(De)Constructing Timbre at NIME: Reflecting on Technology and Aesthetic Entanglements in Instrument Design

Charalampos Saitis*
Centre for Digital Music
Queen Mary University of London
London, UK
c.saitis@qmul.ac.uk

Courtney N. Reed
Institute for Digital Technologies
Loughborough University London
London, UK
c.n.reed@lboro.ac.uk

Ashley Noel-Hirst
Centre for Digital Music
Queen Mary University of London
London, UK
a.l.noel-hirst@qmul.ac.uk

Giacomo Lepri
InfoMus - Casa Paganini
University of Genoa
Genoa, Italy
giacomo.lepri@edu.unige.it

Andrew McPherson
Dyson School of Design Engineering
Imperial College London
London, UK
andrew.mcpherson@imperial.ac.uk

Abstract

Timbre, pitch, and timing are often relevant in digital musical instrument (DMI) design. Amongst the three, timbre is the most difficult to define and discretise when negotiating audio representations and gesture-sound mappings. We conduct a corpus-assisted discourse analysis of "timbre" in all NIME proceedings to date (2001–2024). Combining this with a detailed review of 18 timbre-focused papers, we examine how definitions of timbre and timbre interaction methods are constructed through, for instance, Wessel's numerical timbre control space, synthesis tools and programming languages, machine learning and AI approaches, and other trends in digital lutherie practices. While acknowledging the practical utility of technical constructions of timbre in NIME (and other digital music research communities), we contribute discussion on the entanglement of technology and aesthetics in instrument design, which constitutes what "timbre" becomes in NIME research, and reflect on the tension between technoscientific and constructivist understandings of timbre: how DMIs and musical practices have been reconstituted around particular timbral values operationalised in NIME. In response, we propose ways that the community can embrace more critical approaches and awareness to how our methods and tools shape and co-create our notions of timbre, as well as other musical concepts, connecting more openly with diverse types of sonic phenomena.

Keywords

Timbre, Entanglement, Metaphor, Reification, Constructivism

> … that kind of zone that exists between noise and pitch … I think that there's a big spectrum there, in terms of being able to define the pitch of something as it moves from white noise to a sine wave … and that's the world, for me, that's the world of timbre …
>
> Anonymous, interview with first author, January 2024

*Corresponding author.

NIME '25, June 24–27, 2025, Canberra, Australia

1 Introduction

Timbre matters in musical practice but is difficult to work with. It has been called a sonic "wastebasket," full of what remains beyond pitch, loudness and timing when we experience musical sound [90]. Compared to these aspects, timbre can be defined in many ways; for example, the "quality" or "colour" of a sound (aesthetics), an instrument's identity (perception), individual performers' characteristics (articulation), frequency distribution (acoustics), and vocal formants (physiology). When creating digital musical instruments or interfaces (DMIs), the designer is required to manipulate digital technology as an additional medium for crafting sonic material and interactions [66]. The International Conference on New Interfaces for Musical Expression (NIME) and related communities of digital lutherie and sonic creativity have researched or engaged with timbre in the pursuit for new sounds and musical experiences. However, timbre's ambiguity brings the question of how "successful" DMI designers have been at designing with and for timbre.

We examine timbre in NIME research to better understand how its different views and conceptualisations have been enacted in DMI design and, as a result, musical practices. We analyse previous NIME proceedings (2001–2024) through a corpus-assisted discourse analysis of "timbre." Parts of the analysis draw comparisons with past proceedings of the International Computer Music Conference (ICMC 1975–2024), one of the main pre-NIME (before 2001) and off-NIME (since 2001) venues for disseminating research on the use of computers in music [89]. We further review a subcorpus of 18 NIME papers that specifically denote timbre in their titles to understand timbre's associations, the tools and parameters used to interact with timbre, and some of the timbre goals in the NIME community.

This analysis exhibits two main findings about timbre and NIME. First, the community has focused largely on *control*: there is no clear aesthetical approach to timbre; however, NIME has very specific ways to work with and manipulate timbre in musical interaction (cf. [24]). These methods originate from technoscientific viewpoints and values originating in music information retrieval (MIR), psychoacoustics, and digital systems research that have codified timbre's ambiguity into particular features [57]. These features have come to act as *proxies* for timbre and its perceptual experience. Second, the way that NIME operationalises timbre is *entangled* with musical practices, aesthetics, and digital tools. This reflects epistemological and methodological *tensions* between musicology-sound-timbre studies and perception-technology-research [58]; for example, between manipulating timbre and what Fales [27] calls "perceptualising" timbre.

Through reflection on timbre within the NIME corpus, we here examine how timbre has been constructed within the NIME community and digital musical interaction more broadly. We offer three main considerations for the community and its relationship with timbre. Examining and potentially challenging the way we conceptualise timbre, we suggest some avenues for DMI designers to pursue novel dimensions of timbre interaction:

(1) We unpack the NIME corpus to understand influences in timbre's use in our community. We identify the influence of conceptual timbre frameworks, links to MIR research trends, and tools used in designing with/for timbre.
(2) We discuss how *timbre thinking* is entangled with *design thinking* and how timbre has been codified through the influence of technological practices. We suggest how NIME research is at risk of being *both* reductive and vague about timbre, trapped in technology advancements that see timbre reified as a handful of extractable features.
(3) We provide *six points of tension* between timbre ideologies as new avenues for research at NIME and beyond (music psychology, sound studies, MIR). We discuss whether timbre is an object or property to be treated by mapping, or something that emerges from mappings and the behaviour they engender. We suggest future work to expand our ideas of what timbre is and what we can do with it.

2 Conversation Starters

2.1 Timbre Defined, Timbre Ambiguous

The concept of timbre emerged as a distinct musical parameter during the 18th century, initially described through metaphorical language that emphasized its ineffable qualities. Early theorists, particularly Rousseau in his 1768 *Dictionnaire de musique* [69], characterised timbre as the distinctive "colour" or "quality" of a sound that distinguished one instrument from another, even when playing the same pitch at the same volume. Almost two centuries later, the American National Standards Institute's (ANSI) formal 1960/1994 definition of timbre described it as encompassing all perceptual attributes of sound other than pitch, loudness, and duration [3]. This definition-by-negation, while precise in what it excludes, reflects the historical challenge of positively defining what timbre is (cf. [78]).

The challenge of defining timbre extends beyond mere technical description into fundamental questions about musical aesthetics and sonic creativity. Barrière [8] positions timbre as a critical point of aesthetical and philosophical divergence—"an inevitable breaking-point"—in musical discourse, where different conceptualisations of timbre represent opposed musical worldviews and aesthetical positions. Chion [16] has gone so far as to call for the dissolution of timbre (timbre survived this attack). While the ANSI definition provided a formal framework, the challenges of meaningfully negotiating and working with timbre persist, particularly in the contexts of electroacoustic music [80], computer music [22], and digital instrument design [71] where technology and aesthetics are entangled. Specifically concerning NIME, the problem that timbre is ambiguous and underspecified might also be symptomatic of a difficulty of talking about musical aesthetics. It is much easier to talk about control structures, after all.

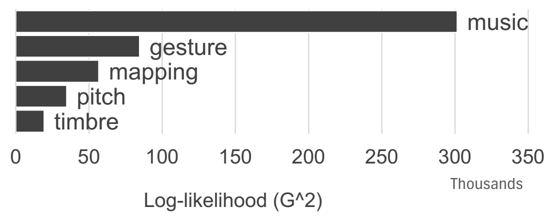

Figure 1: How *key* is timbre in NIME? Keyness evaluated using log-likelihood (G^2), comparing observed versus expected frequencies of five terms in the NIME and Elsevier corpora (χ^2, $df = 1$). All $p \ll 0.0001$.

2.2 Timbre Described, Timbre Prescribed

Technology has been central in the formation of timbre perception as a research field. Computational methods for multidimensional scaling developed in the 1960s enabled researchers to consolidate responses from listening tests into spatial representations which they could also visualise (e.g., the timbre of the English horn "being closer" to that of the bassoon than that of the trumpet when the three instruments are played at the same pitch and dynamic) [41]. Advances in digital signal processing and sound synthesis around the same time [68] made it possible to link spectral, temporal, and energy features of recorded or synthesized audio signals with the dimensions of what came to be known as "timbre space" (see [54] for a review). Since then, the timbre space model and metaphor has dominated scientific discourse on timbre [79]. It has also been built into new technologies for music creation and performance as well as into audio classification schemes and formatting standards like MPEG-7, which support the design of DMIs and music software applications [49, 57, 60].

One of the first attempts to instrumentalise timbre space was by Wessel [91]. He was interested in how its *descriptive* (cor)relations between audio features and listeners' perceived timbre similarity might be turned into a *prescriptive* and *interactive* "musical control structure" that would allow musicians to measure, navigate, and create new timbres. Mapping real-time gestures to timbre in high-dimensional control spaces has since been a major staple of computer music and instrument design research [88], including more recently "latent timbre spaces" learned by generative artificial intelligence (AI) models for audio synthesis [28, 56, 62, 77]. Other approaches include generating musical materials directly from timbre space [74], or else using timbre space to condition generative AI models [26].

2.3 Timbre Represented, Timbre Entangled

In investigating the usages and possible definitions of the word *timbre*, we highlight but neither endorse nor reject the representationalism widely found in the sciences: "the power of words to mirror preexisting phenomena" [7]. Representationalism underpins the common assumption that there already exists in the world some property of sound (whether physical or perceptual) that we might control or manipulate, and that the word "timbre" (and/or a set of numeric descriptors) act as pointers to that property. Numerical timbre space models [54] and semantic descriptors [73] are just two of many possible representations that are proposed to stand in for that ideal underlying property.

Alternatively, relational ontologies such as Barad's agential realism [7] consider timbre (and any other concept) to be a "specific material configuration", not something preordained at all, but enacted through designing, performing, listening, and other acts of musicking. If we follow that line of thought, we could propose that the instrument doesn't manipulate timbre as some independent, well-defined entity, nor even *have* timbre as a context-independent property, so much as the instrument *is* timbre; or rather, "timbre" acquires its meaning in relation to the specific *phenomenon* (in Barad's sense of the term) encompassing the relationships between physical materials, code, player, listener, musical context, etc. The implication would be that we should not go chasing universals about what a term could mean for a community overall, but rather should attend to specificities and effects of difference across practices, a point aligned with recent critical discourse in NIME and the broader human-computer interaction (HCI) community [50, 67, 83, 97].

3 When NIME Talks Timbre

We considered the entire collection of NIME proceedings from 2001 to date, which are freely available for download in PDF format from the conference's permanent website.[1] The total number of papers listed on the NIME website is 2,204. For some years, these include keynote talks and music or installation submissions, which we opted to omit from analysis. PDF files were converted to plain text using the python library `PyMuPDF`. Any text before the first heading (usually Introduction), including the Title and Abstract, was discarded, as were headers and any References or Acknowledgments sections. Thirty PDF documents could not be parsed, leaving a total of 2,132 text files to be analysed, or a lexical corpus of 7M tokens (6,870,080). Corpus-assisted discourse analysis [1] was conducted using LancsBox[2] [12] and Sketch Engine[3] [46], as well as manual keyword search.

3.1 Keyness

A keyword is a word that is statistically more frequent in a "focus" corpus in comparison to another, typically much larger "reference" corpus [6]. We first performed a keyness analysis of the terms "timbre," "pitch," "gesture," "mapping," and "music" to establish how distinctive the NIME bibliography (focus corpus) is in terms of the frequency of occurrence of these words.

"Music" was used as a control term; in keyword search (see next section), we expected it to show up in all of the papers [44]. Pitch has long been favoured over timbre in music scholarship (though this is changing [25]), which often informs research practices in digital instrument design, music informatics, and machine listening. Therefore, we expected "pitch" to occur more frequently than"timbre" in the two corpora. This is indeed what we observed. Finally, "gesture" and "mapping" are two key concepts in DMI design that are similarly difficult to negotiate, their existence being lodged between technical-physical and metaphorical-musical definitions and tensions [44, 56]. This is not to suggest that pitch is a well-defined musical concept that resists epistemological and methodological tension—think of tuning [75] and noise [21], for example—and even "music" itself is constantly negotiated across disciplines [92] and cultures [86]. While our focal point is timbre, its ambiguity connects to a broader pattern of technology and aesthetic entanglements in music and HCI [58], including timing/rhythm [19] and genre [37].

[1]https://www.nime.org/archives/
[2]https://lancsbox.lancs.ac.uk/
[3]https://sketchengine.eu/

Table 1: Percentage of papers containing keyword terms at each NIME conference to date (body text only).

Year	Papers #	Music %	Timbre %	Pitch %	Gesture %	Mapping %
2001	12	91.7	50.0	75.0	75.0	91.7
2002	46	100.0	52.2	65.2	67.4	67.4
2003	48	100.0	47.9	60.4	72.9	75.0
2004	54	98.1	22.2	55.6	55.6	55.6
2005	75	91.9	29.7	43.2	60.8	60.8
2006	74	100.0	39.2	66.2	70.3	67.6
2007	90	95.6	38.9	57.8	64.4	58.9
2008	82	96.3	23.2	53.7	62.2	56.1
2009	85	95.3	24.7	48.2	54.1	60.0
2010	110	98.2	34.5	63.6	60.9	56.4
2011	129	97.7	37.2	49.6	64.3	59.7
2012	128	98.4	43.8	54.7	64.8	64.1
2013	118	96.6	32.2	56.8	59.3	65.3
2014	148	98.0	37.8	56.8	65.5	56.1
2015	102	100.0	33.3	57.8	55.9	58.8
2016	86	100.0	41.9	61.6	57.0	62.8
2017	104	99.0	37.5	55.8	64.4	54.8
2018	91	97.8	38.5	52.7	65.9	50.5
2019	88	100.0	36.4	58.0	61.4	64.8
2020	126	96.8	42.9	61.9	64.3	65.1
2021	88	98.9	30.7	58.0	53.4	59.1
2022	55	100.0	33.9	58.9	66.1	62.5
2023	99	99.0	40.4	50.5	57.6	50.5
2024	94	100.0	37.2	56.4	63.8	62.8
mean	89	97.9	36.9	57.4	62.8	61.9
stdev	30	2.3	7.5	6.4	5.5	8.3

As reference, we considered a corpus of 40k (40,001) open access (OA) CC-BY articles from across Elsevier's journals.[4] Keyness was evaluated using the log-likelihood (G^2) test, comparing the observed versus expected frequencies of each of the five terms above in the NIME and Elsevier corpora to a χ^2 distribution with one degree of freedom. Frequency counts included plurals (e.g., "timbres") and other variations (e.g, "timbral", "timbrally"; "remapping"); for NIME, counts were obtained from LancsBox, while Sketch Engine was used for Elsevier. The G^2 threshold for keyness is typically 15.13 for $p < .0001$ [64], hence the values reported in Figure 1 are significant by orders of magnitude. We can see that "timbre" is key in NIME, but its keyness is the lowest among the tested terms.

3.2 Keyword Occurrence

We then searched through each year of NIME to return the number of papers containing the five terms above (i) generally in the main body text (i.e., excluding titles) and (ii) explicitly in the title (only the first four terms). For comparison purposes, we also searched for the same terms in the ICMC proceedings as a community more interested in aesthetics of computer music, with NIME more interested in instrument design and HCI. ICMC proceedings since 1975 are freely available (except for 1976, 1979, and 2020) but less straightforwardly downloadable. However, it is possible to manually perform keyword search online.[5]

Table 1 reports the percentage of papers containing each of the keyword terms at each NIME conference (body text only). The cumulative number of NIME and ICMC papers containing

[4]https://elsevier.digitalcommonsdata.com/datasets/zm33cdndxs/2
[5]https://quod.lib.umich.edu/i/icmc/; https://www.fulcrum.org/icmc

Table 2: Keywords in title and body text, comparing NIME vs ICMC.

Conference	Papers #	Search in	Timbre % (#)	Pitch % (#)	Gesture % (#)	Mapping % (#)
NIME 2001–2024	2,132	body text	36.5 (778)	56.5 (1205)	62.2 (1326)	60.3 (1286)
		title	0.84 (18)	0.94 (20)	7.18 (153)	3.66 (78)
ICMC 2001–2024	2,769	body text	39.9 (1104)	58.8 (1629)	34.5 (955)	34.4 (952)
		title	1.77 (49)	1.59 (44)	2.85 (79)	1.41 (39)
ICMC 1975–2024	4,862	body text	38.8 (1885)	57.6 (2799)	28.1 (1365)	27.7 (1345)
		title	1.95 (95)	1.40 (68)	2.28 (111)	1.01 (49)

all but the control terms are presented in Table 2 (body text only and title only). In NIME there are some, but very few (2%), papers that do not contain the word "music,". "Gesture" is the most commonly occurring of the terms, after "music," in agreement with previous literature [44]. "Timbre" is the least frequently occurring of the terms, used on average in 37% of all papers, with some fluctuations from year to year. On average, "pitch" as a keyword appears about 1.6x more often than "timbre" and at similar levels with "gesture" and "mapping."

In contrast, the most frequent term in ICMC is "pitch," appearing about 1.5x more often than "timbre," which in this corpus is used between 1.14x (2001–2024) and 1.4x (1975–2024) more than "gesture" and "mapping." This was to be expected, as ICMC is less focused on instruments and interaction than NIME. The distribution of "mapping" generally follows that of "gesture" across both corpora. When narrowing down the search by paper titles only, "timbre" comes up about as frequently as "pitch" in NIME (0.84% versus 0.94%). At ICMC 2001–2024 the trend is similar but reversed (1.77% versus 1.59%) and overall "timbre" appears at least twice more often than in NIME. When considering the full ICMC corpus, timbre is even more frequently used than pitch (1.95% versus 1.40%), at similar levels with gesture, and almost 2x more often used than mapping.

The fact that there are almost twice as many papers mentioning "pitch" than there are containing "timbre" in both NIME and ICMC corpora could appear to reflect what Diduck [23] calls *claviocentrism*—that current musical interaction design practices encode assumptions about musical space that are primarily based on the *analytical* idea that music is made by discrete onset and release events. This further connects [58] to the hegemony of keyboard paradigms of interaction in NIME (and ICMC as well as in commercial practice), moreover encapsulated in the MIDI (Musical Instrument Digital Interface) protocol. Puckette [63] has disclosed that the early work behind the Pd and Max software packages was inspired by the "piano metaphor... a collection of tasks running in parallel", whose timing is controlled by "wait functions and triggers."

3.3 Full Corpus Collocations

To get a sense of what conceptions, definitions, and tools NIME authors usually associate with timbre, we continued with analyses of collocation (complete NIME corpus) and concordance (focused subcorpus, see below). To identify collocates (terms appearing together with the keyword), we looked for co-occurrences within five words to the left of "timbre" and five to the right, thus identifying looser word associations than multiword expression approaches, such as n-grams or word clusters [1].

Collocates were initially determined by the logDice statistic in Lancsbox. logDice is a standardised measure for identifying co-occurrence. It expresses the typicality (or strength) of the collocation rather than its frequency, and operates on a scale with a fixed maximum value of 14, which makes it directly comparable across different corpora [33]. Comparing two scores, +1 point indicates twice as often collocation. Setting a minimum logDice value of 6 and a minimum collocation frequency of 1, a list of 444 collocates was initially obtained. These were then filtered manually to remove generic or irrelevant terms (e.g., prepositions, plurals, variations, synonyms).

Table 3 lists 72 collocates alongside their logDice statistic, co-occurrence frequency (Freq.), and distribution of textual position around "timbre," ranging from -1 (five words to the left) to +1 (five words to the right). We find several interesting observations:

- Timbre is referred to as a *(multi)dimensional space* and by its many facets (*acoustic*, *perceptual*, *semantic*, *musical*, *sonic*);
- *Space* is a key conceptual metaphor at the crossroads of timbre understanding (representation, description) and timbre interaction (control, mapping);
- Technical descriptors from psychoacoustics and music technology (e.g., *frequency*, *similarity*, *brightness*, *synthesiser*, *parameters*, *features*)—in Section 4.1.1, we identify a risk in NIME practices for parameters and features to become (technical) proxies for timbre;
- Musical-aesthetical descriptors referring to performative and artistic practice (e.g., *tone*, *noise*, *texture*, *samples*, *soundscape*, *vibrato*, *morphing*), which may also act as (aesthetical) proxies for timbre [29, 70, 80];
- Words related to timbre (space) *control* or *exploration*, through *gesture* and *(re)mapping*, and the design of *expressive* instruments with a *range* of timbral *nuances* and *possibilities*;
- Descriptors of "identity," of pursuing a *different* and *original* sound—van Elferen [87, p. 72] notes that "timbral difference is a crucial factor in musical individuation and identity" and Gooley [35] demonstrates a case of timbral difference as black identity in the adoption of specific pedalling techniques by African-American jazz pianists;
- Terms that diverge from views of timbre as something static or categorical (*dynamic*, *temporal*, *continuous*).

3.4 Focused Subcorpus Concordances

Keyword and collocate findings provide a high-level understanding of trends in corpora. Moving towards an in-depth understanding of timbre's use and conceptualisation in NIME, we next focused on a subcorpus of 18 papers (59,364 tokens) where "timbre" (also "timbres" and "timbral") is explicitly used in the title. Using LancsBox, a total of 587 concordances were found, listing all found examples of timbre within corresponding textual contexts spanning 10 words to the left and 10 to the right. Findings are detailed in Table 4 and summarised below.

Table 3: Seventy-two collocates for timbre in the full NIME corpus.

Collocate	logDice	Freq.	Position
space	9.3	183	0.63
changes	9.0	93	0.05
remapping	8.9	44	0.73
control	8.8	256	0.06
dynamic	8.5	58	0.21
variation	8.5	43	0.44
range	8.4	75	-0.41
variety	8.4	50	-0.08
features	8.3	68	0.35
parameters	8.2	84	0.05
synthesizer	8.2	38	0.42
soundscape	8.2	31	-0.55
sound	8.0	219	0.16
acoustic	8.0	53	0.09
tone	8.0	33	0.27
articulation	8.0	28	-0.14
produced	7.9	36	0.39
envelope	7.9	29	-0.03
descriptors	7.9	25	0.28
different	7.8	63	-0.52
complex	7.8	40	-0.75
harmonic	7.8	28	-0.14
frequency	7.7	43	0.07
manipulation	7.7	26	0.23

Collocate	logDice	Freq.	Position
nuances	7.7	21	0.62
instrument	7.6	117	0.18
continuous	7.6	30	-0.2
recognition	7.6	26	0.69
dimensions	7.6	25	-0.28
perceptual	7.6	21	-0.33
palette	7.6	20	0.1
exploration	7.5	26	-0.23
unique	7.5	24	-0.67
morphing	7.5	17	0.65
spatial	7.4	26	0.23
modulation	7.4	21	-0.05
multidimens.	7.4	16	-0.13
musical	7.3	117	-0.06
audio	7.3	75	-0.01
attack	7.3	18	0
generated	7.2	24	0.42
interesting	7.2	24	-0.5
possibilities	7.2	22	0.45
original	7.1	21	-0.62
spectral	7.1	18	0.78
distinct	7.1	15	-0.73
texture	7.1	15	0.07
vibrato	7.1	15	0.6

Collocate	logDice	Freq.	Position
semantic	7.1	14	-0.29
source	7.0	23	-0.39
temporal	7.0	16	0.63
selection	7.0	16	0.38
intonation	7.0	13	-0.54
gesture	6.9	32	-0.25
sonic	6.9	23	0.13
expressive	6.9	20	-0.2
samples	6.9	18	-0.11
similarity	6.9	13	0.38
navigation	6.9	12	0
brightness	6.9	12	0.67
richness	6.9	12	0.5
novel	6.8	15	-0.47
modify	6.8	12	-1
shape	6.7	16	-0.5
neural (net)	6.7	12	0.67
noise	6.6	14	0.29
character	6.6	11	0.45
categories	6.6	11	0.27
matching	6.6	10	0.8
mapping	6.5	27	-0.19
signal	6.5	24	0.33
freedom	6.1	8	-0.75

Half of the papers explicitly adopt feature extraction methods (e.g., spectral centroid, Mel-Frequency Cepstrum Coefficients or MFCCs) to work with timbre through audio representations popular with the MIR and audio signal processing communities [31, 36, 43, 45, 47, 61, 77, 82, 95]. Several papers (six out of 18, or 33.3%) focused on timbre synthesis and control through verbal attributes [45, 47, 48, 82, 84] or notation of spoken/sung vocables [55]. Soraghan et al. [82] explicitly designed a control interface based on the "luminance-texture-mass" semantic model of timbre proposed by Zacharakis et al. [96]. Lam and Saitis's [48] timbre synthesiser instrumentalises the psychoacoustical timbre space model (see Section 2). Moving from controlling pre-labelled parameters (whether signal- or word-based), timbre interaction in more recent works involves learned features and latent space manipulation, reflecting current advances in deep learning and AI technology [39, 40, 77].

Most of the papers (12 out of 18 or 66.6%) make no reference to what timbre *is* or might be, rather focusing on the control of relevant parameters from related work. Those that do attempt to describe what timbre is acknowledge the difficulty of the task: Shier et al. [77] open their paper reflecting that "Timbre is a musical concept that has distinctly resisted precise definition in psychoacoustics and music psychology research." Elsewhere, timbre is outlined via its function as a critical aspect of expression [43, 47, 77] and beneficial for the development of instrumental or sonic skills [48, 61, 81]. We find a single paper where timbre explicitly "had priority over pitch" in the design of an electroacoustic instrument [32]. The authors note that "the objects that do *not* have clear pitched sonority are the most musically interesting. This is because they allow for experimentation with more complex timbral and textural nuances." (p. 290, emphasis in original).

4 Implications/Contributions

4.1 Assumptions and Risks

The findings from the corpus-assisted discourse analysis demonstrate the entanglement between timbre and available technology and trends in MIR research at NIME. Although most of the 18 subcorpus papers do not directly define it, the framing and nature of timbre is tied to the method or technique used to examine it. Our working definitions of and interaction with timbre is therefore a result of the tools and approaches we have at our disposal. The availability of techniques to extract particular facets of timbre from audio signals (e.g., MFCCs) give cause to utilise and associate them with what timbre is for us and our musical practices. As a result, our understanding of timbre is tied to technoscientific achievements, a proper sociotechnical construction [51].

This use of particular views of timbre is not inherently problematic but can become so when these features assume the whole of timbre. As well, when the available technology changes, our conceptualisation changes. For example, the recent rapid proliferation of deep neural networks as audio synthesisers has remobilised our conceptualisation of and interaction with timbre to move away from feature extraction towards implicitly learned latent spaces that defy conceptual explanation. To that end, the scientific approach is intolerant of multiple truths about what timbre is and how it can be operationalised in DMI design. The scientific principle that, sooner or later, we will get to a concrete understanding of timbre through small advancements comes into contrast with artistic and creative ideals, through which there are many paths and no "correct" way through. This tension between technoscientific solutionism and phenomenological and constructivist approaches is present in broader HCI as well [56, 65]. Working to further advance control over timbre, as opposed to asking broader questions about context, creativity, and exploration, poses a risk of diluting timbre down to a handful

Table 4: Outline of the 18-paper NIME subcorpus with "timbre" featured in the title. Timbre objective(s) and goals are highlighted alongside tools and approaches used and their respective timbre associations from concordance analysis.

Corpus ID			Timbre Objective(s)		Tool(s) and Material(s)	Timbre Association(s)
Year	ID	Ref	Overall Goal(s)	Purpose		
2005	242	[31]	Control	Gestural timbre control, vowel-like formant synthesis	Speech-like vowel formants, weighted frequency and partials, tristimulus model: pitch, amplitude, timbre	frequency; gesture; shape
2006	376	[43]	Retrieval	Classification, transfer, control, expression for saxophone (bespoke interaction for saxophonist John Butcher)	Butcher's saxophone technique and performance, timbre categories from listener perspectives, frequency extraction model of auditory roughness [75], Lebanese mijwiz and other non-Western music, Pd: Puckett's fiddle~, Jehan's analyzer~	brightness; control; envelope; flutter; gesture; harmonicity; noise; parameters; roughness; shape; tone
2006	101	[45]	Retrieval	Classification, association of timbral parameters to human-perceptual descriptors	Chosen timbre descriptors (10 - bright, warm, harsh, hit, plucked, constant, thick, metallic, woody), frequency extraction techniques (extracted f0 and partials)	brightness; control; envelope; parameters; spectra
2007	270	[18]	Control	Timbral shaping and spectral manipulation	Wavetable/waveform timbre, physical shape of device to "shape" in the sound; frequency extraction	control; gesture; sound; spectra; structure
2009	276	[55]	Control	Connecting vocal syllables to percussive hits for percussion modeling	Vocable words specified in written form, Karplus-Strong physical percussion modeling and synthesis: tension and dampening of the surface, drumstick parameters for stiffness and mass; parameters: downward velocity, starting x/y position, angle and velocity of travel across the drum skin	control
2010	287	[32]	Exploration	Examining timbre as opposition to pitch, e.g., pitch vs sound in the physical materiality of instruments (glass, wood, stone)	Glass instruments: shape, size of vessels, glass tempering, structural, material qualities of different instruments; opposition to pitch and pitch serialisation to spectromorphologies (e.g., Schoenberg - Messiaen - Babbitt - Schaeffer, etc.)	space
2013	23	[47]	Control	Linking timbre and affect, understanding emotional aspects of timbre as structural component	MIRToolbox [49]—"features that relate to timbral and dynamic qualities", sounds and tags from Freesound.org folksonomy [30]	affect; genre; harmonicity; sound (musical vs non-musical); structure
2014	440	[95]	Control & Transfer	Spectral modelling synthesis, timbral morphing between source characteristics/timbral attributes and user-defined target timbral attributes	Audio morphing: brightness [13], softness [93], warmth [94], spectral modeling, FFT audio analysis for spectral and temporal characteristics: rhythmic density, amplitude, spectral flux, spectral centroid	control; noise; parameters; spectra; temporality; transfer; warmth
2016	81	[82]	Control & Visualisation	Aligning perceptual features with either spectral or harmonic timbral features	Spectral features (centroid, spread, and flatness) and harmonic features (harmonic energy ratio and inharmonicity) correlated in previous research using Timbre Toolbox [60].	articulation; control; gesture; harmonicity (harmonic energy ratio, inharmonicity); instrumentation; noisiness; parameters; sound; space; spectral centroid; texture
2017	30	[36]	Retrieval & Visualisation	Modelling timbre as a 3D space for embodied performance in virtual reality	Timbre space model [41], Brent's timbreID tools [11] extracting spectral brightness, spectral spread, spectral centroid, spectral flatness, and waveform slope, Bark-frequency applied to guitar technique classification, guitar pedagogy's relationship to timbre	affect; articulation; control; envelope; frequency (FFT-based); gesture; parameters; sound; space; spectra; structure
2017	12	[84]	Control	Clustering sounds based on timbre semantic descriptors	Use of "most common" timbral adjectives - warm and bright; audio clustering techniques; Blues and Metal training data; equaliser curves (frequency and volume dB) for for each cluster to map other samples into these representations	control; frequency (EQ curve); parameters; spectra
2018	78	[61]	Retrieval	Evaluating "tone-quality" on the cello; study of tone on the cello compared to other string instruments as a facet of pedagogy	Audio features that "correspond to faults in the fundamentals of bow control", FFT, Essentia library [10], various spectral and cepstral features initially, then focus on harmonic centroid [15, 42]	parameters; quality; sound; tone
2019	85	[39]	Control & Generation	Augmenting synthesis with timbre-based control parameters	parametrised latent space, timbre parameter extraction methods from MIR, inferred parameters from latent encodings, TimbreMap [38]	articulation; control; frequency; gesture; parameters; space
2021	50	[40]	Control	Investigating/evaluating control using six chosen and mapped timbre parameters	6 parameters controlled with physical knobs/sliders in physical synthesiser: frequency ratio, detune, duty cycle, modulation index, cutoff frequency, Q	control; expression; frequency; harmonicity; parameters; qualities; randomness; space; transposition
2021	38	[48]	Exploration & Visualisation	Using a GUI to visualise user-controlled interaction with timbre, perceptual salience, and user understanding of timbre	Sound spectrum: attack time, brightness, spectral flux, and spectral density recreated as frequency filter (low-or high-pass depending on parameter, ADSR envelope, physical keyboard controller, Euclidian space multi-dimensional scaling, cites previous model/interface [88]	articulation; brightness; differentiation; envelope; frequency; parameters; perception; sound; space; spectra; temporality; transposition; visual-auditory
2021	34	[4]	Control & Transfer	Generating novel sound with ambient sound tone, timbre and volume envelope mapped from source onto synthesised sound	Ambient sound timbre as tones, frequency-based timbre features, subtractive synthesis to a pitched tone, cascaded peak and notch filters, filter bandwidths as major "control" mechanism, ADSR volume envelope, MIDI/mod wheel controller	control; frequency; noise; shape; tone; transfer
2023	50	[81]	Exploration (Influence on Music Learning)	Timbre as a dimension of auditory feedback, as in music pedagogy/ learning associations between action and sound, for learning new DMIs	Physical percussion modelling, related to technical aspects re-configuration of a virtual-acoustic string-bridge-plate instrument model, mechano-acoustic relationships, gestures (i.e., hitting, scratching, pressing, tapping), "pink noise" object in MAX/MSP with audio follower, tracked sound intensity envelope of physical model, other spectral components blocked out	frequency
2024	55	[77]	Retrieval & Transfer	Real-time timbre remapping, using percussion instruments as example	Timbre remapping [85], differentiable digital signal processing (DDSP), sound pressure level, temporal centroid, spectral centroid [20], motivated by the MPEG-7 standard, spectral flatness	control; harmonicity; parameters; space; structure; transfer

of features and extraction methods. Without consideration, research at NIME and related fields can treat abstract concepts such as timbre (and pitch, rhythm, and even music itself) as problems to be resolved, as opposed to exploring their entanglement with and construction via sociocultural valuation.

4.1.1 Features as Proxies. Within the 18-paper subcorpus, it is clear that there is no one way that timbre is defined. Authors make associations to frequency, spectra, filters, modulations and other technical features of or operations on audio signals. Full corpus collocations indicate similar timbre associations. Selected features then become a way to work with timbre and are dependent on the technology being used and the relevant instrumentation and pedagogies in which the interaction is grounded. For example, Pond et al. [61] quantify tone quality in cello playing via the harmonic centroid of a note, a feature motivated by the authors in prior work by Charles [15] and Hermes et al. [42]. It is worthwhile to note that Charles identified features for violin timbre analysis, and the findings of Hermes et al. were based on listening rather than playing. The feature of harmonic centroid, while practically useful, is applied in such a way that it stands in for timbre, when in actuality it merely reflects a particular nuance of a larger conceptual space.

Pond et al. outline the origins of their chosen feature, which provides a necessary context for their system, as do Graham et al. [36] with Bark-frequency spectral features' relationship to guitar technique, and Hsu [43] with respect to communicative priorities in saxophone improvisation. We can observe how audio features such as these become a proxy for timbre as a whole. Without centering the chosen timbre proxies in previous research, particular tools and approaches (even down to the distinctions of violin and cello bowing technique) risk *becoming* timbre. Then, they may be used to objectively evaluate, say, performance as "good" or "bad" without consideration of the bias or constructed values through which these proxies originate. Such paradigms already exist in other HCI contexts, leading to data being taken out of context and potentially reducing human experience down to a handful of relatively arbitrary dimensions [17, 65].

4.1.2 Operationalising Features (or Not). If particular features have become a proxy for timbre, then it is important to reflect on how these features have come to be used. That NIME practices are inherently linked with trends in MIR research is logical: toolboxes for audio feature extraction from the late 2000s and early 2010s (e.g., [10, 49, 60]) make it simple to decompose timbre into distinguishable, readily controllable features. Today, deep learning and AI methods have greatly expanded the set of timbre generation and control tools that are possible, including some that would be unimaginable or infeasible with traditional audio signal processing techniques, such as replacing the timbre of one sound (e.g., a person singing) in an existing audio recording with that of another (e.g., a violin) while preserving pitch and dynamics, an approach known as "timbre transfer" [14, 77].

Feature extraction, timbre transfer and other machine listening techniques *augment* human listening and therefore our relationship to an instrument or timbre constructed with these methods. In this way, artefacts like aliasing [14] might become *real* aspects of a reified timbre because they are tied to an algorithm, rather than to human perception. In this sense, MIR research and DMI design both involve an analytical premise that timbre was always there and the current goal is to "extract" what it is and use it. When we define timbre in terms of these readily available technical representations, they become the point of interaction: timbre description is prescription, and DMIs and musical practice reconstitute themselves around what this operationalised definition of timbre is. This approach entangles the aesthetic valuation of commercial and solutionist approaches in DMI design, wherein accessible, easy-to-implement values are prioritised [59]. Conversely, there is less space for individual, subjective valuation of timbre and aesthetic consideration of the messy, emotional, and complex relationship we have with sound.

Our research practices fit into this valuation: having done so repeatedly through subsequent research, other experiences and previous knowledges are rewritten as a result of a particular approach—consider Pond et al.'s work (Section 4.1.1), where cello timbre is evaluated using a feature-as-proxy derived from findings in different contexts. Our *timbre thinking* is redefined by our *design thinking*—the solutions we come up with [59]. As Magnusson [53] describes, this premise operates as a "migration" of musical instruments around sociotechnical conditions. The evolution of particular valuations of what matters in musical expression are transported from one context to another. In NIME, timbre and our working definitions of it become embedded in this migration process; our research carries forward a valuation of certain approaches to understanding and manipulating timbre and instructs others, including the musicians using the tools we design, that they should value these approaches as well.

4.2 Six Points of Tension and Future Work

NIME is suspended in this tension between technoscientific and constructivist approaches to timbre. We must decide what notion of timbre we constructed based on the tools we use and create. First, we must acknowledge that the NIME community's conceptualisation of timbre is a construction of the digital age [57] and an entanglement of human and computer. The tools we use and create are part of the reification of timbre. Second, to examine this construction is not to suggest that instrument designers should abandon current approaches to timbre, or that alternative technologies could achieve timbre neutrality [51]. Instead, we might embrace and even emphasise the unresolved contradictions embedded in our tools, acknowledging their influence on the design of new timbral and musical artefacts.

In our analysis, we have identified six such contradictions as points of epistemological and methodological tension—*paradoxes of timbre* as famously outlined by Fales [27] and van Elferen [87]:

(1) **Static** vs. **dynamic** – Timbre can refer to fixed, objective features of sound, or timbre can be nuanced as an emergent property of human-computer interaction. Contemplating the timbral agency of piano pedaling in African-American jazz aesthetics (see Section 3.3), Gooley [35, p. 121] articulates that instrumental timbres "are the products of an encounter between a person and a technology" (see also [72]).

(2) **Categorical** vs. **continuous** – Categorical timbre arises from traditional music research (violin vs. trumpet) and language ("bright," "breathy," "airy") as well as the digital synthesis era (MIDI protocol, synthesiser presets). Acoustics and our musical experience recognise infinite gradations between these categories. Digital tools simultaneously reinforce both perspectives—offering categorical presets while also providing continuous control.

(3) **Controlled** vs. **explored** – While DMI design places central importance on interacting with timbre as something that can be controlled, manipulated, mapped, and transferred, it also "works directly contrary to the efforts of perceptualisation" [27, p. 66]

necessary to support timbre as something to be explored, for example, through utilising ambiguity over labelled features [65].

(4) **Identity** vs. **features** – Perceptually, we can understand the timbral difference between a violin tone and a cello tone based on context and experience (cf. [35, 87]). Timbre might also be portioned into explicit features as representations, which may act as a proxy for timbre altogether, for example, a "cello tone" is ascribed to particular harmonic partials [61]).

(5) **Ascribed** vs. **generated** – Timbre can be treated as something ascribed through listening and context. Feature extraction and timbre transfer (and generative AI more broadly) approaches ascribe timbre qualities to an existing signal. Timbre can also be viewed as something generated through measurable physical properties, as in physical modelling synthesis approaches [81]. When we construct representations and analytical spaces which also generate timbre (e.g., Wessel's numerical timbre control space [91]), we force the ascribing and generating perspectives to be the same thing when they are not.

(6) **Perceived by humans** vs. **by machines** – Humans use abstract representations and metaphorical language to conceptualise timbre, for instance calling a sound "bright," "breathy," "airy." Such descriptors are neither uniform nor easily defined; rather, digital systems mean that designers constantly negotiate between what is useful for machine processes and what is perceptible or meaningful to humans [5, 37, 76, 92].

This is a non-exhaustive and of course non-orthogonal list of contradictions-paradoxes in the way timbre is used and understood in DMI design. The NIME community can further explore how particular aspects of each tension might be encoded in our tools and practices and what they instruct users—musicians, audiences, scholars, scientists, engineers—to value. We suggest NIME to utilise these paradoxes and explore timbre through personal, situated, and entangled knowledge [2]. Within the NIME community's focus not on defining timbre but on control, manipulation, mapping, and transfer, there is potential to be open to other aspects and novel forms of timbre and timbre interaction. We argue that, if unpacked (but not necessarily resolved), these contradictions could be rather generative for both instrument design and timbre research.

Reflecting on our view of timbre as a site of epistemological and methodological tension between technoscientific and constructivist approaches, we ask DMI designers to consider: *If timbre is a "thing" that exists, then what form does it exist in? A numerical dimension space? Is it categorical or continuous? Is it something ascribed or generated?* and so on. Given the six points of tension outlined above, it is clear that this question is difficult to answer in any generalisable way, and we suggest that maybe timbre is *not* a "thing" to begin with, but a term that has been used itself as a proxy to the massive, ambiguous contexts that comprise sound and our experience of it. Thinking back to NIME's conceptualisation of timbre as a construction of the digital era [57], we may provoke a little further: *What is left when we remove the digital? Does "timbre" survive in these forms?*

An invitation to challenge and expand current notions of timbre might extend beyond the two main tendencies of technoscience and constructivism. In resonance with the metaphor of "timbre space" we draw an analogy from geography: the distinction between space (a property of the natural world, a location describable by coordinates) and place (which carries meaning, personality, or connections to cultural or personal identity). This could lead us to begin exploring alternative ways of conceptualizing and using timbre—such as in relation to personal emotional states, memories, shared identities, or cultural contexts [34]. Further methodologies or frameworks for interacting with social entanglements may then emerge.

We also aim to revisit the NIME corpus, including a more systematic analysis of ICMC and other pre- and off-NIME sources, to extract more nuanced information on digital music communities and timbre-based practice, and expanding it on other relevant literature (e.g., Computer Music Journal, Organised Sound, International Society for Music Information Retrieval) and also on non-academic sources (e.g., online communities about modular synthesisers, digital electronics). Doing so might elicit a better understanding of how DMI communities construct micro-conceptualisations of timbre based on community interest, communication, and collaboration [9].

Furthermore, as one anonymous reviewer of this work suggested, analysing how timbre discourse has shifted over time within NIME and ICMC (and related communities and sources) would provide valuable insights into the changing epistemologies of digital music research [52]. Such as the shifting interest from controlling features *extracted* from audio signals to manipulating latent spaces of neural networks *learned* from audio signals mentioned earlier, other trends and influences will have shaped the nuanced conceptualisation of timbre. Understanding these shifts over time can provide insight or even predictions about future shifts that will inevitably emerge with innovation in the digital and AI space.

Ethical Standards

This work did not involve experiments with other human participants, hence no institutional ethics board review was required. The authors have no known conflicts of interest.

Acknowledgments

We are grateful for the thoughtful feedback from the anonymous NIME reviewers, which has greatly improved this paper. This research is supported by the UKRI Centre for Doctoral Training in Artificial Intelligence and Music (EP/S022694/1), a UKRI Frontier Research (Consolidator) Grant (EP/X023478/1, "RUDIMENTS"), by the Royal Academy of Engineering under the Research Chairs and Senior Research Fellowships scheme, and by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101150317 ("DECODMI").

References

[1] Clyde Ancarno. 2020. Corpus-Assisted Discourse Studies. In *The Cambridge Handbook of Discourse Studies*, Anna De Fina and Alexandra Georgakopoulou (Eds.). Cambridge University Press, 165–185. https://doi.org/10.1017/9781108348195.009

[2] Kristina Andersen and Ron Wakkary. 2019. The Magic Machine Workshops: Making Personal Design Knowledge. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow, UK, 1–13. https://doi.org/10.1145/3290605.3300342

[3] ANSI. 1960/1994. *Psychoacoustic Terminology: Timbre.* New York, NY: American National Standards Institute.

[4] Lior Arbel. 2021. Aeolis: A Virtual Instrument Producing Pitched Tones With Soundscape Timbres. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Shanghai, China. https://doi.org/10.21428/92fbeb44.64f66047

[5] Jean-Julien Aucouturier and Emmanuel Bigand. 2012. Mel Cepstrum & Ann Ova: The Difficult Dialog Between MIR and Music Cognition. In *Proceedings of the International Society for Music Information Retrieval (ISMIR) Conference*. Porto, Portugal, 397–402. https://doi.org/10.5281/zenodo.1417179

[6] Paul Baker. 2004. Querying Keywords: Questions of Difference, Frequency, and Sense in Keywords Analysis. *Journal of English Linguistics* 32, 4 (2004), 346–359. https://doi.org/10.1177/0075424204269894

[7] Karen Barad. 2003. Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter. *Signs: Journal of Women in Culture and Society* 28, 3 (2003), 801–831. https://doi.org/10.1086/345321

[8] Jean-Baptiste Barrière. 1991. Introduction. In *Le timbre, métaphore pour la composition*, Jean-Baptiste Barrière (Ed.). Paris: IRCAM/Christian Bourgois, 11–13.

[9] S. M. Astrid Bin. 2021. Discourse is critical: Towards a collaborative NIME history. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Shanghai, China. https://doi.org/10.21428/92fbeb44.ac5d43e1

[10] Dmitry Bogdanov, Nicolas Wack, Emilia Gómez, Sankalp Gulati, Herrera Boyer, Oscar Mayor, Gerard Roma, Justin Salamon, José Zapata, and Xavier Serra. 2013. Essentia: An Audio Analysis Library for Music Information Retrieval. In *Proceedings of the International Society for Music Information Retrieval (ISMIR) Conference*. Curitiva, Brazil, 493–498. https://doi.org/10.5281/zenodo.1415016

[11] William Brent. 2010. A Timbre Analysis and Classification Toolkit for Pure Data. In *Proceedings of the International Computer Music Conference (ICMC)*. New York, NY, 224–229.

[12] V. Brezina and W. Platt. 2024. #LancsBox X 5.0.3 [software package].

[13] Tim Brookes and Duncan Williams. 2007. Perceptually-motivated audio morphing: Brightness. In *Audio Engineering Society Convention 122*.

[14] Franco Caspe, Jordie Shier, Mark Sandler, Charalampos Saitis, and Andrew McPherson. 2025. Designing Neural Synthesizers for Low Latency Interaction. *Journal of the Audio Engineering Society* (2025). https://doi.org/10.48550/arXiv.2503.11562

[15] Jane Charles. 2010. *Playing Technique and Violin Timbre: Detecting Bad Playing*. Ph. D. Dissertation. Faculty of Engineering, Dublin Institute of Technology (Technological University Dublin). https://doi.org/10.21427/D7HC8P

[16] Michel Chion. 2011. Dissolution of the Notion of Timbre. *differences* 22, 2–3 (2011), 235–239. https://doi.org/10.1215/10407391-1428906

[17] Caroline Claisse. 2024. Designing for Spiritual Informatics: Exploring a Design Space to Support People's Spiritual Journey. In *Companion Publication of the 2024 ACM Designing Interactive Systems Conference*. Copenhagen, Denmark, 140–143. https://doi.org/10.1145/3656156.3663723

[18] Luke Dahl, Nathan Whetsell, and John Van Stoecker. 2007. The WaveSaw : A Flexible Instrument for Direct Timbral Manipulation. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. New York City, NY, United States, 270–272. https://doi.org/10.5281/zenodo.1177079

[19] Anne Danielsen. 2006. *Presence and Pleasure: The Funk Grooves of James Brown and Parliament*. Wesleyan University Press, Middletown, CT.

[20] Anne Danielsen, Carl Haakon Waadeland, Henrik G Sundt, and Maria AG Witek. 2015. Effects of Tnstructed Timing and Tempo on Snare Drum Sound in Drum Kit Performance. *The Journal of the Acoustical Society of America* 138, 4 (2015), 2301–2316. https://doi.org/10.1121/1.4930950

[21] Joanna Demers. 2010. *Listening Through the Noise: The Aesthetics of Experimental Electronic Music*. Oxford University Press.

[22] Agostino Di Scipio. 1994. Formal Processes of Timbre Composition Challenging the Dualistic Paradigm of Computer Music. In *Proceedings of the International Computer Music Conference (ICMC)*. 202–208.

[23] Ryan Diduck. 2018. *Mad Skills: MIDI and Music Technology in the Twentieth Century*. Watkins Media Limited.

[24] Emily I. Dolan. 2012. Toward a Musicology of Interfaces. *Keyboard Perspectives* 5 (2012), 1–12.

[25] Emily I. Dolan and Alexander Rehding (Eds.). 2021. *The Oxford Handbook of Timbre*. Oxford University Press.

[26] Philippe Esling, Axel Chemla-Romeu-Santos, and Adrien Bitton. 2018. Bridging Audio Analysis, Perception and Synthesis with Perceptually-regularized Variational Timbre Spaces.. In *Proceedings of the International Society for Music Information Retrieval (ISMIR)*. 175–181. https://doi.org/10.5281/zenodo.1492373

[27] Cornelia Fales. 2002. The Paradox of Timbre. *Ethnomusicology* 46, 1 (2002), 56. https://doi.org/10.2307/852808

[28] Stefano Fasciani. 2020. Interactive Computation of Timbre Spaces for Sound Synthesis Control. *array. the journal of the ICMA* (2020), 69–78. https://doi.org/10.25370/array.v20152528

[29] Robert Fink, Melinda Latour, and Zachary Wallmark (Eds.). 2018. *The Relentless Pursuit of Tone: Timbre in Popular Music*. Oxford University Press.

[30] Frederic Font and Xavier Serra. 2012. Analysis of the Folksonomy of Freesound. In *Proceedings of the 2nd Computer Music (CompMusic) Workshop*. 48–54.

[31] Jesse Fox and Jennifer Carlile. 2005. SoniMime: Movement Sonification for Real-Time Timbre Shaping. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Vancouver, BC, Canada, 242–243. https://doi.org/10.5281/zenodo.1176741

[32] Ivar Frounberg, Kjell Tore Innervik, and Alexander R. Jensenius. 2010. Glass Instruments – From Pitch to Timbre. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Sydney, Australia, 287–290. https://doi.org/10.5281/zenodo.1177773

[33] Dana Gablasova, Vaclav Brezina, and Tony McEnery. 2017. Collocations in Corpus-Based Language Learning Research: Identifying, Comparing, and Interpreting the Evidence. *Language Learning* 67, S1 (2017), 155–179. https://doi.org/10.1111/lang.12225

[34] Can Gölgelioğlu and Anlı Ataöv. 2025. Timbre of the place: A Deleuzoguattarian inquiry to assemble music and place. *Planning Theory* (2025), 1–26. https://doi.org/10.1177/14730952251331240

[35] Dana Gooley. 2013. Jazz Piano Pedaling and the Production of Timbral Difference. *Keyboard Perspectives* 6 (2013), 101–126.

[36] Richard Graham, Brian Bridges, Christopher Manzione, and William Brent. 2017. Exploring Pitch and Timbre through 3D Spaces: Embodied Models in Virtual Reality as a Basis for Performance Systems Design. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Copenhagen, Denmark, 157–162. https://doi.org/10.5281/zenodo.1176207

[37] Owen Green, Bob Sturm, Georgina Born, and Melanie Wald-Fuhrmann. 2024. A Critical Survey of Research in Music Genre Recognition. In *Proceedings of the International Society for Music Information Retrieval (ISMIR) Conference*. San Francisco, CA, 745–782.

[38] Jeff Gregorio. 2019. *TimbreMap*. https://github.com/JeffGregorio/TimbreMap

[39] Jeff Gregorio and Youngmoo Kim. 2019. Augmenting Parametric Synthesis with Learned Timbral Controllers. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Porto Alegre, Brazil, 431–436. https://doi.org/10.5281/zenodo.3673025

[40] Jeff Gregorio and Youngmoo E. Kim. 2021. Evaluation of Timbre-Based Control of a Parametric Synthesizer. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Shanghai, China. https://doi.org/10.21428/92fbeb44.31419bf9

[41] John M. Grey. 1977. Multidimensional Perceptual Scaling of Musical Timbres. *The Journal of the Acoustical Society of America* 61, 5 (1977), 1270–1277. https://doi.org/10.1121/1.381428

[42] Kirsten Hermes, Tim Brookes, and Chris Hummersone. 2016. The harmonic centroid as a predictor of string instrument timbral clarity. In *Audio Engineering Society Convention 140*.

[43] William Hsu. 2006. Managing Gesture and Timbre for Analysis and Instrument Control in an Interactive Environment. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Paris, France, 376–379. https://doi.org/10.5281/zenodo.1176927

[44] Alexander Refsum Jensenius. 2014. To gesture or not? An analysis of terminology in NIME proceedings 2001–2013. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 217–220.

[45] Colin G. Johnson and Alex Gounaropoulos. 2006. Timbre Interfaces using Adjectives and Adverbs. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Paris, France, 101–102. https://doi.org/10.5281/zenodo.1176933

[46] Adam Kilgarriff, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel. 2014. The Sketch Engine: ten years on. *Lexicography* 1, 1 (2014), 7–36.

[47] Niklas Klügel and Georg Groh. 2013. Towards Mapping Timbre to Emotional Affect. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Graduate School of Culture Technology, KAIST, Daejeon, Republic of Korea, 525–530. https://doi.org/10.5281/zenodo.1178586

[48] Joshua Ryan Lam and Charalampos Saitis. 2021. The Timbre Explorer: A Synthesizer Interface for Educational Purposes and Perceptual Studies. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Shanghai, China. https://doi.org/10.21428/92fbeb44.92a95683

[49] Olivier Lartillot, Petri Toiviainen, and Tuomas Eerola. 2008. *A Matlab Toolbox for Music Information Retrieval*. Springer Berlin Heidelberg, 261–268. https://doi.org/10.1007/978-3-540-78246-9_31

[50] Amanda Lazar, Ben Jelen, Alisha Pradhan, and Katie A Siek. 2021. Adopting diffractive reading to advance hci research: A case study on technology for aging. *ACM Transactions on Computer-Human Interaction (TOCHI)* 28, 5 (2021), 1–29. https://doi.org/10.1145/3462326

[51] Giacomo Lepri and Andrew McPherson. 2021. Embrace the weirdness: negotiating values inscribed into music technology. *Computer Music Journal* 45, 3 (2021), 39–57. https://doi.org/10.1162/comj_a_00610

[52] Thor Magnusson. 2019. *Sonic Writing Technologies of Material, Symbolic, and Signal Inscriptions*. Bloomsbury Academic, New York, NY.

[53] Thor Magnusson. 2021. The migration of musical instruments: On the socio-technological conditions of musical evolution. *Journal of New Music Research* 50, 2 (2021), 175–183. https://doi.org/10.1080/09298215.2021.1907420

[54] Stephen McAdams. 2019. The Perceptual Representation of Timbre. In *Timbre: Acoustics, Perception, and Cognition*. Springer International Publishing, 23–57.

[55] Alex Mclean and Geraint Wiggins. 2009. Words , Movement and Timbre. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Pittsburgh, PA, 276–279. https://doi.org/10.5281/zenodo.1177629

[56] Andrew McPherson, Landon Morrison, Matthew Davison, and Marcelo M. Wanderley. 2025. On Mapping as a Technoscientific Practice in Digital Musical Instruments. *Journal of New Music Research* (2025), 217–220.

[57] Landon Morrison. 2024. Timbre space: On the flat history of a multidimensional metaphor. *Music & Science* 7 (2024), 1–16.

[58] Landon Morrison and Andrew McPherson. 2024. Entangling Entanglement: A Diffractive Dialogue on HCI and Musical Interactions. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Honolulu, HI, 1–17. https://doi.org/10.1145/3613904.3642171

[59] Laurel Pardue and S. M. Astrid Bin. 2022. The Other Hegemony: Effects of software development culture on music software, and what we can do about it. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Auckland, New Zealand. https://doi.org/10.21428/92fbeb44.0cc78aeb

[60] Geoffroy Peeters, Bruno L. Giordano, Patrick Susini, Nicolas Misdariis, and Stephen McAdams. 2011. The Timbre Toolbox: Extracting Audio Descriptors from Musical Signals. *The Journal of the Acoustical Society of America* 130, 5

(2011), 2902–2916. https://doi.org/10.1121/1.3642604

[61] Robert Pond, Alexander Klassen, and Kirk McNally. 2018. Timbre Tuning: Variation in Cello Sprectrum Across Pitches and Instruments. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Blacksburg, VA, 356–359. https://doi.org/10.5281/zenodo.1302619

[62] Nicola Privato. 2024. Mouja: Experiencing AI through Magnetic Interactions. In *Proceedings of the Eighteenth International Conference on Tangible, Embedded, and Embodied Interaction*. Cork, Ireland, 1–3. https://doi.org/10.1145/3623509.3635328

[63] Miller Puckette. 2002. Max at seventeen. *Computer Music Journal* 26, 4 (2002), 31–43.

[64] Paul Rayson, Damon Berridge, and Brian Francis. 2004. Extending the Cochran rule for the comparison of word frequencies between corpora. In *Le poids des mots: Proceedings of the 7th International Conference on Statistical Analysis of Textual Data (Journées internationales d'Analyse statistique des Données Textuelles)*. 926–936. https://api.semanticscholar.org/CorpusID:15579590

[65] Courtney N. Reed, Adan L. Benito, Franco Caspe, and Andrew P. McPherson. 2024. Shifting Ambiguity, Collapsing Indeterminacy: Designing with Data as Baradian Apparatus. *ACM Transactions on Computer-Human Interaction (TOCHI)* 31, 6 (2024), 1–41. https://doi.org/10.1145/3689043

[66] Nathan Renney, Benedict Gaster, Tom Mitchell, and Harri Renney. 2022. Studying How Digital Luthiers Choose Their Tools. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. New Orleans, LA, 1–18. https://doi.org/10.1145/3491102.3517656

[67] Matthew Rodger, Paul Stapleton, Maarten Van Walstijn, Miguel Ortiz, and Laurel Pardue. 2020. What makes a good musical instrument? a matter of processes, ecologies and specificities. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Birmingham, UK, 405–410. https://doi.org/10.5281/zenodo.4813438

[68] Tara Rodgers. 2010. *Synthesizing sound: Metaphor in audio-technical discourse and synthesis history*. Ph. D. Dissertation. Department of Art History and Communication Studies, McGill University.

[69] Jean-Jacques Rousseau. 1768. *Dictionnaire de musique*. Paris: Duchesne la Venue.

[70] Luigi Russolo. 1913. *L'arte dei rumori*. Direzione del movimento futurista.

[71] Charalampos Saitis, Bleiz M Del Sette, Jordan Shier, Haokun Tian, Shuoyang Zheng, Sophie Skach, Courtney N Reed, and Corey Ford. 2024. Timbre Tools: Ethnographic Perspectives on Timbre and Sonic Cultures in Hackathon Designs. In *Proceedings of the 19th International Audio Mostly Conference: Explorations in Sonic Cultures*. Milan, Italy, 229–244. https://doi.org/10.1145/3678299.3678322

[72] Charalampos Saitis, Claudia Fritz, Gary P. Scavone, Catherine Guastavino, and Danièle Dubois. 2017. Perceptual evaluation of violins: A psycholinguistic analysis of preference verbal descriptions by experienced musicians. *Journal of the Acoustic Society of America* 141, 4 (2017), 2746–2757. https://doi.org/10.1121/1.4980143

[73] Charalampos Saitis and Stefan Weinzierl. 2019. The Semantics of Timbre. In *Timbre: Acoustics, Perception, and Cognition*, Kai Siedenburg, Charalampos Saitis, Stephen McAdams, Arthur N. Popper, and Richard R. Fay (Eds.). Springer, Cham, 119–149. https://doi.org/10.1007/978-3-030-14832-4_5

[74] Allan Seago, Simon Holland, and Paul Mulholland. 2008. Timbre space as synthesis space: towards a navigation based approach to timbre specification. In *Spring Conference of the Institute of Acoustics: Widening Horizons in Acoustics*. 516–523.

[75] William A. Sethares. 2005. *Tuning, Timbre, Spectrum, Scale* (2 ed.). Springer, London. https://doi.org/10.1007/b138848

[76] Jordie Shier, Rodrigo Constanzo, Charalampos Saitis, Andrew Robertson, and Andrew McPherson. 2025. Designing Percussive Timbre Remappings: Negotiating Audio Representations and Evolving Parameter Spaces. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Canberra, Australia.

[77] Jordie Shier, Charalampos Saitis, Andrew Robertson, and Andrew McPherson. 2024. Real-time Timbre Remapping with Differentiable DSP. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Utrecht, Netherlands, 377–385. https://doi.org/10.5281/zenodo.13904884

[78] Kai Siedenburg, Charalampos Saitis, and Stephen McAdams. 2019. *The Present, Past, and Future of Timbre Research*. Springer International Publishing, 1–19. https://doi.org/10.1007/978-3-030-14832-4_1

[79] Kai Siedenburg, Charalampos Saitis, Stephen McAdams, Arthur N. Popper, and Richard R. Fay (Eds.). 2019. *Timbre: Acoustics, Perception, and Cognition*. Springer, Cham. https://doi.org/10.1007/978-3-030-14832-4

[80] Denis Smalley. 1994. Defining timbre – refining timbre. *Contemporary Music Review* 10, 2 (1994), 35–48. https://doi.org/10.1080/07494469400640281

[81] Olivia B Smith, Matthew Rodger, Maarten van Walstijn, and Miguel Ortiz. 2023. Sound guiding action: the effect of timbre on learning a new percussive DMI for beginner musicians. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Mexico City, Mexico, 358–363. https://doi.org/10.5281/zenodo.11189208

[82] Sean Soraghan, Alain Renaud, and Ben Supper. 2016. Towards a perceptual framework for interface design in digital environments for timbre manipulation. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Brisbane, Australia, 413–418. https://doi.org/10.5281/zenodo.1176129

[83] Katta Spiel. 2021. The bodies of tei–investigating norms and assumptions in the design of embodied interaction. In *Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction*. 1–19. https://doi.org/10.1145/3430524.3440651

[84] Spyridon Stasis, Jason Hockman, and Ryan Stables. 2017. Navigating Descriptive Sub-Representations of Musical Timbre. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Copenhagen, Denmark, 56–61. https://doi.org/10.5281/zenodo.1176171

[85] Dan Stowell and Mark D Plumbley. 2010. Timbre remapping through a regression-tree technique. In *Proceedings of the 7th Sound and Music Computing Conference*. Barcelona, Spain, 45–50.

[86] Dylan Van der Schyff. 2015. Music as a manifestation of life: exploring enactivism and the 'eastern perspective'for music education. *Frontiers in Psychology* 6 (2015), 345. https://doi.org/10.3389/fpsyg.2015.00345

[87] Isabella van Elferen. 2018. Timbrality: The Vibrant Aesthetics of Tone Color. In *The Oxford Handbook of Timbre*, Emily I. Dolan and Alexander Rehding (Eds.). Oxford University Press, 68—-91. https://doi.org/10.1093/oxfordhb/9780190637224.013.28

[88] Roel Vertegaal and Barry Eaglestone. 1998. Looking for Sound? Selling Perceptual Space in Hierarchically Nested Boxes. In *CHI 98 Conference Summary on Human Factors in Computing Systems*. 295–296.

[89] Marcelo M Wanderley. 2023. Prehistoric NIME: Revisiting research on new musical interfaces in the computer music community before NIME. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 60–69. https://doi.org/10.5281/zenodo.11189104

[90] W. Dixon Ward. 1965. *Psychoacoustics*. Williams Wilkins Co., Baltimore, 48–71.

[91] David L. Wessel. 1979. Timbre Space as a Musical Control Structure. *Computer Music Journal* 3, 2 (1979), 45–52. https://doi.org/10.2307/3680283

[92] Geraint A. Wiggins. 2009. Semantic Gap?? Schemantic Schmap!! Methodological Considerations in the Scientific Study of Music. In *2009 11th IEEE International Symposium on Multimedia*. 477–482. https://doi.org/10.1109/ISM.2009.36

[93] Duncan Williams and Tim Brookes. 2009. Perceptually-motivated audio morphing: softness. In *Audio Engineering Society Convention 126*.

[94] Duncan Williams and Tim Brookes. 2010. Perceptually-motivated audio morphing: warmth. In *Audio Engineering Society Convention 128*.

[95] Duncan Williams, Peter Randall-Page, and Eduardo Miranda. 2014. Timbre morphing: near real-time hybrid synthesis in a musical installation. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. London, United Kingdom, 435–438. https://doi.org/10.5281/zenodo.1178983

[96] Asterios Zacharakis, Konstantinos Pastiadis, and Joshua D. Reiss. 2014. An Interlanguage Study of Musical Timbre Semantic Dimensions and Their Acoustic Correlates. *Music Perception* 31 (2014), 339–358. https://doi.org/10.1525/MP.2014.31.4.339

[97] Eevee Zayas-Garin, Charlotte Nordmoen, and Andrew McPherson. 2023. Transmitting Digital Lutherie Knowledge: The Rashomon Effect for DMI Designers. In *Proceedings of International Conference on New Interfaces for Musical Expression*. Mexico City, Mexico, 350––357. https://doi.org/10.5281/zenodo.11189206