

Musical Meaning Within Super Semantics

Philippe Schlenker

(Institut Jean-Nicod, CNRS; New York University)

Music consultant:

Arthur Bonetto



European Research Council

Established by the European Commission

Special thanks to Paul Egré and Lyn Tieu

This research was funded by the ERC Advanced Grants FRONTSEM and Orisem.

Claims

- Music has a meaning (it can convey information about the extra-musical world), but musical meaning is mostly produced by different means than linguistic meaning.
- Musical meaning is usually abstract (it holds true of very diverse situations), but its effects can be assessed by the method of 'minimal pairs' (i.e. by rewriting excerpts).
- Music semantics benefits from new tools from Super Semantics, the study of meaning beyond spoken language. **Case study:** referential indices in language, pictures, music
- When accompanying other mediums, such as film, music might have the same semantic effects as gestures.

Semantics

- a. Syntax = what is well-formed
 - b. **Semantics** = what the form says about the world
- 'To know the meaning of a sentence (... or another meaning-bearing form) is to know in what situations it is true.'

■ **Compositional semantics**

F and G is true if and only if $F = G = \text{true}$

F or G is true if and only if $F = \text{true}$ or $G = \text{true}$ or both

■ **Iconic semantics**

a. The talk was long.

b. The talk was looong.

c. The talk was loooooong.

d. If the talk is long, I might stay. But if it's loooong, I won't.

A_1 longer than $A_2 \Rightarrow A_1$ denotes a longer period than A_2

Iconic Semantics

- Sign language has the same type of grammatical and logical resources as spoken language + **far richer iconic resources.**
- **POSS-1 GROUP GROW_little/medium/large.**

'My group has been growing.' ([ASL, 8, 263](#))

V_1 broader than $V_2 \Rightarrow V_1$ denotes larger growth than V_2

V_1 faster than $V_2 \Rightarrow V_1$ denotes faster growth than V_2

Narrow endpoints small amount, slowly	Medium endpoints medium amount, slowly	Broad endpoints large amount, slowly
 	 	 

Music Semantics

Schlenker 2017, 2019; Migotti 2019

- a. **Proposal:** a music semantics *can* be developed, but based on a non-standard notion of meaning:
the meaning of an auditory percept is the set of inferences it licenses on its causal sources (corresponding to 'voices')
- b. This semantics resembles but is more general than an iconic semantics: auditory trace vs. visual trace of an event.

■ Source-based semantics for music

- a. normal auditory cognition licenses inferences on 'virtual sources' = Bregman's 'auditory scene analysis';
- b. virtual sources are understood to be located in an abstract space isomorphic to tonal pitch space => stability, attraction, = 'journey through tonal pitch space' (Lerdahl, Steedman, ...)

Signaling the end

■ Inferences from normal auditory cognition

a. Lowering the volume ('decrescendo')

=> the source is losing energy, OR moving away

b. Diminishing the speed

=> the source is losing energy

c. Lowering the frequency

=> fewer events per time unit, hence less energy

■ Inferences from tonal pitch space

a. Cadence: ends in V - I

=> gradual movement to the point of greatest stability

b. Half-cadence: ends in V

=> movement to a point of partial stability

■ Ex.: Chopin's [Nocturne Op. 9/2](https://soundcloud.com/philippeschlenker/chopin-op9-2-better-115-end), last 2 measures.

Stylized Example

- **3 musical events** (= crescendo, maximal stability at edges)
<I, 70db>, <V, 75db>, <I, 80db>

- **Preservation rules** (Schlenker 2017, 2019)
The musical events can denote any triple of events satisfying:
 - a. **Time**: ordering of the denotations = ordering of the notes
 - b. **Loudness**: lower loudness is interpreted as (i) lower energy or (ii) greater distance
 - c. **Harmonic stability**: lower harmonic stability is interpreted in terms of lower (i) physical (ii) emotional stability.

- **The 3 musical events could for instance be true of:**
 - a. a sunrise (but not a sunset: crescendo wouldn't be correct)
 - b. a boat approaching (but not departing – same reason)
 - c. not a car crash (as the most unstable event is at the end)

Bernstein's Challenge

Bernstein's Challenge

■ **What does music mean? (1958)**

The true meaning of music is "the way it makes you feel when you hear it" (Young People's Concerts)
(different take on meaning in Bernstein's 1973 Harvard Lectures)

■ **Argument**

You can take a piece of program music, and tell the 'wrong' story – but things will work just as well.

■ **Example: Strauss's Don Quixote**

a. Variation II is about Don Quixote (with Sancho Pança) mistaking a flock of sheep for an army, attacking it, and ending up very proud of his knightly deed.

b. But it can be retold as a story about Superman!

Response: Bernstein Against Bernstein

- **Music semantics as defined must be very abstract:** we do not expect program music to evoke (alone) specific scenes.
- **But diverse situations that make the music true should be structurally connected** (via preservation rules they obey).
- **Case in point: Bernstein's two interpretations**

Don Quixote departing	Superman charging along
Sancho chuckling	Superman whistles
Sheep going <i>baa-ba</i>	Prisoners snoring away peacefully
A shepherd is playing on his pipe	Imprisoned friend playing his kazoo.
Don Quixote charges at the sheep	Superman charges into the prison yard
The sheep run off baaing wildly (and become more distant)	With the snoring still going on , Superman carries his friend away.
Don Quixote is convinced he has done a truly knightly deed, and is he proud!	Superman (with his friend) at last reaches freedom!

Case Study 2: Dissonances

■ Dissonances evoking chaos (temporal alignment plays a role too)

They see a **flock of sheep** in the field going *baa-baa*.

Superman hears all **the prisoners** snoring away peacefully .

■ Dissonances (original, simplified Midi)

Langsam

■ No dissonances (Bonetto)

Langsam

Case Study 3: Crescendo

■ Crescendo evoking the sheep (+ shepherd) approaching

Langsam

p *cresc.*

■ [Dissonances < (Bernstein, 1943)]

■ All <, as in the score (simplified Midi)

■ All > (simplified Midi)

■ [Dissonances <, melody >] (simplified Midi)

■ [Dissonances >, melody <] (simplified Midi)

Case Study 4: Cadence

■ Cadence evoking a triumphant completion

Don Quixote is convinced he has done a truly knightly deed, and is he proud!

Our hero [= **Superman, with his friend**] at last reaches freedom!

■ Expected chord (I) at the end (original)

Musical score for the expected chord (I) at the end. The score is in 4/4 time and G major. It consists of two measures. The first measure has a piano introduction marked *ff*. The second measure has a piano introduction marked *sfz* and a piano introduction marked *fff*. The piano introduction is a chord of G major.

■ [Diminished 7th] (Bonetto)

Musical score for the diminished 7th chord at the end. The score is in 4/4 time and G major. It consists of two measures. The first measure has a piano introduction marked *ff*. The second measure has a piano introduction marked *sfz* and a piano introduction marked *fff*. The piano introduction is a chord of G major.

■ Cluster (Bonetto)

Musical score for the cluster chord at the end. The score is in 4/4 time and G major. It consists of two measures. The first measure has a piano introduction marked *ff*. The second measure has a piano introduction marked *sfz* and a piano introduction marked *fff*. The piano introduction is a chord of G major.

**Referential indices:
Language, Visual Narratives, and Music**

Referential Indices

■ Referential indices regulate coreference in language

a. **English:** indices are invisible

Sarkozy_b told Obama_a that **he**_{a/b} would win the election.

b. **French Sign Language:** indices are visible

SARKOZY_b CL-b OBAMA_a CL-a a-TELL-b **IX-a/b**
WILL WIN ELECTION

■ Referential indices are needed in visual sequences

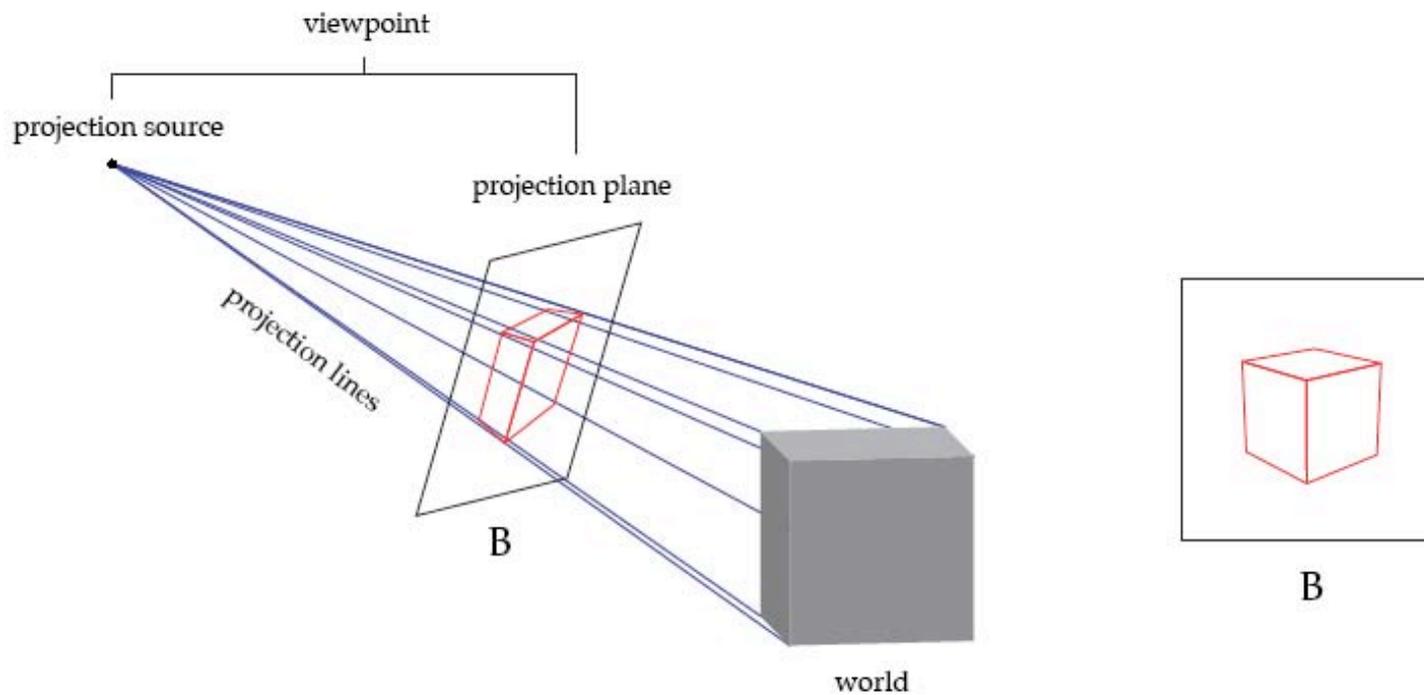
(Abusch 2015)

■ Referential indices might be needed in music as well.

Picture Semantics

Greenberg 2013

■ Perspective projection



Visual Narratives vs. Music

■ Picture sequences true of tuples of events (after Abusch)

A picture sequence $\langle P_1, \dots, P_n \rangle$ is true of events $\langle e_1, \dots, e_n \rangle$ **relative to viewpoint v** along the system of projection S iff

(1) temporally, $e_1 < \dots < e_n$;

(2) $\text{proj}_S(e_1, v) = P_1$ and ... and $\text{proj}_S(e_n, v) = P_n$.

■ Musical sequences true of tuples of events

A musical sequence $\langle P_1, \dots, P_n \rangle$ is true of events $\langle e_1, \dots, e_n \rangle$ **relative to auditory point v** iff

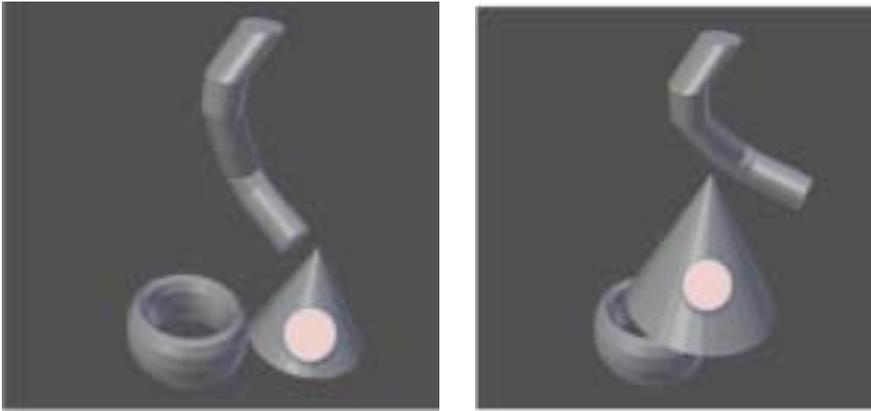
(1) temporally, $e_1 < \dots < e_n$;

(2) the Loudness and Harmonic stability conditions are satisfied for the relevant events relative to auditory point v .

Referential Indices in Visual Narratives

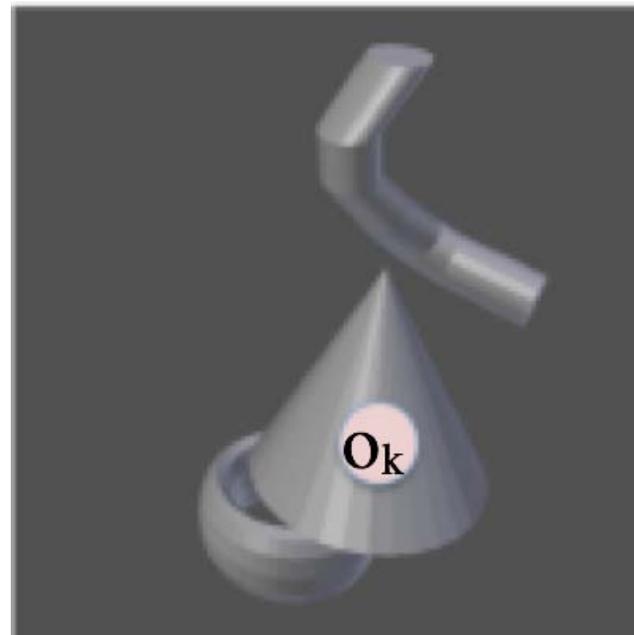
Abusch 2015

- **An ambiguity of coreference in pictures** (Abusch 2015)



- Abusch 2015: This "is consistent with worlds where a single cone moves in front of a torus. It is also consistent with worlds where the cone of the first picture moves out of view, and another cone moves into view."
- **Abusch's solution: referential indices**

Referential Indices in Visual Narratives (Abusch 2015)



[Referential Indices in Music: Don Quixote]

■ Variation II, beginning, Don Quixote departing

Var. II.
Kriegerisch.

ff (3 Solo Celli.)

4

■ Variation II, Don Quixote charging

Wieder doppelt so schnell. 3

3 Soli Celli.

ff

Referential Indices in Music

■ Adding variables

A musical sequence $\langle P_1(x_{i_1}), \dots, P_n(x_{i_n}) \rangle$ is true of events $\langle e_1, \dots, e_n \rangle$ relative to auditory point v and assignment function s (= providing values for referential indices) iff each $s(x_{i_k})$ takes part in e_k and

(1) temporally, $e_1 < \dots < e_n$;

(2) the Loudness and Harmonic stability conditions are satisfied (for all the relevant objects and events) relative to auditory point v .

Case Study: Chopin's Mazurka Op. 33, No 2

■ Structure

A1 - A2

A'1 - A'2

with A1 = A'1, A2 = A'2

■ Flat realization

[\[Midi\]](#)

■ Chopin's dynamics

[\[Midi\]](#)

f **A1 - A2**

pp **A'1 - A'2**

■ Britten's orchestration

[\[Britten\]](#)

Orchestra **A1 - A2**

Oboe+flute **A'1 - A'2**

Case Study: Chopin's Mazurka Op. 33, No 2: Ballet

- **Michel Fokine's ballet Les Sylphides**
(1984, close to Britten's version)
Main ballerina **A1 - A2**
Other dancers **A'1 - A'2**

Case Study: Chopin's Mazurka Op. 33, No 2: Orchestration

- **Orchestration:** Britten-like [\[Simplified Midi\]](#)
 - Strings **A1 - A2**
 - Flute **A'1 - A'2**

- **Alternative orchestrations**
 - a. Anti-Britten [\[Simplified Midi\]](#) (Bonetto)
 - A1** **A2**
 - A'1** **A'2**

 - b. Red assertive - Blue fearful (Bonetto)
Red (+ piano) is faster, louder, more accented (staccato).

 - c. Red fearful - Blue assertive (Bonetto)
Blue (+ piano) is faster, louder, more accented (staccato).

Mazurka.

F. CHOPIN. Op. 33, No 2.

23. *Vivace.*

The musical score is presented in four systems, each with a grand staff (treble and bass clefs). The key signature is D major (two sharps) and the time signature is 3/4. The tempo is marked *Vivace.* The score includes various musical notations such as slurs, fingerings, and accents. The first system includes a red box around a forte (*f*) dynamic marking. The second system includes a red box around a pianissimo (*pp*) dynamic marking. The score features various musical notations such as slurs, fingerings, and accents.

f

pp

Music as Gesture: Cosuppositions

Typology of Inferences

- How does the meaning of music interact with other mediums?
- In combination with other meaning-bearing forms:
 - a. gestures give rise to a rich typology of linguistic-like inferences;
 - b. music does too.
- In particular,
 - a. co-speech gestures give rise to characteristic conditional inferences (called 'cosuppositions');
 - b. co-speech music does to...
 - c. and it might be that co-film music does as well!

Co-speech vs. Pro-speech Gestures

■ Co-speech gestures

Your brother, I will _punish.

■ Pro-speech gestures

Your brother, I will .

Note: some examples discussed below involve objectionable actions.

Co-speech vs. Pro-speech Gestures

■ Co-speech gestures



My brother, will you _punish ?

=> if you punish my brother, slapping would be involved

■ Pro-speech gestures



My brother, will you ?

≠> if you punish my brother, slapping would be involved

Presuppositions and Cosuppositions

■ Presuppositions

a. At 12:05, will the company's plane take off?

=> right before 12:05, it will be on the ground

b. At 12:05, will the boss continue to smoke?

=> right before 12:05, the boss will be smoking

■ Cosuppositions are conditionalized presuppositions that are characteristic of co-speech gestures

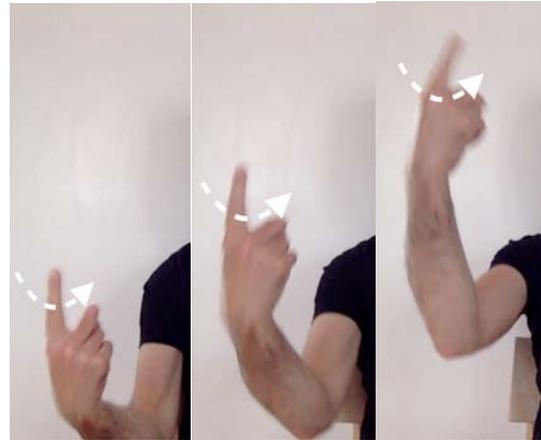


My brother, will you _punish ?

=> if you punish my brother, slapping would be involved

Co-speech vs. Pro-speech Gestures

■ Co-speech: cosuppositions



Will this thing ROTATION_ take off?

=> if this thing takes off, rotation will be involved

■ Pro-speech: assertions, presuppositions, implicatures...

Will this thing take off / ROTATION ?

Presupposition: this thing is on the ground

Assertion: this thing will take off

Pro-speech Music

- The full typology of linguistic inferences (implicatures, presuppositions, supplements, etc.) can be replicated with:
(i) gestures; (ii) visual animations. (Schlenker 2019; Tieu et al. 2019)



- Will this thing ?
=> this thing is on the ground (a presupposition!)



- This light bulb, you shouldn't .
=> the light bulb is on the ceiling (a presupposition!)

- **Guerrini and Migotti 2019** show that these results extend to:
a. pro-speech onomatopoeias;
b. **pro-speech musical snippets.**

Co-speech Gestures and Pictures: Cosuppositions

■ Co-speech cosuppositions

When a gesture or picture G accompanies a verbal construction V , it triggers the cosupposition that

$$V \Rightarrow G$$

■ Will Asterix DRINK [do what's needed]?

$V =$ do what's needed

$G =$ DRINK

Presupposition: if Asterix does what's needed, he'll drink.

What will happen next: will Asterix ...

...do what's needed?



Co-speech Sound [and Music]: Cosuppositions

Pasternak 2019

- The soldier will not BOOM [assassinate his target].
=> if the soldier were to assassinate his target, it would be via explosion
- The soldier will not assassinate his target like this BOOM.
≠> if the soldier were to assassinate his target, it would be via explosion
- The student will not DOWN [adjust the brightness setting of his computer screen].
=> if the student adjusts the brightness, this will involve turning it down

Co-speech Music: Cosuppositions

- Do you think your new student will MUSIC [take part in this afternoon's end-of-the-year competition]? [\[audio\]](#)
- V = take part in this afternoon's 1st year competition
G = MUSIC
Cosupposition:
if you new student takes part, they will play Für Elise.
- [Phlegmatic pianist, to the mayor of a besieged city]
Sir, I am told the enemy is about to enter the city. Will our cavalry MUSIC [do precisely what's needed at the present moment]? [\[audio\]](#)
Cosupposition:
if the cavalry does what's needed, it will come quickly and triumphantly

Co-film Music: Cosuppositions?

■ Co-film cosuppositions

- a. When a musical piece G accompanies a film snippet V (within a sentence), it triggers the cosupposition: $V \Rightarrow G$
- b. Open question: does this extend to normal film music?

■ Target: I wondered: Will it FILM + MUSIC?

Inference: if [film action], then [musical meaning]

■ If the ape breaks bones, this will be ...

a terrible thing? something light-hearted? something positive? an accomplishment? a fateful action?

■ Version 1 Version 2 Version 3 [Version 4 Version 5]

■ ≠ I wondered: Will it break bones in a terrible / light-hearted / positive / fulfilling / fateful fashion?

Conclusion

- Music triggers inferences about the extra-musical reality, albeit **very abstract ones** => solution to Bernstein's challenge
- Music semantics can learn from the semantics of pictorial narratives: **need for referential indices in music**
=> performing or orchestrating a piece, or adding dance to it, involves **resolving some implicit referential indices**.
- Like pro-speech gestures, pro-speech music can trigger diverse types of inferences (Guerrini and Migotti 2019)
- Like co-speech gestures, co-speech music can trigger **cosuppositions** (Pasternak), **but so can co-film music**.
= not a narrowly linguistic effect + new questions about film
- Experimental work is needed to test... **numerous** effects.