

PSYCHOACOUSTIC, COGNITIVE, and PERCEPTUAL TERMINOLOGY AND OTHER CORE-RELATED ISSUES

This document represents the interactions between the nine members of a UCSD Music Department CORE 1 seminar during the Fall of 2020 and its instructor, R. Reynolds. The seminar members were formed into three groups that interacted continuously and collegially throughout. Group 1 considered terminology with Perceptual implications, Group 2 terms with Cognitive relevance, and Group 3 considered Evaluative terminology.

UCSD team: Stephen de Filippo, Erin Graham, Peter Ko, Zachary Konick, Ioannis Mitsialis, Sang Song, Jonny Stallings, Alex Stephenson, Alexander Taylor

McGill edits by: Stephen McAdams, Lindsey Reymore, Joshua Rosner

PART I. PERCEPTUAL TERMINOLOGY (concerning the local phenomena that underlie the character of musical sound)

Introduction

We propose new definitions for foundational *perceptual* terminology. Of particular importance is the notion of *timbral blending*, which we understand as a continuum and which underlies other ancillary terms, such as *timbral augmentation*, *resonance*, and *emergence*. These terms address special cases of timbral blending frequently utilized as orchestrational devices. Further, we propose new definitions of *Klangfarbenmelodie* and *masking*.

DEFINITIONS:

1. Timbral blending

Definition: Given two or more sound sources, *timbral blending* occurs when their timbral identities appear less demarcated, blurred or even erased as a result of such multiple sources sounding simultaneously. Timbral blending is a continuum whose two extremes are “unblended” or “distinct” at one end and “fully blended” or “fused” at the other. If the simultaneous timbral identities maintain their character, they are “unblended” or “distinct.” On the other hand, two or more timbres are “partially blended” if their timbral identities appear less demarcated or blurred, and “fully blended” or “fused” if their timbral identities appear to be erased completely.

Commentary: *Timbral blending is the fusion of different sources of acoustic information into a seemingly unified auditory event. Timbral blending primarily depends on onset synchrony, harmonicity, and the degree of similarity of constituent sound spectra. It is reinforced over time by parallel or similar motion in pitch and dynamics. Timbral blending is the fundamental concept underlying such ancillary concepts as timbral augmentation, timbral extension, and timbral emergence. We will now turn to definitions of these ancillary concepts, each of which relates to a special instance of timbral blending which has numerous musical applications.*

2. Timbral augmentation

Definition: *Timbral augmentation* occurs when two or more timbres are partially blended and when a hierarchical relationship is drawn between the timbres, such that secondary timbre(s) are understood to embellish, highlight or reinforce primary timbre(s).

Commentary: *The underlying concept for timbral augmentation is timbral blending. Timbral augmentation can occur when the constituent timbres are partially blended without some of them completely losing their individual timbral identities.*

3. Timbral resonance

Definition: *Timbral resonance* is a case of timbral augmentation in which, after the primary timbre(s) are no longer perceptible, the secondary timbre(s) remain perceptible.

Commentary: *Timbral resonance is a device where the underlying concept is timbral augmentation (which, in turn, is a subset of timbral blending). Timbral resonance is often implemented in such a way that the behavior of the secondary timbre(s) mimics the resonance created by a piano's sustain pedal, though in some cases the secondary timbre(s) may also sustain or swell.*

4. Timbral emergence

Definition: *Timbral emergence* can occur when there is a fully blended (fused) timbral combination in which the individual constituent timbres are significantly distinct from one another, and whose fusion results in a strikingly novel timbral complex which transcends the constituent timbral identities.

Commentary: *For timbral emergence, the underlying condition is timbral fusion (which, in turn, is a condition within timbral blending). The core idea of timbral emergence is that two or more significantly distinct timbres (e.g., belonging to different instrumental families) produce a strikingly novel timbral complex when sounding together.*

5. Klangfarbenmelodie

Definition: *Klangfarbenmelodie* occurs when timbre changes in a manner akin to constructing a melody by successive pitch changes. If timbre changes occur in a context that lacks prominent melodic implications, it is of a Schoenbergian type. If timbre changes occur in a context with pre-existing prominent melodic elements, it is of a Webernian type.

Commentary: *For Schoenbergian Klangfarbenmelodie to exist, it may be necessary to suppress or background non-timbral elements so as to bring the timbre changes to the foreground. In the case of Webernian Klangfarbenmelodie, timbre changes enhance or reinforce the pre-existing melodic elements rather than creating melodic elements where there were none previously.*

6. Masking

Definition: *Masking* occurs when one sound inhibits, whether partially or fully, the clear perception of another. This can take the form of obscuring, distracting or altering perceptions and can result from loudness differences, textural contrasts, or temporal proximity.

Commentary: Textural contrasts can result in masking if, for example, a sound with a high density of onsets occurs alongside another sound with a lower density of event onsets; clear perception of the second sound may be inhibited by the presence of the first. Temporal proximity can result in masking when an attention-grabbing sound occurs just before or shortly after a less distinctive sound and distracts the perceiver from registering the identity or even the existence of the more subtle event. Several masking-inducing factors may coincide with and reinforce one another.

MUSICAL EXAMPLES for PERCEPTUAL TERMINOLOGY

1. Timbral Blending (Timbre Lingo [Blend](#))

Partial Blending

Olivier Messiaen: *Quartet for the End of Time*, Movt. 6

<https://www.youtube.com/watch?v=csugaLRPdZM>

(Performed by Matthew Schellhorn & Soloists of the Philharmonia Orchestra)

In this excerpt, the individual identities of the clarinet, violin, cello, and piano are somewhat, but not completely, blurred. Factors such as onset synchrony and harmonicity (unison pitch doublings) contribute to this instance of partial blending.

2. Timbral Augmentation

Claude Debussy: *Trois Nocturnes*, Movt. 3 (23:08-23:30)

<https://youtu.be/spXwXLqFLvs?t=1388>

(Performed by Montreal Symphony Orchestra, Charles Dutoit, conductor)

In this example, the clarinets' undulating figures are occasionally doubled by the strings, resulting in the former's timbre being augmented by that of the latter. When these figures are passed on to the strings (esp. the violas), the roles are reversed and the clarinets' timbre augments the strings'.

3. Timbral Resonance

Tristan Murail: *Ethers* (13:46-14:30)

<https://youtu.be/4Y5GlzOPzX4?t=826>

(Performed by Erin Lesser, flute, and Argento Chamber Ensemble, Michel Galante, conductor)

In this example, a secondary timbre (strings) sustains short attacks articulated by a primary timbre (flute). Rather than simply modeling standard acoustical resonance, the strings also perform dynamic swells, creating a more complex, imaginary resonating chamber.

4. Timbral Emergence

Claude Debussy: *La Mer*, Movt. 1 (0:35-1:00)

https://youtu.be/loENgt1h4_A?t=33

(Performed by the Cleveland Orchestra, Pierre Boulez, conductor)

In this excerpt, the sounds of the muted trumpet and the English Horn are fused to create a novel, emergent timbre.

5. Klangfarbenmelodie (Timbre Lingo [Klangfarbenmelodie](#))

Schoenbergian

Arnold Schoenberg: *Five Pieces for Orchestra*, Op. 16, Movt. 3 "Farben" (0:00-0:50)

<https://www.youtube.com/watch?v=uUzab3O7vBA>

(Performed by the Cleveland Orchestra, Christoph von Dohnányi, conductor)

In this example, different groups of instruments with different timbres play the same chords in alternation.

Webernian

J.S. Bach-Anton Webern: *Ricercar a 6 voci* from *Musical Offering* (0:00-)

<https://www.youtube.com/watch?v=UUECC4IGVAw>

(Performed by Berlin Philharmonic Orchestra (Pierre Boulez (cond.))

In this Webern's orchestration of Bach's Ricercar, Bach's original line passes back and forth between different instruments, which reinforces and embellishes Bach's original melody.

6. Masking (Timbre Lingo [Masking](#))

Cassandra Miller: *Duet for Cello and Orchestra* (14:10-)

<https://drive.google.com/file/d/1qPxGTISwGpHx9iq0HIQAgFTLNeh2Zzrk/view?usp=sharing>

(Performed by Charles Curtis, cello, and BBC Scottish Symphony Orchestra, Ilan Volkov, conductor)

*In this example, the solo cello alternates a unison double-stopped G and D (marked **mf**) without a pause. However, one cannot hear the cello when the orchestra's loud outbursts cover it up.*

PART II. COGNITIVE TERMINOLOGY (Concerning larger-scale phenomena that contribute to the perceived structure of music)

DEFINITIONS:

7. Event: a sonic occurrence which can encompass any of the ensuing timbral phenomena

8. Surface texture: a collection of interwoven instrumental parts that are difficult to follow perceptually individually and which cohere into a complex grouping in which it is clear that there is more than one instrument playing, but one tends to focus on the grouping and its emergent textural properties.

9. Stratification: Formation of layers with a single sound source or groups of blended or associated sound sources, and that incorporates changes of its constituent elements linearly over time.

9a. Layer: a layer tends to behave in such a way that its components are perceived as belonging together and are consequently less identifiable. A layer's components can change erratically, and, unlike a stream, it can also be static. A stratum doesn't necessarily have a clear flow or linear direction, which distinguishes it from a stream, although a stream may occupy a layer, such as a solo instrument in a foreground layer or two contrapuntal streams of equivalent prominence may occupy a layer.

9b. Layering: stratification involves looser groupings of events into strata of different prominence or proximity (e.g., foreground, middle ground, background). Individual stratified layers can include several instruments that are not fused, but which group together by

contrast with materials being played in other layers. Such contrasts can, but don't always include timbral differences (for example, a homogeneous ensemble). A certain degree of hierarchical organization is inherent in the constituent elements of each stratum, constituents not restricted to the timbral domain.

Commentary: When diverse and partially differentiated elements coexist such as in the above described case of strata, focal elements are unavoidably perceived as more prominent than others because of their function in their environment. This gives them a kind of independence. In order for this prominence to be recognizable, elements other than timbre need to be taken into consideration such as register, rhythm, dynamics, and pitch. These focal elements tend to form a certain hierarchical organization in terms of perception, leading to the idea of the "dominance" of certain musical elements between musical layers. Still, the term "stratification" involves conceiving of an agreed upon "space" that can be defined in terms of different dimensions. If one thinks of a space in terms of proximity, then near (foreground), middle-ground, and far (background) distinctions are relevant.

10. Grouping / perceptual grouping: Perceptual effects of orchestration practice based on principles of auditory grouping, in three sub-categories:

10a: Concurrent grouping: Concerns the integration of the acoustic information present in each of multiple, ongoing, and concurrently occurring sound sources. This phenomenon is responsible for auditory event formation. In orchestration, an "event" can be formed from several grouped sound sources. There is not always a significant amount of change over time; change can occur so gradually that it's not necessarily easily perceived.

Commentary: This definition seemingly points to homophonic textures that lack significant change or development over time. However, the concept of homogeneity is complicated by sound states in which multiple events occur so frequently that we fail to recognize each particular sound event, begetting a cloud of granularity. Such "chaotic" or stochastic textures arguably make up a state of concurrent grouping that behaves similarly to homophony through its perceived state of stagnation and homogeneity.

10b: Sequential grouping: Involves the grouping of similar sources that occur sequentially over time. These are perceived as separate from other groupings also occurring sequentially over the same time span. Here, "sequential" implies that grouped events evolve over time. Such evolution, in turn, implies that streaming is a necessary factor in the perception of sequential grouping.

10c: Segmental grouping: Segmentation into perceptual units occurs when discrete changes occur to sequential groupings in one or more ongoing musical dimensions; in particular for example, sudden changes in instrumentation or electroacoustic sound create perceptual contrasts that delineate sets of different segments in succession, each with their individual identities.

11. Auditory scene analysis: This is an outlook rather than a phenomenon. The set of perceptual processes by which a sound world – understood as a "scene" such as is established by objects in a visual field – is organized into events, sequences of events (streams, surface textures), layers of differing degrees of perceptual prominence, and segmented units of various temporal scales.

Discussions of such fields often imply a separation into foreground, middle ground, and background events. (Timbre Lingo [Auditory Scene Analysis](#))

MUSICAL EXAMPLES for COGNITIVE TERMINOLOGY

8. Texture

EXAMPLE: Xenakis *Pithoprakta* 0'58- 1'46

<https://youtu.be/nvH2KYYJg-o?t=58>

Dense irregular entries make it difficult to follow individual instruments and result in a zooming out of attention to take in the whole as a texture.

9. Streaming / musical streams

EXAMPLE: Bartok *Sonata for 2 Pianos and Percussion* 4'44 - 5'23

<https://youtu.be/uRoaFfMLx-M?t=282>

Independence in terms of timbre, register, rhythm, and melodic context creates separate streams for two pianos and timpani.

9a & b. Stratum and stratification

EXAMPLE: Messiaen, *Quartet for the End of Time*, Part I - Crystal Liturgy - 0' - 0'27

https://www.youtube.com/watch?v=i6caajUI_p8

Clarinet foreground, other instruments in background. Example of a single stream in a foreground layer.

EXAMPLE: Ligeti *Piano Concerto*, mvt. 1: beginning - 0'53

<https://www.youtube.com/watch?v=P3mDevv2LP0>

Piano foreground, other instruments in background. Example of a complex texture in a foreground layer.

10. Grouping / perceptual grouping: Perceptual effects of orchestration practice based on principles of auditory grouping, in three sub-categories:

10a: Concurrent grouping

EXAMPLE: Ives's *The Unanswered Question* 0:00 -2:30

<https://www.youtube.com/watch?v=WBiLOVEttZw>

Blend of static background sonority.

10b: Sequential grouping

EXAMPLE 1: Beethoven, 7th Symphony, Allegretto 0:00-0:47 ... 1:16 segmental moment; 1:30 another sequential layer enters

<https://www.youtube.com/watch?v=B6Mi3fcVd4M>

Initial melody as a stream, joined at 0:47 by a countermelody and then again by another line at 1:30..

EXAMPLE 2: Beat Furrer's *Aria* (1999) beginning - 0'58 ; 5'46 - 6'25

<https://www.youtube.com/watch?v=xbbrNAPfLfM>

Failure to form streams due to strong discontinuities creating fragmentation at beginning. At 5'46 individual instrumental streams are more apparent.

10c: Segmental grouping.

EXAMPLE: Beethoven 7th Symphony, Allegretto 7:45 segmental sequence

<https://youtu.be/B6Mi3fcVd4M?t=465>

Segmentation of material by orchestral contrasts in recapitulation.

PART III. EVALUATIVE TERMINOLOGY (Concerning larger-scale phenomena [how local phenomena in music can be verbally characterized by perceivers] that contribute to the perceived structure of music)

Introduction:

TIMBRE TERMS (EVALUATIVE): Concepts that combine temporal spectral profiles with perceptual correlates, as captured with the most common semantic labels.

This list initially included: bright, dark, rough, smooth, and muted. We suggest below an expanded toolkit of evaluative terms to cover different sonic qualities that are glossed over or overly simplified by the original list. We also propose pairs of opposing terms encompassing continua, rather than discrete conditions. So, a sound could be perceived as “somewhat bright”, or “extremely limited”, for example. We hope that with further revisions to the evaluative timbre terms, that a language can be arrived at that serves musicians (and others) in identifying and describing sounds with more precision and nuance.

On the issue of evaluating the qualities of sounds, such as brightness/darkness, these terms can be quite subjective and rely substantially on context. A bright sound in one contextual framework (i.e., a piece for 4 contrabasses) might be perceived as a dark sound in another contextual framework (i.e., a piece for 4 piccolos). All of this lies on a spectrum where context of instrumental forces is key.

One theoretical application for these evaluative terms (which would have to be tested) is that they could potentially be utilized as predictors of the perceptual blending outcomes of particular instrumental combinations.

DEFINITIONS:

13. Brightness (continuum bright/dark): The degree to which a sound's spectrum favors high frequencies, strongly correlated with the position of the spectral centroid. Perceptually, it could include (at the bright end) brilliant, edgy, nasal, metallic; (at the dark end) dull, covered. (Timbre Lingo [Brightness/Darkness](#))

14. Roughness (continuum rough/smooth): The degree of sensory dissonance in a sound; related to harmonicity/inharmonicity. Perceptually, it could include (at the rough end) graininess, noisiness, hairiness, wobbliness, shimmer, etc.

15. Richness (continuum rich/limited): The fullness of the sound; the number of partials meaningfully present in the spectrum: greater number of partials (e.g. a string instrument played sul ponticello) produces a more complex, full sound; smaller number of partials (e.g. bowed vibraphone, string harmonics) produces a simple, thin, pure sound; limited timbres are often associated with muting techniques.

16. Focus (continuum focused/diffuse): The “peakiness” of the spectrum around a particular frequency/partial, or the spread of frequencies across the spectrum; the perceived directionality of the sound. A more focused sound will have a cluster of intensity around a single frequency, while a more diffuse sound will have a spectrum that is flatter and less peaky.

17. Tonal volume (continuum large/small): the perceived “size” of the sound. I.e., a sound can seem fatter or more voluminous as produced by one instrument or another. This quality is independent of amplitude or pitch.

18. Spectral Centroid - The “center of gravity” in terms of a sound’s frequency spectrum (a higher centroid implies a brighter sound, etc.).

MUSICAL EXAMPLES for EVALUATIVE TERMINOLOGY

General examples

Using the original terminology with which the seminar began, we could categorize a bowed vibraphone timbre as bright and smooth. But the new terminology allows us to also characterize its richness. The vibraphone timbre is quite spectrally limited, despite its overall brightness: its spectrum peaks around the fundamental and again at the fourth partial, but few other partials are substantially present. A bright sound might be expected to be rich, and brightness and richness will often be correlated, but here they are not.

Using the original terminology, we could categorize the first octave of the flute as dark, but the new terminology allows us to characterize its richness and focus. The flute in its low register is quite limited in partials, and also quite diffuse in comparison to the focus of its third octave.

Examples of Timbral Combinations

Alex Stephenson’s *Etude* - Time-code: 2:04 (meas. 81)

Unison sustain of Trombone [C4] and B. Cl. [E4]

Trombone - moderately bright, slightly rough/’noisy’/’dirty’, mildly limited, focused(?) timbre
Bass Clarinet - moderately bright, smooth, ’rich’ and ’focused’ timbre.

Commentary: Very little blend between these instruments, especially in this register, when spatialized and a major third apart, so it is difficult to sense a combined timbral unit that can be evaluated as a single unit. This is perhaps due to the dominance of the odd partials present in the clarinet spectrum when juxtaposed against the contrasting trombone spectrum. When analyzing the combined unitary character, my approach was to take aspects from each category of evaluation for the two instruments, and decide which is the dominant presence in the combined unit.

Attempted characterization of time-code 2:04: A moderately bright, moderately rough (once any rough sound is present, you can’t completely discount that dimension), fairly rich, focused timbre.

Score and recording in CORE 1 Materials on myCourses.

Sang Song's *Stendhal Etudes* – III. Time code: 1:15 mm.1-2, 9

Sustained chord of Ratchet, overpressured violin (G3), fluttertongued+harmon muted trombone (C#2), bass clarinet multiphonic (D2)

Commentary: 1:15 - Very rough at the high end, and very limited in the middle of the spectrum. High frequencies predominate, with a lot of noisy components. Overall seems quite bright, focused, and noisy with the clarinet multiphonics and fluttertongued trombone taking weight away from the fundamental. The ratchet masks some of the high resonance of the other instruments.

Score and recording in CORE 1 Materials on myCourses.

Anqi Liu's *Etude for Echoes* Time code: 0:00-0:20

Violin, Bass Clarinet, Trombone, and Percussion (Medium Slate) fluctuate around a “node”, oscillating between barely pitched and unpitched sounds, exploring the timbral shadows of each instrument.

Commentary: What we get as a general result is a limited and rough sound from each instrument, with the exception of a striking, glancing attack by the percussion around 0:18, which is quite noticeably rich in contrast.

Both the violin and trombone remain rather focused in their sound, while the bass clarinet and percussion have a diffuse quality. However, the violin and percussion both share a brighter sound quality, despite differences in focus. My hypothesis has to do with the higher spectral centroid of these sounds. The trombone can be characterized as a brighter sound despite its limited/muted characteristics, and the bass clarinet comes across as quite dark.

In terms of tonal volume, the order from largest to smallest is: bass clarinet, percussion, trombone, and violin. There is certainly a correlation between lower pitch and wider volume. However, I would note that while the percussion feels as though it has a wider tonal volume compared to the trombone and violin (a view that might be contested), it certainly is much softer than these other instruments, and is less noticeable due to its diffusion.

Score and recording in CORE 1 Materials on myCourses.