Cooper's event structures in Schlenker's musical semantics

TERM PAPER: SEMANTICS AND GRAMMAR
LECTURER: PAUL DEKKER

Department of Philosophy
University of Amsterdam

Mick de Neeve
9305017/mick@live.nl

DRAFT – November 13, 2019

Abstract Schlenker has proposed a notion of musical meaning based on inferences that may be drawn from virtual sources which are perceived while listening. His approach links coarse specification of musical events to similarly coarsely described world events, in order to have a semantics where music refers to a reality that is external to itself. Cooper on the other hand describes events in the world as well as in music in a structured and fine-grained manner, especially the latter. While Cooper favours a direction in musical semantics not unlike Schlenker, he does not explicitly link his structured descriptions of musical events to correspondingly structured events in the world. This paper provides musical examples which demonstrate that it is possible and viable to do just that.

1 Introduction

As Philippe Schlenker points out in *Prolegomena to Music Semantics* [Sch19], the study of musical semantics has not led to formal developments comparable to the study of music syntax. In fact, the subject has received very little attention, and Schlenker thinks this is down to many authors believing there is no such thing as a semantics of music. He terms this the *"Null Hypothesis"*, which plausibly has its origin in the position that music has neither truth conditions nor denotations. But according to Schlenker, musical semantics is *"a rule-governed way in which music can provide information (i.e. license inferences) about some music-external reality, no matter how abstract"*. So in effect, he takes a view of meaning for language on which semantics is a rule-governed relation between linguistic form and a reality external to language which in that case may be considered as truth conditions or denotations, and then relaxes restrictions concerning this relationship. However, what language and music have in common is that both are assumed to convey information, but the information music conveys is more abstract.

Broadly speaking, Schlenker conjectures that listeners identify so-called virtual sources in music which are then perceived as participating in particular sorts of events. The sources are considered as akin to denotations, while the events are infered. A virtual source may be an instrument in an orchestra but need not be, as a single instrument is capable of suggesting the presence of several sources (e.g. by playing a lower as well as a higher part), while an orchestra may at some point play 'in a single voice'. These sources, once identified, play the same role in perception as real sources in the real world, and the listener may then be led to conclude that the source is undergoing some sort of change, such as moving closer or gaining strength.

As Schlenker indicates, this sort of inference is not unique to music, but is more fundamental in that it relies on auditory perception. For Schlenker, this is the base level of semantic inferences in music, but it is insufficient in itself as it does not cover phenomena such as tonal attraction. As the author points out, the ending of a piece of music may be signalled by an auditory inference of moving away or getting weaker by decreasing loudness, but also by the harmonic device of resolving to the tonic, i.e. going from a state of tension (instability) to resolution (stability). Ultimately, the aim for Schlenker is to formulate a notion of musical truth: *"a voice undergoing a musical movement m is true of an object undergoing a series e of events just in case there is a certain structure-preserving map between m and e"* ([Sch19], page 38) – despite music not being considered as having truth conditions. Note that such events being evoked, or infered, may themselves be silent, such as a sunrise.[1]

Robin Cooper develops a method for representing the perception of events in *Type theory, interaction and the perception of linguistic and musical events* [Coo13]. A central tennet in his approach is to consider the perception of invariance as the perception of a type. Cooper's types can be made to apply to objects and to situations, and may be concatenated as well as embedded to describe complex events that may involve repetitions. They are associated with regular languages, which is not to say that regular grammars suffice to describe music (or language), but rather that these are expressive enough to cover type assignments during the perception of language and music, and to represent the subsequent perception of events.

Cooper gives an analysis of Beethoven's first Rasumovsky string quartet (*ibid*, pages 75-79) in terms of simple events represented by what he terms string types,[2] and more complex so-called record types. The former can represent elementary events and/or actions plus the way agents participate in them, while the latter allow subordinate simultaneously occuring events to be expressed. For instance, the author gives record types expressing motion from a half-strong to a half-weak beat while the pitch moves from $B\flat$ up to d (page 78). This illustrates the level at which the analysis takes place: it is essentially another

[1] Jean Sibelius apparently does this in the finale of his *Symphony No. 5*, but for a 2000 Swedish Radio Symphony Orchestra performance, it was deemed necessary to add visuals. See `http://youtu.be/nkzrSZKA4cM?t=130`.

[2] Or *regular* types – it is actually these which are strictly speaking associated with regular languages.

representation of what is given in the musical score, and hence does not go beyond a description of basic musical sound events. In other words, it falls well short of the sort of semantic representations that Schlenker has in mind.

In fact, by essentially re-representing the musical score, Cooper's approach does not even cover Schlenker's base level of auditory perception semantics. This is somewhat remarkable, since on page 85 (*ibid*), Cooper effectively endorses Schlenker's approach by referencing a lecture by Ginzburg where pieces of music are paired with video scenes, some of which are considered to appropriately reflect what the music expresses, while others are not.

The goal of this paper is to demonstrate that Cooper's type representations are in fact capable of representing to a sufficient degree the level of semantic analysis for music which Schlenker has in mind. As Schlenker indicates, for a series of (infered) events to accurately represent the semantic content of a given musical movement, there needs to be a structure-preserving map between the two. In the remainder, it will be shown that such mappings can indeed be said to exist between Cooper's score rehashings and the sort of Ginzburg scene pairings he mentions, which are effectively what Schlenker's approach boils down to as well.

2 Schlenker and Cooper: events and semantics

This section outlines in greater detail how Schlenker and Cooper view events, as well as their role in perception, cognition, and meaning. It concludes with some brief comments on the role of events in semantics.

2.1 Schlenker

As indicated, Schlenker develops his take on musical meaning by considering a number of possibilities, including that there is no such thing as a musical semantics, which is ultimately rejected on the grounds that certain inferences drawn from music are considered to be semantic in that they refer to extra-musical phenomena ([Sch19], section 3.2, page 44). Beside that, the author looks at theories that consider musical semantics as something internal, in other words that meaning is somehow about the music itself. In this respect he mentions *Emotion and Meaning in Music* by Meyer [Mey56], according to whom musical events have meaning because they refer to and make the listener expect other musical events. For Meyer, it is these expectations which give rise to emotional effects in listeners, as expectations may be fulfilled, postponed, or left unsatisfied. But the emotions are not the meanings itself, not for Meyer nor for Schlenker. For the latter, the semantics can be identified with the inferences which are licensed by the musical form, and these inferences are about events which have a structural similarity to that musical form. Emotional effects arise in virtue of the events which are infered.[3]

As indicated in section 1, music may evoke events such as a sunrise in a lis-

[3] Note that on Meyer's view, familiarity with a piece of music would neutralise emotional response, since the listener, being always 'right', would no longer be surprised or fulfilled by unsatisfied or satisfied expectations. Schlenker's account does not suffer from this drawback.

tener. This is the 'aboutness' relating to something outside the music which is essesntial for a musical semantics proper in Schlenker's view. In the aforementioned section 3.2 of [Sch19], Schlenker considers the famous opening of Richard Strauss' *Also Sprach Zarathustra*, used in the opening sequence of the equally famous Stanley Kubrick film *2001: A Space Odyssey*.[4] The author notes that the film synchronises a sun appearing from behind a planet with the music, and indicates that the event of this appearance is the sort of inference which the music licenses, and hence such an inference is considered as part of the musical semantics. As Schlenker indicates, this is in part down to the music evoking the development of a phenomenon in stages due to its having an antecedent-consequent form, but also to loudness as well as to the way in which the musical harmony plus melodic development leads initially to a first and then to a more assertive second climax. In other words, auditory perception (loudness) and pitch space (harmony) as well as melodic contour, are needed in order for this evocation to work.

2.1.1 Inferences from normal auditory cognition

Schlenker assumes the listener has identified the virtual sources about which inferences are to be drawn, either by contrapuntal or voice leading principles of classical music, or by more fundamental principles of auditory perception. In section 4 (*ibid*), he lists several examples of the latter category on the basis of which this may be achieved. Among these are instrument timbre, alteration of sound and silence, and related to this, the tempo at which a passage is played.[5] Additionally, Schlenker mentions loudness and pitch height.[6] Beside these, the author notes imitation, which generally belongs to the realm of so-called programme music,[7] and repetition. As for the latter, Schlenker mentions cases where repetitions are altered, e.g. diminishing in volume. It is then the fact that there is a repetition which allows the listener to identify a particular virtual source, and the alteration that leads the listener to infer it is (e.g.) moving away. But a repetition may also be interpreted in terms of a dialogue between two sources.

According to Schlenker, inferences are to be based on clearly stated hypotheses, for instance, that higher-pitched sources have greater energy. They should be assessed by constructing so-called minimal pairs, and posing questions such as *"Which of these two pieces evokes a phenomenon with the greater level of energy?"* (*ibid*, section 4.8). Moreover, it then needs to be demonstrated that non-musical cognition triggers the same inferences.

[4] `http://youtu.be/a1pqRbQypqM`

[5] As the author points out, speed tells the listener more than merely which virtual source is active, since acceleration is also associated with an increase in energy emitted by that source.

[6] Like tempo, loudness is associated with energy, while pitch differences relate to differences in the sizes of various virtual sources.

[7] A well-known example is Serge Prokofiev's *Peter and the Wolf*, where instrument imitation, such as the oboe imitating the sound of a duck, is intended to support an explicit story narrative. So-called pure music falls outside this realm, since any narrative is implicit and is to be infered.

2.1.2 Inferences from tonal properties

Schlenker points out that semantic effects which have been discussed before, e.g. in Lerdahl's *Tonal Pitch Space* [Ler01], often make reference to motion, for instance to a concept of magnetism, where during harmonic resolution from unstable to stable chords, at the note level notes move to the closest stable notes.[8] The author mentions Mark Granroth-Wilding and Mark Steedman [GWS14], who implement a system that casts musical meaning as a journey through tonal pitch space. In this respect, Schlenker says there is a temptation to reduce musical meaning to musical tension, but according to him, although tension is crucial to the semantics, it doesn't cover the whole story. Rather, understanding properties of tonal pitch space is important for understanding the events in which the virtual sources partake that have been identified. But ultimately for a full (or fuller) understanding, these properties are to be combined with those infered using normal auditory cognition.

2.1.2.1 Cadences and modulations

A standard way of dealing with tension in music is through so-called cadences, which represent particular ways of moving from points of relative harmonic instability to stability. As Schlenker puts it, cadences are the standard way of marking the end of a (classical) piece of music ([Sch19], section 5.3). More generally, it does not have to be the end of an entire piece; a cadence may just as well end a particular section, or a phrase, after which the music may proceed into a different direction or elaborate on the preceding part by means of variations.

Schlenker refers to Fred Lerdahl and Ray Jackendoff's *A Generative Theory of Tonal Music* [LJ83], and points out that in this work, cadences are viewed as syntactic devices. Lerdahl and Jackendoff use cadences in the construction of syntactic trees, almost as if they were on a par with the full stop that marks the end of a sentence. But Schlenker suggests that cadences may be better understood when considered from a semantic perspective. He compares the motion towards stability to auditory features such as speed and loudness which may both be decreased to similarly signal the end of a piece of music – or indeed of a smaller unit such as a phrase. The author gives the example of the famous *Ah vous dirai-je Maman* theme by Wolfgang Amadeus Mozart, and compares the situation of a so-called perfect cadential motion towards maximal stability to one where the phrase (or rather the consequence of a phrase) ends in a less stable state – called a deceptive cadence.[9] The latter case is incidentally what Meyer [Mey56] might have highlighted as an example of an unfulfilled expectation, or a postponed one – as indicated in section 2.1 here – since it is not difficult to imagine the deceptive cadence to be followed by a perfect cadence. And if one does, then an inference along the following lines may in

[8] The best-known example is the motion from a dominant to a major chord, e.g. from $G7$, consisting of g, b, d, and f, to $C\triangle$, which has the notes c, e, g, and b. The instability of the former chord is due to the presence of the flat fifth interval $[b, f]$ – which splits the octave in half and was called the devil's interval in medieval times – and it is these notes that feature most strongly in the resolution to $C\triangle$, with b attracted to c a semitone up, and f attracted to e a semitone down.

[9] Cf. `http://bit.ly/2DohwYa` and `http://bit.ly/2D7fMEI`.

fact be licensed: a source moves away from where it was, then moves back but hovers around close by at first, and only then returns to its point of origin, or 'home'. Although this sort of motion may well be suggested by purely auditory means, e.g. a gradual change in loudness that only returns to the original volume after lingering at a point close to the initial loudness, it is arguably the case that movement in pitch space is more effective at suggesting movement away and towards – possibly in conjunction with said auditory features.

According to Schlenker, modulation, or key change, can be viewed as movement to a different region. By extension, cadential motion would then be movement within a region, since it occurs within a single key. The author provides an example in the form of *Le cygne* by Camille Saint-Saëns, and gives minimal pairs to illustrate the omission of Saint-Saëns's original modulation from *G major* to *B minor*.[10] Schlenker conjectures that the purpose of the modulation is to suggest the exploration of an area with a different landscape, but arguably, different interpretations are possible, especially considering that this particular piece of music features in Anna Pavlova and Michel Fokine's famous ballet based on Alfred Tennyson's poem *The Dying Swan*. Since *B minor* is less stable than *G major*,[11] the modulation may also be viewed as a change of state of the source within the same region. Arguably, this is then signalled by the note change from *c* to *c*♯, which is the only difference between the keys of *G major* and *B minor*. A more daring conjecture outside the scope of this paper would be to posit that the meaning of the note *c*♯ is that state change.

2.1.3 Musical truth

Schlenker aims for a notion of musical truth based on events. While a musical passage may not be true or false as such, it may be appropriate to say that it is true of a certain type of situation, described by an event or set of events. The author gives an example in [Sch19], section 4.2, using part of *Le carnaval des animaux* (movement VI: *Kangourous*) by Saint-Saëns.[12] He points out that a series of short notes separated by rests are meant to evoke the idea of kangaroos jumping, but stresses (*ibid*, section 6) that it is not supposed to be infered that kangaroos are producing the sounds. Instead, the inference would be more abstract, to the effect that there was a rapid succession of events, and kangaroos jumping would be one instance which satisfies this. In effect, kangaroos jumping would then be part of the semantic content of the *Kangourous* movement, as it is included in the set of objects and associated situations the movement is true of (section 6.3).

As a basis to work from, Schlenker specifies the format of a voice M, regardless of whether it represents an instrument in an orchestra or ensemble, a contrapuntal line, or a virtual source that the listener has identified – which may in fact be several instruments collaborating to produce some intended effect. In section 6.2, he returns to Strauss' *Zarathustra* (cf. section 2.1 here), and out-

[10] See `http://bit.ly/2D6TcNq` for the original with the modulation, and `http://bit.ly/2DqCC80` for a version without it. Note that *G major* to *B minor* is a real modulation, since the major parallel of *B minor* is *D major*.

[11] [Sch19], section 3.2.

[12] `http://bit.ly/2m98kPd`

lines a reduction of musical events to sequences of pairs consisting of chord types coupled with loudness, where the chord type is given as a church scale mode in order to characterise its stability, and loudness is stated in terms of decibels. For instance, he renders the beginning of the piece as in equation 2.1:[13]

Equation 2.1.

$$M = \langle\langle I, 70dB\rangle, \langle V, 75dB\rangle, \langle I, 80dB\rangle\rangle$$
$$M' = \langle\langle I, 70dB\rangle, \langle IV, 75dB\rangle, \langle V, 80dB\rangle\rangle$$
$$M'' = \langle\langle IV, 80dB\rangle, \langle V, 75dB\rangle, \langle I, 70dB\rangle\rangle$$

Here, Schlenker notes that the chord built on the fifth mode, V, is less stable than the one on the first mode, I (see footnote 8 in this paper). In other words, what matters is that as the passage increases in volume, which the listener understands through ordinary auditory cognition, it moves from stability to instability and back again (in pitch space).

Schlenker then specifies what a possible denotation is given a set of musical events such as indicated for the M in *Zarathustra* above:

Definition 2.2. *Let M be a voice, with $M = \langle M_1, \ldots, M_n\rangle$. A possible denotation for M is a pair $\langle O, \langle e_1, \ldots, e_n\rangle\rangle$ of a possible object and a series of n possible events, with the requirement that O be a participant in each of $e_1, \ldots, e_n$.*

The idea behind definition 2.2 is to associate voices that participate in musical events with objects participating in possible events in the world. For the *Zarathustra* example, it is intended to link the situation of sunrise (as aforementioned in section 2.1 here) with harmonic motion from a chord on scale mode I via mode V back to mode I (i.e. from stability to instability and back again), combined with increasing and decreasing loudness, i.e. with the musical events in equation 2.1.

Schlenker then defines what it means for a musical voice to be true of an object in the world given the events that the object participates in.[14]

Definition 2.3. *Let $M = \langle M_1, \ldots, M_n\rangle$ be a voice, and let $\langle O, \langle e_1, \ldots, e_n\rangle\rangle$ be a possible denotation for M. M is true of $\langle O, \langle e_1, \ldots, e_n\rangle\rangle$ if it obeys the following requirements.*

(a) *Time*
The temporal ordering of $\langle M_1, \ldots, M_n\rangle$ should be preserved, i.e. we should have $e_1 < \ldots < e_n$, where $<$ is ordering in time.

(b) *Loudness*
If M_i is less loud than M_k, then either:

(i) *O has less energy in e_i than in e_k; or*

(ii) *O is further from the perceiver in e_i than in e_k.*

[13] [Sch19], page 66.
[14] [Sch19], page 67.

(c) *Harmonic stability*
If M_i is less harmonically stable than M_k, then O is in a less stable position in e_i than it is in e_k.

Any piece of music will be true of a great deal of objects and events in the world, the author points out, but likewise, a sentence such as *"It is raining"* similarly refers to numerous situations. Still, the denotations in the world that music can be true of are extremely heterogeneous, according to Schlenker, but this is due to informational content in musical pieces being abstract and underspecified. The author notes that this has led some authors to conclude that music has no semantics at all, but for him, it simply means that musical semantics is equally abstract and underspecified.[15]

Schlenker then turns to the first part of equation 2.1, i.e. from Strauss' *Zarathustra*, and notes that it is true of ts(a) and *(c)* in equation 2.4, but not of *(b)*, *(d)*, or *(e)*.[16]

Equation 2.4.

(a) $\textit{Sun-rise} = \langle \textit{sun}, \langle \textit{min-luminosity}, \textit{rising-luminosity}, \textit{max-luminosity} \rangle \rangle$
(b) $\textit{Sun-set} = \langle \textit{sun}, \langle \textit{max-luminosity}, \textit{diminishing-luminosity}, \textit{min-luminosity} \rangle \rangle$
(c) $\textit{Boat-approaching} = \langle \textit{boat}, \langle \textit{max-distance}, \textit{approach}, \textit{min-distance} \rangle \rangle$
(d) $\textit{Boat-departing} = \langle \textit{boat}, \langle \textit{min-distance}, \textit{departure}, \textit{max-distance} \rangle \rangle$
(e) $\textit{Car-crash} = \langle \textit{car}, \langle \textit{movement}_1, \textit{movement}_2, \textit{crash} \rangle \rangle$

In other words, although M is true of many different sorts of events, there are clear limits to what it can be true of.

Schlenker links his take on the semantics of music to that of pictures, where a triangle may be said to correctly represent some scene in case there is a mapping from scene to triangle that preserves chief geometric aspects. The author views musical semantics as having a similar structure as other inferential systems in perception (*ibid*, section 6.4), while differing considerably from logical semantics. A point in case is that music may be rendered as a series of concatenated symbols to represent musical events, e.g. as p_1p_2, which is not the same as a logical conjunction $p_1 \wedge p_2$, since has the same meaning as $p_2 \wedge p_1$, while a reversed concatenation of musical events p_2p_1 may be true of an entirely different set of situations in the world than p_1p_2 would. This is witnessed by the fact that Strauss' *Zarathustra* is true of the *Sun-rise* event, but not of *Sun-set*.

2.2 Cooper

Unlike Schlenker, Cooper's focus in [Coo13] is not semantics or musical semantics per se, but the structure of events, how they may be coordinated and how the perception of events by an agent can be represented. That being said, towards the end of his article (in section 3), Cooper does suggest that music

[15] Schlenker does not intend definition 2.3 to be exhaustive. The time condition being critical, loudness and harmonic stability have been picked as cases that are suitable enough for a few examples to illustrate directions for possible future research.
[16] [Sch19], page 68.

represents things outside its own realm, and refers to a lecture by Jonathan Ginzberg [Gin12] in which a strategy of pairing musical pieces with 'appropriate' and 'inappropriate' scenes is proposed that is very similar to Schlenker's approach. However, Cooper does not end up linking this idea to his proposed method of structuring events (which will be the topic of the next section of this paper, section 3).

Cooper's aim is to cast perceived invariance of objects and situations in the perception of events as the attunement of the perceiving agent to particular types. This can be expressed by the judgement $a : T$, meaning that object a is of type T. When objects (or agents) partake in situations, events may be predicates of a given arity, and in this case the author states that the type in question is a so-called *ptype*, where judgments take the form $e : P(o_1, \ldots, o_n)$. A situation, the author emphasises, is really just another way to characterise an event. A central tennet of Cooper's approach is that events can be complex, i.e. the type (or ptype) of an event may be another event, allowing for events to be embedded so that they may be coordinated serially (i.e. sequentially) or in parallel. This will be illustrated in the following section.

2.2.1 Events in the world

Example 2.5 below, which is example 8 on page 5 in [Coo13], represents an event e of type *'game of fetch'* involving a human, a dog, and a stick.[17] It illustrates the concatenation of serial subevents using the '⌢' symbol, as well as parallel coordination with constituent events appearing below eachother. The event e may be repeated any number of times, due to the Kleene plus at the end.

Example 2.5.

$$
\begin{aligned}
e : (&\left[\ e_h : \mathit{pick\_up}(h, s)\ \right] \frown \left[\ e_h : \mathit{attract\_attention}(h, d)\ \right] \frown \\
&\left[\ e_d : \left[\begin{array}{l} e_1 : \mathit{jump}(d) \\ e_2 : \mathit{bark}(d) \end{array}\right]\ \right] \frown \left[\begin{array}{l} e_h : \mathit{throw}(h, s) \\ e_d : \left[\begin{array}{l} e_1 : \mathit{jump}(d) \\ e_2 : \mathit{bark}(d) \end{array}\right] \end{array}\right] \frown \\
&\left[\ e_d : \left[\begin{array}{l} e_1 : \mathit{run\_after}(d, s) \\ e_2 : \mathit{bark}(d) \end{array}\right]\ \right] \frown \left[\ e_d : \mathit{pick\_up}(d, s)\ \right] \frown \\
&\left[\ e_d : \mathit{return}(d, s, h)\ \right])^+
\end{aligned}
$$

Cooper calls types such as those with e_2 below e_1 in example 2.5 *record types*. These, he states on page 4 of section 1, provide a way of specifying types which interact and are coordinated in time. The first two of these cases in example 2.5, the author writes, refer to the dog simultaneously jumping and barking. The third case, with two subevents that are of different durations, is plausibly about the dog barking a number of times while running after the stick. The second case, with $e_h : \mathit{throw}(h, s)$ at the top, underlines that such events do not simply take place in parallel but are correlated.[18]

According to Cooper, that types are structured allows types of events already

[17] For this particular example, Cooper has not put the event variable e outside the main parentheses, but he has done so in example 10 (on page 6), which is the same as 8, except that it contains a specification of the other variables.

[18] Here, it may either be the case that the dog jumps and barks because the stick is thrown, or conversely that the human throws the stick in response to the dog jumping and barking.

present to be reused (*ibid*, section 4, page 21), see for instance $pick\_up(x, s)$ in example 2.5 where x may be either h or d. Moreover, the author alleges, what types are available in the perception and cognition of events can vary from agent to agent. In musical perception, this may explain listeners associating different sorts of (related) events with the same piece of music.

2.2.2 Events in music

Cooper uses the first four bars of the second movement of Ludwig van Beethoven's *String Quartet No. 7* (the first *Rasumovsky* quartet, Opus 59/1) as an example of musical event coordination. The score is depicted in figure 1.

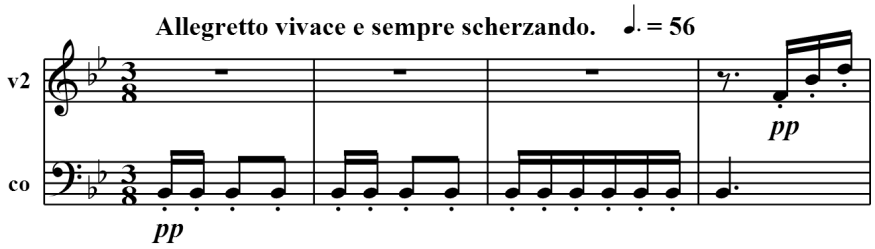

Figure 1: Second violin and cello from the start of Beethoven's first *Rasumovsky* quartet, second movement (extract from example 18 in [Coo13])

"In terms of musical interaction you cannot get much simpler than this", the author writes (*ibid*, section 1, page 9). This may be the case, but note that unlike Schlenker's examples, figure 1 does not represent a complete musical idea. The bass part (the cello) introduces the tonic ($B\flat$) in a series of repeated notes, and after this introduction ends, the second violin plays a $B\flat$ *major* triad, in the second inversion, ending on the 3^{rd}, i.e. in d, in a series of three rapid notes.[19] Example 2.6 below (which is example 24 in [Coo13]) gives Cooper's rendition in terms of coordinated events.

Example 2.6.

$$
e\text{:}\left[\begin{array}{l}
e_{v2}\text{:}\left[\begin{array}{l} beat{:}Strong \\ e{:}Silent \end{array}\right] \frown \\
\quad \left[\begin{array}{l} beat{:}Weak \\ e\text{:}\left[\begin{array}{l} beat{:}HalfStrong \\ e{:}Silent \end{array}\right] \frown \left[\begin{array}{l} beat{:}HalfWeak \\ pitch{:}F \end{array}\right] \end{array}\right] \frown \\
\quad \left[\begin{array}{l} beat{:}Weak \\ e\text{:}\left[\begin{array}{l} beat{:}HalfStrong \\ pitch{:}B\flat \end{array}\right] \frown \left[\begin{array}{l} beat{:}HalfWeak \\ pitch{:}d \end{array}\right] \end{array}\right] \\
e_{co}\text{:}\left[\begin{array}{l} beat{:}Strong \frown Weak \\ pitch{:}B\flat'' \end{array}\right]
\end{array}\right]
$$

As usual, serial coordination is specified for each event by horizontal concatenation (with the '$\frown$' symbol), and parallel coordination by vertical stacking.

[19] See http://youtu.be/f7vleQm1mg4?t=662 to hear the begining of this movement, and note that Cooper's example ends after the first three violin notes. A more definite musical form might be achieved by ending on the (longer) twelfth note that the violin plays, i.e. on the f (which is the 5^{th} of $B\flat$), but even this is at most a kind of antecedent that calls for an addition in order to achieve a completed idea.

An extra 'e :' – for the main musical event – has been added outside the main square brackets, in order to make the notation more consistent.

As mentioned (in section 1 here), this way of specifying events is quite a bit more detailed than Schlenker's, and can be viewed as an alternative representation of a musical score – even if at the same time it omits the dynamics in the score while Schlenker does render this by having louder and softer events.[20] The objective of section 3 here is to use Cooper's way of specifying coordinated events in a 'Schlenkerian' manner, i.e. to use structured events in the world as denotations of musical events. But first the role of events in semantics is highlighted, in the next section.

2.3 Events in semantics

Events are about situations, in particular about actions. Traditionally, actions, or verbs, have been analysed as n-ary predicates, e.g. *"Brutus stabbed Caesar"* would be rendered as $stabbed(brutus, caesar)$. As pointed out by Lucas Champollion in [Cha14] (page 1), a semantics centered on events can help emphasise the theme of actions, e.g. in the so-called neo-Davidsonian approach, the rendition would be $\exists e[stab(e) \wedge agent(e, brutus) \wedge theme(e, caesar)]$. This casts the verb as an event featuring $caesar$, in which $brutus$ is the agent (rather than simply having two otherwise undistinguished literals as in the predicate case). Adding *"at noon"* as the time of the event then allows entailments such as $\exists e[stab(e) \wedge agent(e, brutus) \wedge theme(e, caesar) \wedge time(e) = noon] \models \exists e[stab(e) \wedge agent(e, brutus) \wedge theme(e, caesar)]$, i.e. given that *"Brutus stabbed Caesar at noon"* it is also the case that *"Brutus stabbed Caesar"*. This may seem obvious, but the entailment does not hold as such between the predicated statements $stabbed(brutus, caesar, noon)$ and $stabbed(brutus, caesar)$.

In the above stabbing event, *"stabbed"*, *"stabbed Caesar"*, and *"Brutus stabbed Caesar"* all denote sets – that decrease in magnitude as the roles for the event are assigned. But $theme(e, caesar)$ also denotes a set, namely the set of events which have *Caesar* as theme. This is relevant here, see for instance equation 2.4 *(a)* and *(b)*. Using the notation of [Cha14], both have $theme(e, sun)$. While it is clear that Schlenker's *Zarathustra* rendition is true of *(a)*: $rise(e) \wedge theme(e, sun)$ but not of *(b)*: $set(e) \wedge theme(e, sun)$, it is also arguably true of $appear(e) \wedge theme(e, sun)$ (from behind the clouds). So there are particular sorts of 'sunny events' that *Zarathustra* is true of, which for a musical semantics, one would want to have together in a set, starting with $theme(e, sun)$. Conversely, $appear(e)$ denotes a set – which arguably includes $rise(e)$ – that can be restricted by adding $theme(e, sun)$.[21]

Noting that in Schlenker's approach, large and varied sets of events in the world can be taken as denotations for musical events, event semantics for natural language points to the possibility of zooming in to an appropriate denotational

[20] Of course, in example 2.6, dynamics could be easily added as parallel events to indicate the *pp* marking in the score, but this may actually be redundant since the fragment has an unvaried level of loudness.

[21] It is taken to be appropriate not to have agency on the sun's part, since rising is caused by planetary rotation, and appearing from behind clouds is due to motion of those clouds.

set in a step-wise fashion. In other words, it should be possible to specify more precisely the sorts of world events that musical events denote. This is one the aims of the next section, which is chiefly devoted to allowing more fine-grained structures for events in the world as well as for musical events (that is, using Cooper's [Coo13] structures in a Schlenkerian [Sch19] approach).

3 Structured events as musical denotations

An important difference in the approaches of Schlenker and Cooper is that the former specifies a precise relationship between musical events and events in the world, while only specifying events (in music as well as the world) very broadly. The latter on the other hand specifies events very precisely, but postulates no relationship between musical and world events.[22] Moreover, as noted in section 2.2.2, the only musical event example Cooper gives does not represent anything like a completed musical idea. So it is natural to consider the best of each author, and state musical and world events more precisely, while illustrating things using complete musical ideas – and then to specify the semantic relationship between events in music and in the world.

The example to carry this section is *Waltz No. 2* (Op. 99) by Dimitri Shostakovich, in particular the first twenty measures. The score for these is depicted as a violin and cello arrangement in figure 2.[23]

4 Discussion

5 Concluding remarks

References

[Cha14] L Champollion. Integrating Montague Semantics and Event Semantics. Lecture notes from the 26th European Summer School in Logic, Language and Information, University of Maryland, College Park, United States, June 23-27, 2014. https://ling.auf.net/lingbuzz/002143.

[Coo13] R Cooper. Type Theory, Interaction and the Perception of Linguistic and Musical Events. In M Orwin, C Howes, and R Kempson, editors, *Language, Music and Interaction*, pages 67–90. London: College Publications, 2013.

[Gin12] J Ginzberg. False starts, jokes, and music: semantics in the 21st century. Lecture at the 24th European Summer School in Logic, Language and Information, University of Opole, Poland, August 14, 2012.

[22] At most, Cooper hints at this, as mentioned in section 2.2.

[23] Arranged by the author of this paper; a sound rendition can be found at https://deneeve.github.io/ac/edu/sgr/sho.waltz2.mp3 (in a slightly slower version).

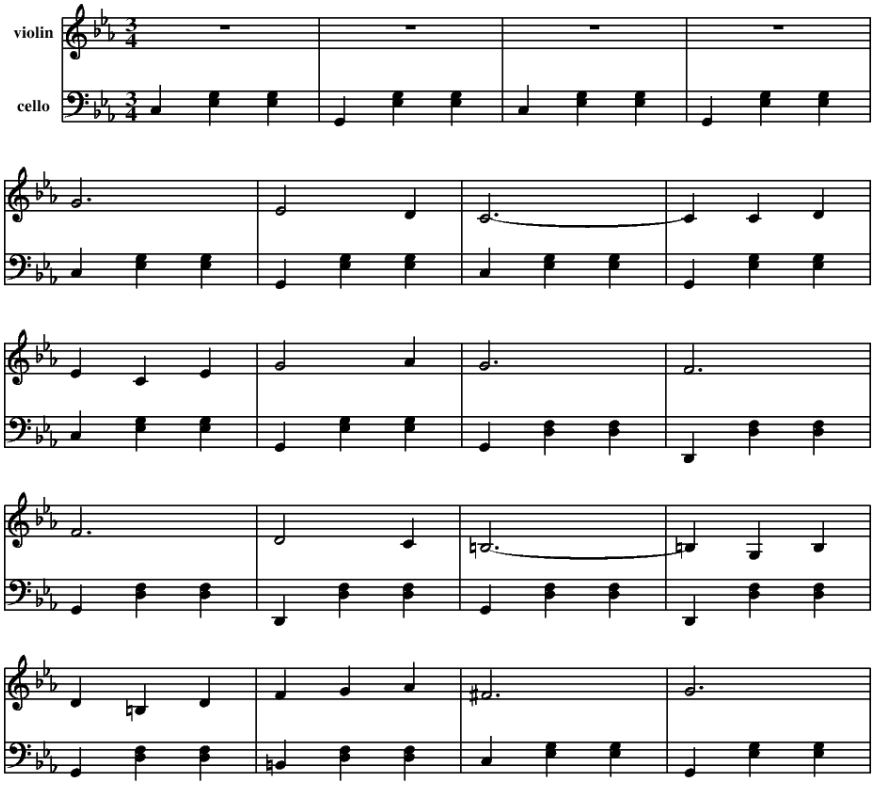

Figure 2: Opening bars from Shostakovich's *Waltz No.2*

[GWS14] M Granroth-Wilding and M Steedman. A Robust Parser-Interpreter for Jazz Chord Sequences. *Journal of New Music Research*, 43(4):355–374, 2014.

[Ler01] F Lerdahl. *Tonal Pitch Space*. Oxford University Press, 2001.

[LJ83] F Lerdahl and R Jackendoff. *A Generative Theory of Tonal Music*. MIT Press, 1983.

[Mey56] L Meyer. *Emotion and Meaning in Music*. University of Chicago Press, 1956.

[Sch19] P Schlenker. Prolegomena to Music Semantics. *Review of Philosophy and Psychology*, 10(1):35–111, 2019.