    http://www.mtosmt.org/issues/mto.17.23.2/mto.17.23.2.zayaruznaya.php

ABSTRACT: The medieval composers of polytextual motets have been charged with rendering multiple texts inaudible by superimposing them. While the limited contemporary evidence provided by Jacobus's comments in the *Speculum musicae* seems at first sight to suggest that medieval listeners would have had trouble understanding texts declaimed simultaneously, closer scrutiny reveals the opposite: that intelligibility was desirable, and linked to modes of performance. This article explores the ways in which 20th-century performance aesthetics and recording technologies have shaped current ideas about the polytextual motet. Recent studies in cognitive psychology suggest that human ability to perform auditory scene analysis—to focus on a given sound in a complicated auditory environment—is enhanced by directional listening and relatively dry acoustics. But the modern listener often encounters motets on recordings with heavy mixing and reverb. Furthermore, combinations of contrasting vocal timbres, which can help differentiate simultaneously sung texts, are precluded by a blended, uniform sound born jointly of English choir-school culture and modernist preferences propagated under the banner of authenticity. Scholarly accounts of motets that focus on sound over sense are often influenced, directly or indirectly, by such mediated listening.

[0.1] The growth of commercial aviation in the 1950s brought with it a new problem for air-traffic controllers. Presented with the voices of multiple pilots emanating from one loudspeaker, they had trouble segregating these simultaneous streams of information, and found themselves at risk of misunderstanding important messages. Their difficulties led to the first research on what has since been termed the "cocktail party effect" or "cocktail party problem"—the isolation of one sound-source in the presence of many. The brains of most humans, and many animals, are able to

navigate complex social soundscapes without much conscious effort. That is the effect. But how exactly we do this, and why it should have been difficult for those air traffic controllers—and later for computers—constituted the problem.[1]

[0.2] For music historians, a parallel effect—or perhaps problem—is inherent in polytextual motets, whose superimposition of multiple strands of texts has led us to inquire whether audiences could have attended to, or ever been meant to understand, any of the words during performance. Scholars have answered this question in various ways, and their stances about intelligibility have in turn shaped their views of the genre, the compositional priorities of its composers, and the role of text in its aesthetics.[2] The present study focuses on an aspect of the problem that has not been foregrounded in existing discussions: the roles played by modern performance aesthetics and listening conditions—and especially by recorded sound—in shaping these views. As a case study, I focus on Christopher Page's work with Gothic Voices and his own arguments about how motets would have been heard by their medieval audiences. I then return to the cocktail party effect and to theorizations of acousmatic sound to see what cognitive science and sound studies can tell us about the listening conditions under which the words of polytextual motets might be intelligible.

[0.3] I offer two caveats before proceeding. First: in exploring the acoustics of ideal listening conditions, I do not imply that those conditions ever existed, or were the conditions under which motets were heard. It is undoubtedly the case that some singers or listeners might have had motet voices or even entire compositions memorized (Busse Berger 2005), rendering the whole intelligibility issue moot. Furthermore, the frequent use of refrains and other quotations in French motetes from the thirteenth (Saltzstein 2013a, 2013b) and fourteenth centuries (Boogaart 2001, Zayaruznaya 2015b) would have rendered some text immediately intelligible, even upon first hearing. And it seems clear that the poetry of polytextual motets was sometimes experienced in monotextual contexts. For example, the early 14th-century Bodleian manuscript Douce 308 transmits 64 motet texts without notation (Atchison 2005). For a later generation, the literary circulation of motet texts by Philippe de Vitry, as documented by Wathey (1993, 1994), must have accorded some listeners and patrons the opportunity to study the poetry on the page, and perhaps to commit it to memory. It is also clear from a comment in the *Speculum musicae*, as well as from fourteenth-century literary and musical sources that motet voices were sometimes sung separately.[3] But references to voices sounding on their own do not obviate questions about intelligibility: fourteenth-century motets *not* by Vitry seem to have been transmitted as poetry very rarely, and it is unlikely that all listeners would have had access to all motet texts ahead of hearing the works performed. Polytextuality is a defining aspect of the medieval motet genre, and it is worth asking what listeners might have heard (might have been expected to hear, might have been able to hear, or might have chosen to hear) in a *tutti* performance.

[0.4] A second caveat concerns the oft-quoted evidence that medieval listeners might have had problems with comprehension when motets were sung. This stems from an anecdote in the *Speculum musicae* in which a listener at a "gathering of discerning people . . . asked what language the singers were using: Hebrew, Greek, or Latin." If this person could not even identify the language, it seems impossible that s/he could have understood the words, and accordingly this passage has been used to support the claim that "the very nature of double motets with their simultaneous presentation of two texts jeopardized the 'message' of either [text]" (Pesce 1986, 91–92). But a contextual reading of this comment gives an entirely different impression of its meaning. The passage is accordingly worth quoting in full:

> In a certain gathering in which skilled singers and discerning laypersons were gathered together and where modern motets were sung according to the modern manner, and some old [motets] as well, I observed that the old motets and also the old manner of singing were more adequately pleasing, even to the laypersons, than the

new. And even though the new manner [initially] pleased in its newness, this is no longer the case, for it begins to displease many.

Therefore let the old songs and the old manner of singing and notating be called back to the native land of singers. Let these things come back to use, and let the rational art flourish once more. It has been exiled, along with its [corresponding] manner of singing, as though violently thrown down by the fellowship of singers. But the violence need not be perpetual. Why does such wantonness of singing give pleasure—this curiosity in which, as may be seen by anyone, *the words are lost*, the harmony of the consonances diminished, the length of notes changed, perfection disparaged, imperfection elevated, and measure confounded?

I saw how in a great gathering of discerning people, when motets were sung according to the new manner, *it was asked what language the singers were using, Hebrew, Greek, or Latin, for it was not understood what they were saying. Thus the moderns, although they write many beautiful and good texts in their songs, lose them in their manner of singing, since they are not understood.*[(4)]

[0.5] In context, the meaning of the final sentence is unambiguous, serving as a summary of what we should get from the anecdote: it is unfortunate that the well-written texts of the moderns are masked by the new manner of performance. The implication is that if the same good texts were delivered in the old manner of singing, *they would not be lost*. In complaining that texts are obscured, Jacobus tacitly acknowledges that the clear projection of text is a merit of the older repertory that he champions. Far from proving that medieval listeners could not understand the words of motets in performance, then, this passage in fact suggests that there was an expectation that text *would* be intelligible.

[0.6] Furthermore, there is an element of humor here (Hentschel 2001, 120). As Suzannah Clark has pointed out, "Jacques's complaint that they could be singing in Hebrew or Greek is, of course, spurious, for the upper voices of motets were either in Latin or French or both" (2007, 32n6). Hence the observer is making fun of a delivery of Latin text which was, to his mind, muddled by the new manner of singing. His mocking observation would hardly make any sense unless understanding the text was a recognized aspect of listening to motets.

[0.7] Jacobus's comments are useful in part because they bridge the two repertories—French *ars antiqua* and *ars nova*—that contain the bulk of surviving polytextual motets, signaling that the same or similar questions might reasonably be asked of both.[(5)] And while he is not a trustworthy source for reporting on *ars nova* performances in any but a negative way, Jacobus tells us some useful things about contemporary performance contexts and expectations. On the former front, his recollection provides evidence that it is not anachronistic for us to speak of an "audience" for motets, using the word in a sense not very different from our modern one: there were listeners distinct from the singers themselves who attended to motet performances.[(6)] Further, Jacobus confirms that he and the learned person whom he paraphrases were listening in part for the texts and would agree that in an ideal situation they should be intelligible. Finally, he gives us the key to understanding why they sometimes are not. Towards the end of the passage he uncouples composition and performance, which had been conflated up to this point in the chapter, focusing exclusively on the moderns' manner of singing as the cause of incomprehension. In doing so, he reminds us of an obvious but important fact: performances can make words audible, or they can obscure texts, causing them to become lost. In what follows, this passage serves as an impetus to turn from the new manner to a newer manner still, and consider the effect that twentieth- and twenty-first century performances of *ars nova* motets have on the intelligibility of their texts.

Medieval Listeners, Modern Sounds

[1.1] To talk of motet texts being intelligible or remaining indistinct, and if intelligible of being attended to or ignored, is to evoke a complicated web of interrelations among singers, sounds, and listeners. The nodes of this web are only potentially joined: singers may send crisp consonants into a booming acoustic, or perfectly comprehensible sung text may reach the ear of a listener ignorant of its language, uninterested in its contents, or distracted. But nevertheless the temptation is strong, given the dearth of evidence on all fronts, to use information about listeners to make conjectures about medieval sounds and singers. This was the approach of Christopher Page, who set out to democratize the motet in his monograph *Discarding Images*, arguing that it was intended not only for the medieval intelligentsia, but for a much broader cut of society. In the course of diversifying the intended audience for the genre, Page flattened the hypothetical listening experiences available to that audience by downplaying intellectual engagement in favor of sensory pleasure. Objecting to "the destabilizing quality that much modern criticism tends to discern and explore" (1993, 96), he depicted motets as integrated wholes, their separate voices merged into a unified mass of sound by congruences in theme, vowel sound, and diction.

[1.2] Such reframing has implications for the importance of individual texts—or of any text—as a salient aspect of composition or reception. Indeed, Page suggested that part of the pleasure gained from motets by their diverse medieval audiences may have been derived from *not* understanding the texts, an experience he characterized as "'intellectual' in this sense: it produced the exhilaration of knowing that a piece contains more than one can ever hope to hear" (101):

> Some listeners in the thirteenth century may have
> enjoyed the sound-patterns of Latin motet poetry
> without deriving any significant understanding of the
> sense, much as some modern listeners do today when
> hearing these motets (or operatic arias, for that matter)
> . . . it can scarcely be doubted that many motets were
> designed to induce an exhilarating impression of
> words leaving sense behind (86).

Noteworthy in the above-quoted excerpt is the elision of reception with compositional design. Not only were the sound-patterns of foremost interest to some listeners, but the motets were *so designed* as to foreground sound over sense.

[1.3] Other theorizations of the motet, both before and after that in *Discarding Images*, have carved out a semantic as well as a sonic role for motet texts in performance. Dolores Pesce suggested in 1986 that in 3-voice Latin *ars antiqua* motets, music serves texts by giving "musical prominence to . . .figures of repetition and rhetorical expressions" and that in some cases this kind of emphasis may represent "at least an attempt to give musical expression to the words" (94, 97). Margaret Bent, the most vocal advocate for the importance of text to the analysis of motets, while conceding an "impossibility of understanding [motet texts] at a first or unprepared hearing" (1993, 632), preserved an intellectual dimension in motets by invoking the "prepared listener": an "experienced listener who, like Boethius's *musicus*, 'exhibits the faculty of forming judgments according to speculation or reason'" and is already familiar with the composition (2003, 387). And Suzannah Clark (2007) has argued that in a French motet from the Montpellier codex, some level of understanding is made possible—even assured—by compositional choices such as the placement of refrains, the use of musical echoes, and the coordination of rests between voices.

[1.4] Alongside these more sympathetic approaches to the semantic contents of motet texts, Page's characterization of the genre in fundamentally sonic terms lives on, not only in textbooks (a point to which I return below), but also in Emma Dillon's recent sound-centered account of some thirteenth- and early fourteenth-century motets as evoking an effect she dubs "supermusical" (2002). This "sonic superabundance" (57), defined as "a sound of words lost in the mêlée of music"

and a "surplus of language" (6–7) has clear roots in *Discarding Images*.

[1.5] There is no doubt that the understanding of a piece by a listener prepared in Bent's sense—intimately familiar with a motet's texts, proportions, and numerical structures, as well as the liturgical context of its tenor—would be deep.[7] No doubt also that a listening experience such as the one suggested by Clark—attuned to intertextual and intermelodic references—would be meaningful. But what of the less-prepared listener? Must we assume that understanding was beyond their grasp, and that something like Dillon's "supermusical" effect was of most relevance to their experience? Jacobus's reminder about the importance of the manner of singing can serve as an invitation to evaluate the performance aesthetics underpinning Page's incredulity about intelligibility.

[1.6] A noteworthy aspect of Page's claim is that it purports to rest on the broadest, most defining elements of the genre. It is the very "*existence* of triple motets with three simultaneous poems" that "*proves* that the aesthetic of the poet is one which allows verbal communication to decline as metrical, musical, and structural ambitions mount" (1993, 101). Opposing "verbal sound" to "verbal sense," despite the dependence of the latter on the former, Page uses a poetic aesthetic to support a claim about the motet *qua* genre:

> Of all the musical genres known to the Middle Ages, it is the motet which most candidly acknowledges the importance of verbal sound over verbal sense by placing two or even three texts together, minimizing their intelligibility but maximizing their phonic contrast (1993, 85–6).[8]

[1.7] The link between words and sounds is bound up in the category of "phonic contrast"; this too is sometimes located in the repertory itself, for example when analysis highlights the coordination of rhyme-sounds and vowels between voices. But ultimately Page, like Jacobus, knows that varying modes of performance can minimize or maximize sonic contrast. And here it helps that Page happens to be a scholar as well as a performer, and thus someone whose performance ideal we can both apprehend and evaluate.

[1.8] Page's debt to recordings—especially to his own work with Gothic Voices—is self-avowed:

> The primary inspiration for [*Discarding Images*] has been provided by performance . . . recordings are sometimes superseded by advances in knowledge, and are often vanquished by changes in taste, yet innovative or challenging performances can none the less disturb a wide range of preconceptions that we may unwittingly hold about the interest and scope of a repertory. (1993, xx)

Indeed Page is hardly alone in being influenced by Gothic Voices. Fueled by—and fueling—his research into the question of instrumental versus vocal performance of medieval song, Page's recordings allowed an entire scholarly community to reconceive of a wide repertory of music as primarily vocal (Page 1982 and 1988, Leech-Wilkinson 2002, 104–50). But if Page's re-imagination of the medieval sound was fresh and radical with respect to its instrumentation, another aspect of his aesthetic was rather traditional: the timbre of Gothic Voices was inherited perhaps too readily from the English choral tradition.

[1.9] This is not to say that Page was not thoughtful about questions of timbre. If, as John Potter has argued, "the position of authority [held by the British early music sound] has been achieved with the minimum change to the techniques that the singers learned to gain their choral scholarships" (Potter 1998, 116), this is less a matter of laziness than of agenda in Page's case. He was reacting to the bright, mixed sound of instrumental performance current in the first stage of the early-music revival—a sound, in his words, of "boldly individualized lines collaborating in a polyphony which

is extrovert and almost heraldic in colour, which is candid, even naive, in its directness of address to the listener" (Page 1982, 441). The ascetic English choral ideal evocatively described by Daniel Leech-Wilkinson as "freshly cleaned Anglicanism" would have been the perfect antidote (2002, 206).

[1.10] This pure sound with its equal timbres, matched vowels, and lack of vibrato stands behind Page's conception of the motet. As his performance aesthetic blends sounds together into a unified whole where sonority reigns supreme, so his arguments about *ars antiqua*—and *ars nova*— aesthetics ultimately stress unity over contrast. For me, "synchronization of vowels" and "verbal identity" that lead to sonorities "of exceptional clarity and forthrightness"—characteristics that Page argues are inherent to the compositional fabric of motets (1993, 103, 105)—bring to mind an Oxbridge college choir, or the Tallis Scholars. Theirs is a sound in which we may, borrowing again Page's account of the motet, "[enjoy] the sound-patterns of Latin motet poetry without deriving any significant understanding of the sense" (1993, 86). And in the phrase "there is variation and contrast, but there is neither conflict nor tension," which Page uses to describe the relationship between a motetus and a triplum (1993, 97), I cannot help but read also a description of the sound-world cultivated by Gothic Voices. Bent has warned us of the "danger of drawing circular conclusions from particular styles of performance that emphasize some rather than other features of the music" (1993, 630).[9] And indeed, Page's arguments about the priorities of the genre could double as an aesthetic manifesto: "The complexity is purely phonic and not semantic. As I read them, the texts of this motet are designed to be so similar to one another that there is little or no contrast between them" (1993, 110). The texts, or the voices that execute them?

[1.11] Page himself has cautioned that his reconstructions of medieval sounds are not "wholly objective," in part because they are "based upon . . . experience with professional British singers, most of whom received their training in the choir of some cathedral or collegiate establishment in England" (1988, 149). But he marshals various kinds of evidence to suggest that the sound of Gothic Voices is at least mostly objective, and that the late-medieval sound might have been similar to the modern British one. For example, the similar range of motet voices is taken to suggest that they were "performed by identical (or by nearly identical) voices":

> Many of the motets in the Bamberg manuscript (probably dating from c. 1285) exploit
> a range of around eleven notes from top to bottom, while some individual voice parts
> also employ this compass. It would therefore seem that 13th-century motets were
> conceived to exploit the resources of two, three or four equal voices; one singer, in
> other words, could normally perform any part in a motet—tenor, duplum, triplum, or
> quadruplum—without undue difficulty. The result would have been a well-blended
> sound with each part distinguished by its text rather than by its colour. . . . This is
> quintessentially music for singers, designed to exploit the sounds of Latin and Old
> French as put into song by voices of very similar type singing in very much the same
> kind of way. (1988, 153–4)

[1.12] Such a reading of the evidence blatantly conflates range with tessitura, and by the following century there is evidence that, despite the small ambitus of many motets, some singers would have been more used to singing tenors, and some, upper parts.[10] But more basically, nothing guarantees that two medieval singers with the same range would necessarily have had "voices of very similar type." The unity of sound achieved by English singers is the result of a huge system of choir schools and choral scholarships, with many professional performers emerging from either Oxford or Cambridge (Potter 1998, 116–17). Arguably, only such an infrastructure can produce singers so similarly trained as to perform music in this way. That is not to say that those singing motets in the thirteenth and fourteenth centuries were not experienced: the degree of specialized knowledge needed to sing mensural notation was great, and the singers able to execute motets

would probably have begun as choirboys or, in the case of the nobility, would have been taught by clerics. [(11)]

[1.13] While the medieval performers of motets would all have been schooled, they would by no means have been products of the same school or the same aesthetic (McGee 1998, 17–18). We can imagine a scenario where a Latin motet was sung by three university men, each from a different part of Europe and with a different way of pronouncing Latin. Would their vowels be matched? We know that even different regions within France had their own local ways of singing: the canons at Lyons cathedral, for example, were known to shout (Page 2000, 248). On the international level, these differences were found worthy of—and subject to—rivalry and ridicule. Thus an Italian rhetorician wrote in 1226 that

> the Greeks say the Latins bark like dogs and the Latins say the Greeks growl like foxes. The French claim that the Italians groan . . . the Italians, on the other hand, say that French and Germans emit tremulous sounds like someone suffering from fever, and that they sing so loudly they must think God is deaf. (Vecchi 1967, 266–73; trans. Page 2000, 348)

Motets were performed in precisely those places where singers from different corners of Europe mingled—in universities, at courts, and on occasions of state. Might some performances not have had in them, then, a trace of those exaggerated differences? If not exactly emitting howls, groans, and feverish wails, it would nevertheless be surprising if such diverse performers could be found "singing in very much the same kind of way."

[1.14] Other scraps of evidence support the idea that the sound-world of motets may have been a varied one. One unlikely source of insight comes from a description of dog-song in Gace de la Buigne's *Roman de Deduis* (c. 1359–77). In arguing that hunting with dogs is more noble than hunting with falcons, the allegory Love of Dogs describes his pleasure at hearing dogs singing a motet:

> Les uns vont chantans le motet,
> Les autres font double hoquet.
> Les plus grans chantent la teneur.
> Les autres la contreteneur.
> Ceulx qui ont la plus clere guile
> Chantent les tresble sans demeure,
> Et les plus petis le quadouble
> En faisant la quinte sur double.
> (la Buigne 1951, ll. 8081-8)

> Some of [the dogs] go singing the motet;
> the others make a double hocket:
> the largest [dogs] sing the tenor
> and the others the contratenor.
> Those that have the clearest voice
> sing the triplum without delay
> and the smallest [sing] the quadruplum,
> making a fifth on the duplum
> (Leach 2007, 181–2).

[1.15] This is hardly informative from the perspective of performance practice. But even parody contains potentially useful information. Indeed, Page has cited this passage in support of all-vocal performance of chansons (1977, 487). We may therefore note the variety of look and sound evoked

in it. Disparate groups produce sounds varying in both timbre and ambitus: the dogs singing the triplum have penetrating voices, the big dogs sing tenor, and the quadruplum is left to the puppies. This variety is the more significant when we remember that, as Page has noted, many motets are relatively narrow in range, all their voices together often covering no more than a twelfth—and this in turn can be the range of a single voice in a more virtuosic work such as a *chasse* (Kügle 1993, 153n134). Thus Love of Dogs presents us with a more varied picture than we might expect (insofar as we have any expectations about canine motet performance).

[1.16] This variety is not dissimilar from that cultivated in the image of performance which opens the motet section of F-Pn fr. 1584 (Machaut's manuscript A; see **Figure 1**). Here clerics and laypersons gather around a cask, drinking wine and singing a motet from a scroll.[12] Their diverse ages and backgrounds, coupled with their presumed state of inebriation, all but guarantee that their voices will not sound similar. This group is a far cry from the singing angels of van Eyck's Ghent altarpiece, whose young identical faces seem to imply a painstakingly blended sound. This group of drinkers hints instead at a smorgasbord of timbres and volumes.

[1.17] This, too, is not the kind of evidence we would like. Rather than recording anything akin to a performance, the image may be serving as an allegorical depiction of motets, with their mixture of the sacred and secular.[13] And Love of Dogs is at least as biased an observer as Jacobus. But choices were made in both cases—the choice to use a motet as the example to be barked; the choice to depict a diverse group of performers. Also noteworthy is the joyful, easy mood that characterizes both depictions. Nothing could be more natural, Love of Dogs wants us to believe, than this pack's rendition, and "there isn't a responsory or an alleluia, even if it were sung in the Kings' chapel . . . that would cause the same tremendous pleasure that one experiences in hearing such a hunt" (trans. Leach 2007, 181).[14]

[1.18] Some of the joy and spontaneity that seem to have been aspects of motet performance are lost when a unified sound is the goal. Indeed, Page has suggested that singing motets should be difficult, even uncomfortable for the singer:

> Something is often to be gained by pitching these motets in a way that lifts the singers
> on the upper parts away from the comfortable range where they can coast or croon.
> When the music is lifted above this range the relatively higher pitch brings many
> things under closer control . . . [and] the singers, recognizing that danger is only about
> a tone away, begin to work hard in a fashion that is always beneficial to this music.
> (1988, 158)

Elsewhere, the same author has written that in the Middle Ages "to succeed [at composition] was to create something that singers could control and dominate," but here the music is instead allowed to control the singers, eliciting a unified sound from them with the threat of nearby danger.[15]

[1.19] The reason that all of this is pertinent to an argument about intelligibility is that differences in timbre are among the variables that allow the human ear to focus on one sound-source in the presence of others. A number of other aspects of performances—including acoustics, listeners' ability to see the performers, and the role of recording technology—also bear on this issue. In what follows I will not suggest that performance conditions that maximize intelligibility are either authentic or desirable. I will only review the existing evidence on the acoustic and cognitive phenomena that allow a person to navigate a busy soundscape—whether it be a hunt, a cocktail party, or a "gathering of discerning people."

The Cocktail Party Phenomenon

[2.1] Given the high likelihood that a contemporary listener interested in motets would hear multiple performances of the same work, the question of whether it is possible to hear multiple texts at the same time wanes in importance next to a more basic one: under what conditions is it possible to isolate any text in a polytextual performance?[16] Rephrasing the query in this way clarifies the link between the psychological "cocktail party phenomenon" and the polytextual motet. Whether in a busy room where a person behind us is speaking louder than our collocutor, or in a composition where a triplum is singing higher and faster than the motetus which we are perhaps desirous of hearing, the cognitive task is the same: to isolate one sound-source in the presence of other, sometimes more prominent ones. But the assumptions have been very different in the two domains: while in musicology the superimposition of multiple texts has called their audibility and intelligibility into question (recall Page's claim that superimposing multiple texts is tantamount to "minimizing their intelligibility"), in psychology it is readily agreed that the cocktail party phenomenon is "one of our most important faculties"—"such a common experience that we may take it for granted" (Cherry 1957, 278). And indeed, most listeners who walk into a busy café and shift attention from one conversation to another will not give a second's thought to the considerable powers of segregating sounds that undergird such an activity. Try it.

[2.2] Understanding speech in a complex environment involves the segregation of simultaneously occurring sounds into streams originating from different sources, a process known as auditory scene analysis (Bregman 1990). The ability to do this successfully rests on a number of connected factors pertaining to the auditors, the sound sources, and the environment in which listening is taking place. Although typically people carry out auditory scene analysis automatically, there is individual variability among listeners.[17] Auditory scene analysis may be especially difficult for listeners who are deaf or hard of hearing, even with auditory access technology (Bayat et. al 2013; Kerber and Seeber 2012), as well as those on the autism spectrum (Lin et. al 2017). On the other hand, musicians have been found to have an advantage in discerning components within complex auditory scenes (Pelofi et. al 2017).

[2.3] Where the sound sources are concerned, differences in the fundamental frequencies and spectral profiles of sounds—that is, differences in timbre—are vital to the segregation of concurrent sounds in the same register.[18] And although this kind of segregation is most often employed in social settings (for example, to distinguish two simultaneously speaking voices), the implications for musical audition are clear: "sounds of similar timbres will group together, so that the successive sounds of the oboe will segregate from the harp, even when they are playing in the same register" (Bregman 1990, 19). When sounds are separated by range as well as timbre, lower and higher sounds will naturally separate into groups, or perceptual streams. One study found that different vowels sung simultaneously can be parsed more easily if they are at least 4 semitones apart.[19]

[2.4] The timbral and registral aspects of sound sources are monaural cues—that is, they can be perceived with one ear. Binaural processes are even more crucial to solving the cocktail party problem. The distance between the ears allows for accurate identification of the location of a sound-source along a horizontal axis, since the emanating sound waves enter the two ears at different times. A more subtle process known as "head-related transfer function" localizes sounds along the other axes.[20] It has long been recognized that such "spatial hearing plays a major role in the auditory system's ability to separate sound sources in a multiple-source acoustic environment" (Haykin and Chen 2005, 1878–9), such that an increase in distance between the sources of simultaneous sounds aids in their auditory differentiation (Drennan, Gatehouse, and Lever 2003). And motion on the receiving end helps: Bregman (2015, 17) has recently suggested that listeners' abilities to move their heads and bodies—even slightly—also aid auditory scene analysis, noting that studies which immobilize the subject impose a set of unhelpful constraints.

[2.5] Environmental and visual factors also affect the difficulty of auditory scene analysis.

Excessively reverberant room acoustics can interfere with localization, though the auditory system is powerful enough to suppress even moderate amounts of echo and focus instead on the source sound (Haykin and Chen 2005, 1885). This focus is aided by visual cues such as facial expression and the movement of the lips in speech (Summerfield 1992). Even under ideal listening conditions, lip-reading can help resolve certain ambiguities—for instance by giving visual cues to amplify the rather small phonetic difference between a "b" and a "v" (McGurk and MacDonald 1976). In noisy situations such cues become much more important, and an ability to see the speaker is thus a significant asset in understanding them (Ma et. al 2009).

[2.6] Together these monaural and binaural processes enable auditory scene analysis, the subconscious prerequisite to solving the cocktail party problem. Once sound sources are localized, the role of attention comes into play. Attention is, in William James's classic definition, "the taking possession of the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others" (1890, 403–4). In the context of the cocktail party problem, several kinds of attention have been studied.

[2.7] With "selective attention," listeners choose to focus on one source and ignore others. Here, the comprehension level of the primary stream is high, but the secondary stream is hardly attended to. In an early study, subjects had different texts played into each ear, and were asked to speak the right-ear text back while they listened (a process referred to as "shadowing"). Colin Cherry found that the "rejected" text got very little attention—indeed, listeners did not notice when that left-ear text switched to German (1953, 978). But more recent studies have shown that while shadowing or listening actively to one text, listeners will still hear or experience aspects of the other: they may be able to recall some words from the "rejected" stream, and may also be subconsciously affected by text that they consciously choose not to attend to (Wood and Cowan 1995; Dupoux, Kouider, and Mehler 2003). And in listening to melodies, an unattended voice can influence a listener's perception of an attended one if the two are close in pitch (Davison and Banks 2003).

[2.8] Listeners can also engage in "divided attention," focusing simultaneously on more than one sound source, or "switched attention," alternating between channels. These different types of focus can be produced by telling the listener what to pay attention to, although certain sonic events, such as hearing one's own name in a "rejected" stream, can produce shifts of attention even during selective attention. Furthermore, factors that interfere with auditory scene analysis can make some kinds of attention more difficult to sustain. For example, as Wood and Cowan have noted, "selective attention depends heavily on physical (e.g. voice) differences between channels," and thus dichotic studies that use similar monotone voices in both ears may not allow listeners to fully concentrate on one channel or to fully ignore the other (Wood and Cowan 1995, 255).

[2.9] If there are any similarities we can count on between ourselves and medieval listeners, they are in the makeup of our auditory systems and the cognitive mechanisms that support listening.[21] For the cocktail party problem predates cocktails—solving it is a basic necessity for human society. But these similar biological systems and processes receive radically different musical stimuli today than they did 600 years ago.

Medieval Songs Remixed

[3.1] Listening to medieval music in the twentieth and twenty-first centuries often means listening to a recording, experiencing the original performance in an acousmatic guise—that is, made by a source which cannot be identified. While such sound is so common as to be taken for granted, Michel Chion warned already in 1983 that it produces a listening experience "whose consequences are more or less unrecognized": "the acousmatic situation changes the way we hear" in that "it creates favorable conditions for reduced listening which concentrates on the sound for its own

sake, as sound object, independently of its causes or meaning" (Kane 2014, 4).[22] As Brian Kane has demonstrated, the "reduced listening" afforded by acousmatic sounds "motivates a reification of the sonic effect" in a way that influences the kinds of analysis to which the sound is subject, and the kinds of narratives in which it plays a part: "the autonomous sound, bereft of its source, is... integrated into the virtual world of musical composition; shedding its source, it can fully participate in the virtual connection of tone to tone, in the metaphorical gravity of tonal-harmonic organization, or in the expressive analogies of musical sound with emotional states" (2014, 8). We see precisely this in Page's account of motets as "designed to induce an exhilarating impression of words leaving sense behind and beginning to *skirl*" —that is, to sound out in a shrill cry (1993, 86). Here is as good an example as any of the way in which "the isolation of hearing from seeing in the acousmatic experience directs the listener's attention toward the sound as such" (Kane 2014, 5).

[3.2] The passiveness of Kane's account of attention in the acousmatic context ("the isolation of hearing from seeing . . . directs the listener's attention") is tellingly different from the more active modes discussed in the literature on the cocktail party effect ("selective," "divided," "switched"). These kinds of attention are, after all, predicated on being able to localize the source of sound, and recordings ratchet up the difficulty of auditory scene analysis considerably. The three-dimensional localization of sound sources usually carried out by head-related transfer function is rendered ineffective when the auditory scene is flattened to two channels, and even the information potentially contained in these is often altered or erased by editing and mixing. Listeners are also often faced with excessive reverb, whether original to the recording space (often a large church) or added in post-production to evoke a churchy "medieval sound." Such surroundings and post-production choices inevitably muddy the waters, making it even more difficult to pick out the texts of individual voices. The challenges created by reverb can be compounded with those posed by timbral blend to create recordings in which is it virtually impossible to tell voices apart.

[3.3] The 2004 Hilliard Ensemble recording of Machaut's motets offers a good example of such challenges. A large amount of reverb cloaks each motet with a blanket of overtones and echoes.[23] Under this cover lies the ensemble's characteristic and oft-lauded blend—a prime example of the careful control and masterly listening that are the products of English training. The singers on the recording are at times difficult to differentiate by any cue other than range—and range can be a fickle indicator.

[3.4] Consider their rendition of Machaut's *Dame/Fins cuers doulz* (M11). Listening to mm. 11–31 (**Audio Example 1**), the impression is of a tenor and two texted voices almost identical in timbre and separated by anywhere from a 3rd to a 6th. The highest voice is especially striking. It begins with an antecedent-consequent figure, and moves down twice, first to an E supported by an unstable inflected sonority, then to a more stable G (see **Example 1**). Soon afterward we hear an oscillation of oscillations, as semibreve groups on G-F-G and A-G-A alternate, A-G-A finally seeming to win out, its repeated emphasis on A causing a mounting tension that is finally resolved by a directed progression to C (see **Example 2**). In a context where repetition on this scale is exceedingly rare, *Dame/Fins cuers doulz*, and this passage in particular, easily number among the most remarkable and haunting moments in Machaut's oeuvre.

[3.5] But a look at the score reveals that the melodies excerpted above are not the top voice, but the composite results of a series of voice crossings: motetus and triplum alternate the antecedent and consequent parts of the descent from A, and then trade off the increasingly insistent oscillations (see **Example 3**). Knowing this, it is possible to teach ourselves the difference between the very similar voices of countertenors David James and David Gould, so as to begin to hear their crossing lines. But this is a task requiring focused work and a score. How can we consciously direct our attention when the first step—auditory scene analysis—is so difficult? And of course, the visual aspects of performance, which would be a great aid in a situation where such similar sound

sources vie for attention, are as unavailable here as in any audio recording. In this case, trying to understand the text of a given voice is like trying to hear a neighbor at a party held in a vaulted hall where the guests wear opaque masks over their faces, while their voices are re-mixed and projected through two speakers at the front of the room with extra reverb added by the capricious hosts. It is much more intuitive to listen to this blanket sound in a reduced way, giving in to its acousmatic reality and concentrating on it "as sound object, independently of its causes or meaning."

[3.6] Of course, not all recordings have the same homogeneity of sound. In a Musik performance by the Studio der Frühen Musik, the upper voices of *Dame/Fins cuers doulz* are differentiated by voice type: the triplum is sung by mezzo-soprano Andrea von Ramm (Andrea Ramm von Marnov), and the motetus by contratenor Richard Levitt. A drier acoustic also helps the listener, as does unadulterated stereo: we hear Levitt firmly at right, and von Ramm at left (**Audio Example 2**). A listener still need not focus on one singer or the other, but such attention is not precluded. Nor are differentiated timbres solely a typical property of older recordings. The 2002 album of Machaut's motets by the French group Ensemble Musica Nova also uses two rather different voices to render *Dame/Fins cuers doulz*. The triplum is sung by a bright tenor with forward resonance, while the motetus is breathier and has a slight buzzing quality; these distinct timbres make the voice crossings in mm. 11–28 rather more discernible (**Audio Example 3**). Once again, a relatively dry acoustic adds to the intelligibility of these voices, which are pleasantly resonant on their own. But even—or rather, especially—in these latter cases, how much more is gleaned in a live performance, where the ability to distinguish sound sources from one another is rendered all-but-automatic by their location in three-dimensional space, and the faculty of sight can aid our comprehension of texture and text, showing us who is singing what, when.

[3.7] Commenting on recordings in an article is a relatively simple matter in these days of online content, but it remains difficult, if not impossible, to convey in writing the experience of hearing music live (Abbate 2004). A chance to rehearse this argument to a live audience in 2012 brought with it the opportunity for a demonstration of the different listening opportunities afforded by performances and recordings (Zayaruznaya 2012). The attendees were presented with two complete renditions of *Dame/Fins cuers doulz*. The first was a recording engineered to have a mixed and resonant "early music sound." The second was live, sung by the same performers and in the same way. Afterward listeners were asked whether they heard anything in the live performance that they did not hear in the recording.

[3.8] By way of conclusion, I would like to offer the reader an analogous experience, insofar as the analogy can be carried out without recourse to liveness. Linked below are two recorded versions of *Dame/Fins cuers doulz*. Both represent the same single take. The first (**Audio Example 4**) is audio engineered to match the resonant acoustic on the Hilliards' Machaut CD.[24] The second is a video (**Video Example 1**) of this same take being recorded in the dry studio, which preserves binaural as well as visual cues. *Before reading on any further*, I invite the reader to listen to the recording, and then view the video, preferably through decent speakers or using headphones.

[CLICK TO CONTINUE READING]

Anna Zayaruznaya
Yale University
PO Box 208310
New Haven, Connecticut 06520-8310
anna.zayaruznaya@yale.edu

Works Cited

Abbate, Carolyn. 2004. "Music—Drastic or Gnostic?" *Critical Inquiry* 30 (3): 505–36.

Alain, Claude, Karen Reinke, Yu He, Chenghua Wang, and Nancy Lobaugh. 2005. "Hearing Two Things at Once: Neurophysiological Indices of Speech Segregation and Identification." *Journal of Cognitive Neuroscience* 17 (5): 811–18.

Atchison, Mary. 2005. *The Chansonnier of Oxford Bodleian MS Douce 308: Essays and Complete Edition of Texts*. Ashgate.

Bayat, Arash, Mohammad Farhadi, Akram Pourbakht, Hamed Sadjedi, Hesam Emamdjomeh, Mohammad Kamali, and Golshan Mirmomeni. 2013. "A Comparison of Auditory Perception in Hearing-Impaired and Normal-Hearing Listeners: An Auditory Scene Analysis Study." *Iranian Red Crescent Medical Journal* 15 (11).

Bent, Margaret. 1990. "A Note on the Dating of the Trémoïlle Manuscript." In *Beyond the Moon: Festschrift Luther Dittmer*, ed. Luther A. Dittmer, Bryan Gillingham, and Paul Merkley, 217–42. Institute of Mediaeval Music.

—————. 1993. "Reflections on Christopher Page's *Reflections*." *Early Music* 21 (4): 625–33.

—————. 1998. "Fauvel and Marigny: Which Came First?" In *Fauvel Studies: Allegory, Chronicle, Music, and Image in Paris—Bibliothèque Nationale de France, MS français 146*, ed. Margaret Bent and Andrew Wathey, 35–52. Clarendon Press.

—————. 2003. "Words and Music in Machaut's 'Motet 9'." *Early Music* 31 (3): 363–88.

Boogaart, Jacques. 2001. "Encompassing Past and Present: Quotations and Their Function in Machaut's Motets." *Early Music History* 20: 1–86.

Bragard, Roger, ed. 1955–1973. *Jacobi Leodiensis Speculum musicae*. 7 vols. American Institute of Musicology.

Bregman, Albert S. 1990. *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press.

—————. 2015. "Progress in Understanding Auditory Scene Analysis." *Music Perception* 33 (1): 12–19.

Busse Berger, Anna Maria. 2005. *Medieval Music and the Art of Memory*. University of California Press.

Cardiff, Janet. 2001. *The Forty Part Motet (A Reworking of* Spem in alium nunquam habui *by Thomas Tallis, 1556)*. Collection of The Museum of Modern Art (MoMA).

Cherry, Edward Colin. 1953. "Some Experiments on the Recognition of Speech, with One and with Two Ears." *Journal of the Acoustical Society of America* 25 (5): 975–79.

—————. 1957. *On Human Communication: A Review, A Survey, and A Criticism*. MIT Press.

Clark, Alice V. 2004. "Listening to Machaut's Motets." *Journal of Musicology* 21 (4): 487–513.

Clark, Suzannah. 2007. "'S'en dirai chançonete': Hearing Text and Music in a Medieval Motet." *Plainsong and Medieval Music* 16 (1): 31–59.

Cramer, Alfred. 2013. "The Harmonic Function of the Altered Octave in Early Atonal Music of Schoenberg and Webern: Demonstrations Using Auditory Streaming." *Music Theory Online* 9 (2).

Coulter, John. 2010. "Electroacoustic Music with Moving Images: The Art of Media Pairing." *Organised Sound* 15 (1): 26–34.

Davis, Dan. 2010. Editorial review of Hilliard 2004, Amazon.com. http://www.amazon.com/Machaut-Motets-Guillaume/dp/B0001A7D3G.

Davison, Lynn L., and William P. Banks. 2003. "Selective Attention in Two-Part Counterpoint." *Music Perception* 21 (1): 3–20.

Dillon, Emma. 2002. *Medieval Music-Making and the "Roman de Fauvel."* Cambridge University Press.

Drennan, Ward R., Stuart Gatehouse, and Catherine Lever. 2003. "Perceptual Segregation of Competing Speech Sounds: The Role of Spatial Location." *Journal of the Acoustical Society of America* 114 (4.1): 2178–89.

Dupoux, Emmanuel, Sid Kouider, and Jacques Mehler. 2003. "Lexical Access Without Attention? Explorations Using Dichotic Priming." *Journal of Experimental Psychology: Human Perception and Performance* 29 (1): 172–84.

Earp, Lawrence. 1995. *Guillaume de Machaut: A Guide to Research*. Garland.

Harrison, Gregory A. 1963. "The Monophonic Music in the *Roman de Fauvel*." PhD diss., Stanford University.

Haykin, Simon, and Zhe Chen. 2005. "The Cocktail Party Problem." *Neural Computation* 17 (9): 1875–1902.

Hentschel, Frank. 2001. "Der Streit um die 'ars nova'—nur ein Scherz?" *Archiv für Musikwissenschaft* 58 (2): 110–30.

Huot, Sylvia. 1987. *From Song to Book: The Poetics of Writing in Old French Lyric and Lyrical Narrative Poetry*. Cornell University Press.

—————. 2003. "Reading Across Genres: Froissart's *Joli buisson de jonece* and Machaut's Motets." *French Studies* 57 (1): 1–10.

Huron, David. 1989. "Voice Denumerability in Polyphonic Music of Homogeneous Timbres." *Music Perception* 6 (4): 361–82.

—————. 2001. "Tone and Voice: A Derivation of the Rules of Voice-leading from Perceptual Principles." *Music Perception* 19 (1): 1–64.

Iverson, Jennifer. 2011. "Creating Space: Perception and Structure in Charles Ives's Collages." *Music Theory Online* 17 (2).

James, William. 1890. *The Principles of Psychology*. H. Holt.

Kane, Brian. *Sound Unseen: Acousmatic Sound in Theory and Practice*. Oxford University Press.

Kerber, Stefan, and Bernhard U. Seeber. 2012. "Sound Localization in Noise by Normal-Hearing Listeners and Cochlear Implant Users." *Ear and Hearing* 33 (4): 445–57.

Kondo, Hirohito M., Anouk M. van Loon, Jun-Ichiro Kawahara, and Brian C. J. Moore. 2017. "Auditory and Visual Scene Analysis: An Overview." *Philosophical Transactions of the Royal Society B: Biological Sciences* 372 (1714). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5206268/

Kügle, Karl. 1991. "Die Musik des 14. Jahrhunderts: Frankreich und sein direkter Einflussbereich." In *Die Musik des Mittelalters*, ed. Hartmut Möller, Rudolf Stephan and Dorothea Baumann, 352–84. Laaber-Verlag.

——————. 1993. *The Manuscript Ivrea, Biblioteca capitolare 115: Studies in the Transmission and Composition of Ars Nova Polyphony*. Institute of Mediaeval Music.

la Buigne, Gace de. 1951. *Le Roman des Deduis*. Edited by Åke Blomqvist. E.G. Johanssons Buchdruckerei.

Leach, Elizabeth Eva. 2006. "'The Little Pipe Sings Sweetly while the Fowler Deceives the Bird': Sirens in the Later Middle Ages.' *Music and Letters* 87 (2): 187–211.

——————. 2007. *Sung Birds: Music, Nature, and Poetry in the Later Middle Ages*. Cornell University Press.

Leech-Wilkinson, Daniel. 1984. "Machaut's 'Rose, Lis' and the Problem of Early Music Analysis." *Music Analysis* 3 (1): 9–28.

——————. 2002. *The Modern Invention of Medieval Music: Scholarship, Ideology, Performance*. Cambridge University Press.

——————. 2003. "*Rose, lis* Revisited." In *Machaut's Music: New Interpretations*, ed. Elizabeth Eva Leach, 249–62. Boydell Press.

Leo, Domenic. 2005. "Authorial Presence in the Illuminated Machaut Manuscripts." PhD diss., New York University.

Lerch, Irmgard. 1987. *Fragmente aus Cambrai: Ein Beitrag zur Rekonstruktion einer Handschrift mit spätmittelalterlicher Polyphonie*. Bärenreiter.

Lin, I-Fan, Aya Shirama, Nobumasa Kato, and Makio Kashino. 2017. "The Singular Nature of Auditory and Visual Scene Analysis in Autism." *Philosophical Transactions of the Royal Society B: Biological Sciences* 372 (1714).

Lorris, Guillaume de, and Jean de Meun. 1994. *The Romance of the Rose*, trans. Frances Horgan. Oxford University Press.

Ma, Wei Ji, Xiang Zhou, Lars A. Ross, John J. Foxe, and Lucas C. Parra. 2009. "Lip-Reading Aids Word Recognition Most in Moderate Noise: A Bayesian Explanation Using High-Dimensional Feature Space." *PLoS ONE* 4 (3): e4638.

McAdams, Stephen, and Albert S. Bregman. 1979. "Hearing Musical Streams." *Computer Music Journal* 3 (4): 26–43.

McGee, Timothy James. 1998. *The Sound of Medieval Song: Ornamentation and Vocal Style According to the Treatises*. Clarendon Press.

McGurk, Harry, and John MacDonald. 1976. "Hearing Lips and Seeing Voices." *Nature* 264: 746–48.

Page, Christopher. 1977. "Machaut's 'Pupil' Deschamps on the Performance of Music: Voices or Instruments in the 14th-Century Chanson." *Early Music* 5 (4): 484–91.

——————. 1982. "The Performance of Songs in Late Medieval France: A New Source." *Early Music* 10 (4): 441–50.

——————. 1988. "The Performance of Ars Antiqua Motets." *Early Music* 16 (2): 147–64.

——————. 1993. *Discarding Images: Reflections on Music and Culture in Medieval France*. Clarendon Press.

——————. 1994. "A Reply to Margaret Bent." *Early Music* 22 (1): 127–32.

——————. 2000. "Around the Performance of a 13th-Century Motet." *Early Music* 28 (3): 343–58.

Pelofi, Claire, Vincent de Gardelle, Paul Egré, and Daniel Pressnitzer. 2017. "Interindividual Variability in Auditory Scene Analysis Revealed by Confidence Judgements." *Philosophical Transactions of the Royal Society B: Biological Sciences* 372 (1714).

Pesce, Dolores. 1986. "The Significance of Text in Thirteenth-Century Latin Motets." *Acta Musicologica* 58 (1): 91–117.

Potter, John. 1998. *Vocal Authority: Singing Style and Ideology*. Cambridge University Press.

Saltzstein, Jennifer. 2013a. "Ovid and the Thirteenth-Century Motet: Quotation, Reinterpretation, and Vernacular Hermeneutics." *Musica Disciplina* 58: 351–72.

Saltzstein, Jennifer. 2013b. *The Refrain and the Rise of the Vernacular in Medieval French Music and Poetry*. D. S. Brewer.

Sanders, Ernest H. 1975. "The Early Motets of Philippe de Vitry." *Journal of the American Musicological Society* 28 (1): 24–45.

Schrade, Leo. ed. 1956. *The Works of Guillaume de Machaut*. Vols. 2–3, Polyphonic Music of the Fourteenth Century. Éditions de l'Oiseau-Lyre.

Smith, Christine and Joseph O'Connor. 2006. *Building the Kingdom: Giannozzo Manetti on the Material and the Spiritual Edifice*. Arizona Center for Medieval and Renaissance Studies.

Stone, Anne. 2003. "The Composer's Voice in Late-Medieval Song: Four Case-Studies." In *Johannes Ciconia: Musicien de la transition*, ed. Philippe Vendrix, 169–94. Brepols.

Summerfield, Quentin. 1992. "Lipreading and Audio-Visual Speech Perception." *Philosophical Transactions: Biological Sciences* 335 (1273): 71–8.

Taruskin, Richard. 2005. *The Oxford History of Western Music*. 5 vols. Oxford University Press.

Vecchi, Giuseppe. 1967. "Musica e scuola delle artes a Bologna nell'opera di Boncompagno da Signa (saec. XIII)." In *Festschrift Bruno Stäblein zum 70. Geburtstag*, ed. Martin Ruhnke, 266–73. Bärenreiter.

Wathey, Andrew. 1993. "The Motets of Philippe de Vitry and the Fourteenth-Century Renaissance." *Early Music History* 12 (1993): 119–50.

——————. 1994. "The Motet Texts of Vitry in German Humanist Manuscripts." In *Music in the German Renaissance: Sources, Styles, and Contexts*, ed. John Kmetz, 195–201. Cambridge University Press.

Wegman, Rob C. 1995. "Reviewing Images." *Music & Letters* 76 (2): 265–73.

——————. 1996. "From Maker to Composer: Improvisation and Musical Authorship in the Low Countries, 1450-1500." *Journal of the American Musicological Society* 49 (3): 409–79.

Wimsatt, James I. 1991. *Chaucer and his French Contemporaries: Natural Music in the Fourteenth Century*.

University of Toronto Press.

Wood, Noelle, and Nelson Cowan. 1995. "The Cocktail Party Phenomenon Revisited: How Frequent are Attention Shifts to One's Name in an Irrelevant Auditory Channel?" *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21 (1): 255–60.

Wright, James K., and Albert S. Bregman. 1987. "Auditory Stream Segregation and the Control of Dissonance in Polyphonic Music." *Contemporary Music Review* 2 (1): 63–92.

Zayaruznaya, Anna. 2004. "Lies, Damned Lies, and Hockets: Words and Music in Machaut's Motet 14." Paper presented at the fall meeting of the New England chapter of the American Musicological Society, Amherst.

—————. 2009. "'She has a Wheel that Turns. . .': Crossed and Contradictory Voices in Machaut's Motets." *Early Music History* 28: 185–240.

—————. 2012. "Motets and the Modern Medieval Sound: *Ars nova* Intelligibility Redux." Paper presented at "After the End of Music History: An International Conference in Honor of Richard Taruskin." Princeton University, February 9–12.

—————. 2013. "Hockets as Compositional and Scribal Practice in the *ars nova* Motet—A Letter from Lady Music." *Journal of Musicology* 30 (4): 461–501.

—————. 2015a. *The Monstrous New Art: Divided Forms in the Late Medieval Motet*. Cambridge University Press.

—————. 2015b. "Quotation, Perfection and the Eloquence of Form: Introducing *Beatius*/*Cum humanum*." *Plainsong & Medieval Music* 24 (2): 129–66.

Discography

Ensemble Musica Nova. 2002. *Guillaume de Machaut, Intégrale des motets*. Compact disc. Zig-Zag Territoires ZZT 021002.

Hilliard Ensemble. 2004. *Motets: Guillaume de Machaut*. Compact disc. ECM Records ECM 1823.

Studio der frühen Musik, dir. Thomas Binkley. 1972. *Guillaume de Machaut: Chansons II*. Recording released in three formats: EMI "Reflexe" 1C 063-30 109 [LP-Stereo]; EMI "Reflexe" 1C 263-30 109 [Cass.-Stereo]; EMI "Reflexe" CDM or 555 7 63424 2 [CD].

Footnotes

* My most profound thanks to those who have helped render this study intelligible: Ardis Butterfield, Jeannette DiBernardo Jones, Henry Parkes, Brian Kane, the members of the Medieval Song Lab at Yale, two thoughtful anonymous reviewers for this Journal, and my three wise singers —Ariadne Lih, Sylvia Leith, and William Watson.
Return to text

1. The Cocktail Party Effect, discussed in detail below, has been an object of study in physiology, neurobiology, psychophysiology, cognitive psychology, biophysics, computer science, and engineering. For a review, see Haykin and Chen 2005. Primates, bats, fish, amphibians, and birds as diverse as penguins and owls have all been found capable of segregating intermingled sounds (Bregman 2015, 14).
Return to text

2. See, e.g., Page 1993 and 1994, Bent 1993 and 2003, Clark 2007, and Zayaruznaya 2015a.
Return to text

3. In the course of defining discanting [i.e. upper] voices in comparison and in contradistinction to tenors, Jacobus notes that "they can be considered by themselves, not with respect to the notes of the tenor and sung at the same time as them, but separately and in turn, one after the other, *as when someone sings some motetus or triplum or quadruplum by himself without a tenor*" (Possunt autem voces discantus ad voces comparari tenoris. . .vel possunt per se considerari, non per respectum ad voces tenoris et ut simul dicuntur cum illis, sed divisim successive una post aliam cum aliquis per se cantat motetum aliquem, triplum vel quadruplum, sine tenore), Bragard 1955–73, 7:9. Clearly the practice of singing voices on their own was nothing too out of the ordinary, since Jacques is citing it here not for its own sake, but to support a broader claim that the voices of polyphonic compositions can each be considered as independent songs on their own terms. Literary support for this practice comes from Froissart's *Joli Buisson de Jonece* (1373), whose narrator gives voice to his joyful mood by singing "a new motet" that he had been sent from Reims ("En chantant un motet nouviel/Quon mavoit envoiiet de Rains"); see Huot 2003. Further evidence of this independence of upper voices comes from the interpolated *Fauvel*, where the triplum voice of the motet *Floret/Florens*, which survives in three voices in the Brussels rotulus and Cambrai fragments, is reworked into a monophonic prose beginning *Carnalitas luxuria*. The prose is edited in Harrison 1963, 188–90; *Floret/Florens* in Sanders 1975, 37–45 and Lerch 1987, 2: 205–14. For a hypothesis on why the prose and not the motet appears in *Fauvel*, see Bent 1998, 46–47.
Return to text

4. "Vidi ergo, in quadam societate, in qua congregati erant, valentes cantores et laici sapientes. Fuerunt ibi cantati moteti moderni et secundum modum modernum, et veteres aliqui. Plus satis placuerunt, etiam laicis, antiqui quam novi, et modus antiquus quam novus. Etsi enim modus novus in sua placuerit novitate, non sic modo, sed incipit multis displicere. Ergo iam placeat cantus antiquos et cantandi notandique modum antiquum ad patriam revocare cantorum. Revertantur haec ad usum refloreatque rationabilis ars. Exul fuit et eius cantandi modus. Quasi violenter deiecta sunt haec a cantorum consortio. Violentum autem non debet esse perpetuum. Ad quid tantum placet cantandi lascivia, curiositas in qua, ut aliquibus videtur, *littera perditur*, harmonia consonantiarum minuitur, valor notarum mutatur, perfectio deprimitur, imperfectio sublimatur mensuraque confunditur? Vidi in magna sapientium societate, cum cantarentur moteti, secundum modernum modum, *quesitum fuit, quali lingua tales uterentur cantores, Ebrea, Greca, vel Latina, vel qua alia, quia non intelligebatur quid dicerent. Sic moderni, licet multa pulchra et bona in suis cantibus faciant dictamina, in modo tamen suo cantandi cum non intelligantur, perdunt ea*." Bragard 1955–73, 7:95; emphasis mine.
Return to text

5. That is: most modern listeners would agree that text is easier to understand in *ars nova* than in *ars antiqua* motets, given the differentiated rates of text delivery in the former (Zayaruznaya 2009 and 2013). The fact that Jacobus makes the opposite argument suggests to me that intelligibility of text is not (or not primarily, and certainly not exclusively) a function of style, but of delivery—hence my focus here on performance and the deliberate blurring of lines between the two repertories.
Return to text

6. Of course there would have been other kinds of interactions available to the medieval consumers of motets (e.g. Dillon 2002), such that their "medieval audience" is a much broader category than those of "medieval listeners" or "medieval singers." But below I mostly use the term "audience" in its most literal sense, as involving audition. See also the theorization and defense of listening as an activity in Clark 2004, where the chief concern is with the intelligibility of structure rather than of text.

7. For example, reading Bent 2003 makes us into listeners prepared, in her sense, to hear Machaut's *Fons/O livoris* (M9).

8. See also the evaluation of Machaut on p. 14: "Machaut's music leaves no doubt that his sensations when composing were as indifferent to moral or intellectual persuasions as those of any composer at any period in history when genuinely engrossed." I find Page's arguments more persuasive for *ars antiqua* than for *ars nova* motets, and admittedly the former are his explicit subject. But at times, for example in the longer section surrounding the passage quoted above, he shifts from discussing thirteenth-century listeners to the broader "medieval motet" almost seamlessly.

9. On the fraught question of authenticity and its role in Page's argument, see also Wegman 1995, 269–70.

10. On "tenorista" as function rather than range, see Wegman 1996, 445–8.

11. That there was some non-clerical performance of mensural polyphony is indicated by the depictions of lay singers in the Machaut manuscripts (e.g. the singers gathered around a barrel, discussed below). The narrator of Machaut's *Voir dit* sends notated music to the noble Toute Belle, expecting her to perform it. For further evidence that women performed mensural polyphony, see Leach 2006.

12. That the composition being performed is a motet is clear not only from the miniature's position at the head of the motet section, but also from the words "tenor" and "dame" that are visible on the scroll. This image is usually interpreted as a group of singers (Huot 1987, 300–1; Earp 1995, 187; Wimsatt 1991, plate 13; Leo 2005 44). Kügle has argued that the four men on the right are singing, while those on the left (a layman, a cleric, and a servant) are drinking and listening (1991, 361; note that in Kügle's essay the image descriptions and images are reversed, so that the description labeled "Bild links" pertains to the motet image, which is on the right). Kügle's analysis is convincing as pertains to the direction of the scroll, but the nobleman at left does seem to be keeping tactus on the arm of the cleric to his left, and the latter's holding the scroll seems to imply his participation.

13. The same iconographic motif—singers around a wine keg—is also used to illustrate the singing of a rondeau in Bibliothèqe nationale de France, MS français 9221 (Machaut MS E), fol. 16. There the singers seem all to be laypersons, but they still vary in age and class; see Kügle 1991, 361.

14. Though the sound is natural, it is not rational, and Leach has argued (2007, 212–19) that this passage is part of a broader discourse about *vox confusa* and *vox discreta*—the dogs are *cantores* but not *musici*.

15. Page 1982, 442. Cf. Jerome of Moravia's advice that "a song [should] never be begun very high, especially by those possessing head voices. . .but they should always begin moderately, that is 'sing'; that is so that the song may not rule the voice, but the voice rule the song; beautiful notes

cannot be produced otherwise" (trans. Randall A. Rosenfeld, quoted in McGee 1998, 28).
Return to text

16. This is suggested by the brevity of many 13th- and 14th-century motets, as well as their long careers; see, for example, the inclusion of *Les l'ormel/Mayn* and *Clap/Sus Robin* in the late Ivrea codex, and more generally the retrospective nature of that collection and of the Trémoïllé index (Kügle 1993, 90–91; Bent 1990). See also the discussion of a motet sung during a courtly procession, and thus presumably sung by heart, in Huot 2003.
Return to text

17. Kondo et al. (2017) note that although "perception is a private process for each individual, and perceptual experiences may differ across individuals even when the physical environment is the same," studies often ignore such differences: "Individual differences in human behaviour and perception are often considered to be 'noise' and are therefore discarded through averaging data from a group of participants."
Return to text

18. Bregman 1990, 19-20, 85–6, and 92–103. Though "timbre" is an inexact term—Bregman calls it "an ill-defined wastebasket category," that "can include fundamental frequency, spectral component frequency, the balance of the spectrum, and possibly individual peaks in the spectrum" (1990, 92)—it is also an intuitive one. The difficulty of quantifying timbre and of producing maximally different timbres is cited by Huron as a reason to explore the limits of perception for polyphony performed in identical timbres (on solo organ); but he notes that "entries employing the pedal division of the organ were significantly easier to hear than matched nonpedal entries. . . because the difference in timbre between the pedal and manual parts is musically rather modest, this result suggests that timbre may be a very strong factor in the identification of concurrent polyphonic voices" (1989, 379).
Return to text

19. Alain et. al 2005. Studies that explicitly connect music theory with auditory scene analysis include McAdams and Bregman 1979, Wright and Bregman 1987, Huron 2001, Cramer 2013, and Iverson 2011.
Return to text

20. Let the record state that Head-Related Transfer Function would be an excellent band name.
Return to text

21. The claim of a common biological makeup is, I think, less subject to accusations of universalism than invocation of a "trans-historical humanness" or trans-historical aesthetics; see Page 1993, 190, and Wegman 1995, 270.
Return to text

22. The counterpart to acousmatic listening is not live, but "direct" listening, which affords the listener-viewer a correlation between the visual and the aural, so that we do not find ourselves guessing about the bodies whose voices we hear.
Return to text

23. As one reviewer (Davis 2010) wrote, "the engineers beautifully capture the glow of the church acoustic." Indeed they do, but these are secular works—many of them love songs—that would probably have been sung not in a church, but in a chamber or small hall. The recording includes eighteen motets, of which four can be classified as ceremonial, and two more as, perhaps, paraliturgical. These latter works might conceivably have been sung in a larger space, but that is a matter of conjecture.

24. The recording and post-production processing were done by Greg DiCrosta, who provided the following comments on his approach: "The recording approach was a combination of a typical *a cappella* setup, combined with a classic acoustic instrument technique. Large diaphragm condenser microphones (AKG 414) were used as the close-up spot microphones for the individual voices. In addition, an AB omni stereo pair of microphones (DPA 4052) were used to capture the live blend of the singers at a slight distance. The cleanest possible preamplifiers were used (GML 8304 and Millenia HV-3D) in order to capture the cleanest, clearest signal possible. Recording was to Pro Tools HDX via Prism Dream ADA-8XR converters. In mixing, no equalization or dynamics processing was used, again to keep processing to a minimum. An amount of reverb was added in order to simulate spaces. Two versions were created, simulating what would be considered "wet" (more reverb) and "dry" (less reverb). The wet mix simulates... a large church or concert hall-type environment. The dry version used as little reverb as possible, with extreme panning (left to right placement) of the voices. Altiverb was used." E-mail communication, 13 March 2016.

25. Page's reading of Jacobus's recommendation for the voices of those singing together to be "voces. . . concordantes" as "voices that blend" notwithstanding (Page 2000, 355); this "concordantes" could as easily refer to the voices' abilities to be in tune with each other, and indeed the context of the passage encourages the latter interpretation (Bragard 1955–73, 1965, iv.12).

26. An example of this approach on a much larger scale is Janet Cardiff's 2001 installation *The Forty-Part Motet*, which played a version of Thomas Tallis's *Spem in alium* through 40 speakers.

27. Kane rightly points out that "acousmatic listening was alive and well, even in the eras when the term 'acousmatic' did not exist" (2014, 9) and that it certainly predates recording technology. Acousmatic sounds were demonstrably present for medieval listeners, but that they were not the norm, or perhaps not inherently sufficient, can be gleaned from some reactions to them. When the Amant of the *Rose* hears the birds singing inside the garden which he has yet to penetrate, he comments on "the harmony of their moving songs. . . .most beautiful to hear" at which "the whole world must rejoice." But the beauty of these sounds only makes him want to see their sources: "I was filled with such joy when I heard it that not for a hundred pounds, if the way in had been open, would I have failed to enter and see the birds assembled there" (de Lorris and de Meun 1994, 9). Similarly, when the narrator of Ciconia's *Sus un fontayne* hears sweet singing that captivates his body and mind, his immediate desire is to see the singer (see the text printed in Stone 2003, 172). Acousmatic sounds may have excited less curiosity in churches, whose layout throws sounds around the corners of chapels, and where screens occluded some singers. For example, Giannozzo Manetti's description of the dedication of the Florentine Duomo, on which occasion Du Fay's *Nuper rosarum flores* was performed, is reminiscent of the sonic reification that results from indirect listening: "melodies were raised by so many and varied singing voices, alternating with songs made with such symphony lifted up to heaven, that to the audience they appeared for sure angelic and divine" (tantis tamque variis canoris vocibus quandoque concinebatur, tantis etiam symphoniis ad celum usque elatis interdum cantabatur ut angelici ac divini cantus nimirum audientibus apparerent), Smith and O'Connor 2006, 320–21. On angelic choirs, see also Kane 2014 108–13.

Copyright Statement

Prepared by Samuel Reenan, Editorial Assisitant