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

March 8, 2020

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 specifications 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's, he does not focus on semantics, and does not use examples which are clear musical ideas, nor does he 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. This is done by transforming sound events into harmonic events first, and in turn transforming the latter into world events, or denotations, using Schlenker's methods. Moreover, it is suggested that a dynamic approach to interpretation may help to make musical semantics compositional.

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, in *Type theory, interaction and the perception of linguistic and musical events* [Coo13], develops a method for representing the perception of events. A foremost feature in his approach is that the perception of invariance can be considered 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

[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.

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 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 main 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 listener. 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

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

[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.

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.

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

[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`.

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 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 $c\sharp$, which is the only difference between the keys of *G major* and *B minor*. A more daring conjecture would be to posit that the meaning of the note $c\sharp$ is that state change.[12]

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.[13] 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

[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] See also section 4.2.

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

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 outlines 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:[14]

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, or perhaps better, a candidate denotation is, given a set of musical events such as the Ms in equation 2.1:

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.[15]

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:

[14] [Sch19], page 66, equation 23.
[15] [Sch19], page 67.

(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.

(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.[16]

Schlenker then turns to the first part of equation 2.1, i.e. from Strauss' *Zarathustra*, and notes that it is true of *(a)* and *(c)* in equation 2.4, but not of *(b)*, *(d)*, or *(e)*.[17]

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 this 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

[16] 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.

[17] [Sch19], section 6.2, page 68.

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 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. Central to 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

Equation 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.[18] It illustrates the concatenation of serial subevents using the '$\frown$' 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.

Equation 2.5.

$$e\text{:} \left[\begin{array}{l} \left[\, \mathrm{e}_h : \mathit{pick\_up}(h,s) \,\right] \frown \left[\, \mathrm{e}_h : \mathit{attract\_attention}(h,d) \,\right] \frown \\ \quad \left[\mathrm{e}_d : \left[\begin{array}{l} \mathrm{e}_1 : \mathit{jump}(d) \\ \mathrm{e}_2 : \mathit{bark}(d) \end{array} \right] \right] \frown \left[\begin{array}{l} \mathrm{e}_h : \mathit{throw}(h,s) \\ \mathrm{e}_d : \left[\begin{array}{l} \mathrm{e}_1 : \mathit{jump}(d) \\ \mathrm{e}_2 : \mathit{bark}(d) \end{array} \right] \end{array} \right] \frown \\ \quad \left[\mathrm{e}_d : \left[\begin{array}{l} \mathrm{e}_1 : \mathit{run\_after}(d,s) \\ \mathrm{e}_2 : \mathit{bark}(d) \end{array} \right] \right] \frown \left[\, \mathrm{e}_d : \mathit{pick\_up}(d,s) \,\right] \frown \\ \quad \left[\, \mathrm{e}_d : \mathit{return}(d,s,h) \,\right] \end{array} \right]_{+}$$

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.[19]

[18] 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.

[19] 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.

According to Cooper, that types are structured allows types of events already 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.[20] Equation 2.6 below (example 24 in [Coo13]) gives Cooper's rendition in terms of coordinated events.

Equation 2.6.

$$
e{:}\left[\begin{array}{l}
e_{violin2}{:}\begin{bmatrix}beat{:}Strong\\ e{:}Silent\end{bmatrix}\frown\\
\left[\begin{array}{l} beat{:}Weak\\ e{:}\begin{bmatrix}beat{:}HalfStrong\\ e{:}Silent\end{bmatrix}\frown\begin{bmatrix}beat{:}HalfWeak\\ pitch{:}F\end{bmatrix}\\ beat{:}Weak\\ e{:}\begin{bmatrix}beat{:}HalfStrong\\ pitch{:}B\flat\end{bmatrix}\frown\begin{bmatrix}beat{:}HalfWeak\\ pitch{:}d\end{bmatrix}\end{array}\right]\frown\\
e_{cello}{:}\begin{bmatrix}beat{:}Strong\frown Weak\\ pitch{:}B\flat_{,,}\end{bmatrix}
\end{array}\right]
$$

[20] 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. Ending on the 3^{rd} is incomplete since the music either suggests wanting to return to the tonic (1^{st}), or to tentatively resolve to $e\flat$ (the 4^{th}) by semitone upward attraction. Alternatively, a more definite musical form might be achieved by ending on the (longer) twelfth note 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 musical idea.

As usual, serial coordination is specified for each event by horizontal concatenation (with the '$\frown$' symbol), and parallel coordination by vertical stacking. 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.[21] 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 $theme(e, sun)$.[22]

Noting that in Schlenker's approach, large and varied sets of events in the world

[21] 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.

[22] 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.

can be taken as denotations for musical events, event semantics for natural language points to the possibility of zooming in to an appropriate denotational 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.[23] 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 music central to this section is *Waltz No. 2* (Op. 99) by Dimitri Shostakovich, in particular the first 20 measures, although the formal semantic analysis focuses on bars 5 to 8 (as well as 13 to 16). The rest is discussed informally on the basis of this. The score for measures 5 to 8 is depicted in figure 2, and the first 20 measures of the piece are shown in figure 3.[24]

Figure 2: Bars 5 to 8 (last note extended) from Shostakovich's *Waltz No. 2*

3.1 From sound events to world events

Equation 3.1 gives the sound pitch events of the violin and cello for bars 5 to 8 of *Waltz No. 2*. The representation is based on how Cooper renders coordinated events in [Coo13], which uses the ABC letter notation standard for the pitches.[25]

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

[24] Arranged by the author of this paper for violin and cello; the respective sound renditions can be found at `https://deneeve.github.io/ac/edu/sgr/sho.waltz2.5-8.mp3`, and `https://deneeve.github.io/ac/edu/sgr/sho.waltz2.mp3`.

[25] See `http://abcnotation.com/blog/2010/01/31/how-to-understand-abc-the-basics/`, but note that unlike here, sharps (♯) and flats (♭) in ABC are represented using the '^' and '_' characters, respectively, since it is an ASCII format. Furthermore, z is used to represent a rest.

Figure 3: Opening bars from Shostakovich's *Waltz No. 2*

Equation 3.1.

$$e_{pitch}\colon \begin{bmatrix} e_{violin}\colon \begin{bmatrix} G \end{bmatrix}_3 \frown \begin{bmatrix} E\flat \end{bmatrix}_2 \frown \begin{bmatrix} D \end{bmatrix} \frown \begin{bmatrix} C \end{bmatrix}_6 \\ e_{cello}\colon \begin{bmatrix} \begin{bmatrix} C, \end{bmatrix} \frown \begin{bmatrix} B\flat, \\ G, \end{bmatrix}_2 \frown \begin{bmatrix} G,, \end{bmatrix} \frown \begin{bmatrix} G, \\ E\flat, \end{bmatrix}_2 \end{bmatrix}_2 \end{bmatrix}$$

Rather than explicitly stating rhythmic events as Cooper does, the above equation uses concatenations of crotchets, or quarter notes (♩), to represent the musical meter. For instance, the first violin note is a G of length 3, which is a half note (𝅗𝅥, which has length 2, i.e. a quarter note times 2) with a dot (𝅗𝅥.) to add half its length to make it a total length of 3. But even with all that information, these letters do not say very much in themselves. So in equation 3.2, a transformation has been made into the harmonic functionality of the notes.

Equation 3.2.

$$e_{harmony}\colon \begin{bmatrix} e_{violin}\colon \begin{bmatrix} v \end{bmatrix}_3 \frown \begin{bmatrix} iii \end{bmatrix}_2 \frown \begin{bmatrix} ii \end{bmatrix} \frown \begin{bmatrix} i \end{bmatrix}_6 \\ e_{cello}\colon \begin{bmatrix} \begin{bmatrix} i \end{bmatrix} \frown \begin{bmatrix} vii \\ v \end{bmatrix}_2 \frown \begin{bmatrix} v \end{bmatrix} \frown \begin{bmatrix} v \\ iii \end{bmatrix}_2 \end{bmatrix}_2 \end{bmatrix}$$

Since the piece is in *C minor*, lowercase Roman numerals are used in equation 3.2 to indicate the note roles. The important thing about the violin is that it descends from the fifth to the first degree of this minor chord. The cello has a dual role; it plays the first and fifth degrees as a bass line, and in between these it lays down a *C minor* 7^{th} chord in two parts. In Schlenker's terms the cello produces two voices, and these have been separated in equation 3.3.

Equation 3.3.

$$e_{harmony}\colon \left[\ e_{cello}\colon \left[\ \begin{array}{l} e_{chord}\colon \left[\ \left[\ z\ \right] \frown \left[\ \begin{array}{c} vii \\ v \end{array}\ \right]_2 \frown \left[\ z\ \right] \frown \left[\ \begin{array}{c} v \\ iii \end{array}\ \right]_2\ \right]_2 \\ e_{bass}\colon \left[\ \left[\ i\ \right] \frown \left[\ z\ \right]_2 \frown \left[\ v\ \right] \frown \left[\ z\ \right]_2\ \right]_2 \end{array}\ \right]\ \right]$$

How the cello plays *C minor* 7 in two parts is shown as the top event in equation 3.3. The first part is the last two notes of this four-note chord, and the second part consists of the middle two. This means firstly that there is an overlap since the fifth degree is played both times, but more importantly that the second part of the chord need not be interpreted as a 7^{th} chord, i.e. as a tetrachord, but may be viewed as a triad instead, in other words as an 'ordinary' *C minor*. This has consequences for the denotation, since a *C minor* is more stable than a *C minor* 7.

Equation 3.4 shows a possible denotation for the music of figure 2. Informally it can be described as a swan being seen to be flying above the sea with waves moving, in the direction of an observer, and concluding its movement with a glide. That it approaches the observer is derived from the descending violin notes but the inference is aided by their harmonic function (viz. equation 3.2). The descending notes imply that the source (i.e. the swan) becomes larger (cf. [Sch19], section 4.5, page 52), while their harmonic roles, i.e. the movement from v down to the tonal centre i, suggest completion (*ibid*, section 5.3, page 59), with the source 'arriving home' so to speak. The descent to the stable harmonic function i moreover implies a situation of greater stability in the denotational world (section 1.2, page 38) – hence the final glide.

The three voices of the piece, i.e. the violin plus the bass as well as the chord voice of the cello, each denote a different aspect of this situation. As indicated, the violin captures the swan's motion and the way it moves through the air; this is the top event e_{swan} in equation 3.4. The two voices played by the cello capture both the animal's wing motion and the movement of the waves, given as the bottom events.

Equation 3.4.

$$e\colon \left[\ \begin{array}{l} e_{swan}\colon \left[\ \begin{array}{l} e_{motion}\colon \left[\ towards\ \right]_{12} \\ e_{mode}\colon \left[\ flying\ \right]_{10} \frown \left[\ gliding\ \right]_2 \end{array}\ \right] \\ e_{swan}\colon \left[\ e_{wings}\colon \left[\ up\ \right] \frown \left[\ still\ \right]_2 \frown \left[\ down\ \right] \frown \left[\ still\ \right]_2\ \right]_2 \\ e_{sea}\colon \left[\ e_{waves}\colon \left[\ \left[\ still\ \right] \frown \left[\ up\ \right] \frown \left[\ down\ \right]\ \right]_4\ \right] \end{array}\ \right]$$

While in equation 3.4 events are separated to reflect the instrument's voices via which they are mediated, in equation 3.5 the swan wings event has instead been moved up into the main e_{swan} event.

Equation 3.5.

$$e\!:\left[\begin{array}{l} e_{swan}\!:\left[\begin{array}{l} e_{motion}\!:[\ towards\ ]_{12} \\ e_{mode}\!:[\ flying\ ]_{10} \frown [\ gliding\ ]_{2} \\ e_{wings}\!:[\ [\ up\ ] \frown [\ still\ ]_2 \frown [\ down\ ] \frown [\ still\ ]_2\ ]_2 \end{array}\right] \\ e_{sea}\!:[\ e_{waves}\!:[\ [\ still\ ] \frown [\ up\ ] \frown [\ down\ ]\ ]_4\ ] \end{array}\right]$$

3.1.1 Minimal pairs

As indicated in sections 2.1.1 and 2.1.2.1, Schlenker proposes the construction of minimal pairs to demonstrate the presence of a particular semantic effect in the denotation of a musical piece. The idea is to minimally alter the music such that the infered effect disappears. As a minimal pair for the music of figure 2 and its denotation in equation 3.5, consider figure 4.[26]

Figure 4: Bars 5 to 8 (first note throughout) from Shostakovich's *Waltz No. 2*

Here, the only difference is that while the violin plays the same rhythm, it does not descend from its initial note. This neutralises the semantic effect of the approach towards an observer of the violin voice's virtual source since there are no longer any lower notes to suggest an increase in size of the source, and hence no suggestion of increasing closeness to the observer. Moreover, the source no longer attains a greater degree of stability – and hence the inference of gliding is no longer valid – as it does not move towards the tonal centre (i.e. to the harmonic degree i).

Pairs for other aspects of the denotation have been omitted, but might be constructed by having the cello repeat its first bar, i.e. play $\left[\ [\ C,\ ] \frown \left[\begin{array}{c} B\flat, \\ G, \end{array}\right]\ \right]_4$. This would neutralise both the wave motion and the wings flapping.

3.1.2 Tetrachordal instability

The rationale behind the wings flapping is the repeated bass pattern of the cello that moves from i to v (descendingly) in the bottom event of equation 3.3. The top event of that equation, representing the cello's chords, denotes the wave's motion due to the progression of the less stable *C minor* 7 chord (wave up) to the more stable ordinary *C minor* (wave down).

As indicated in section 2.1.3, definition 2.3, a less stable harmonic event corresponds to a less stable denotational situation in the world according to Schlenker. However, section 5.3 of [Sch19] only gives an ordering $IV < V < I$ for the stability of chords, while Lerdahl on which Schlenker bases his work does not include the level of 7^{th} or tetrachords ([Ler01], chapter 2, page 47). Hence the reason

[26] `https://deneeve.github.io/ac/edu/sgr/sho.waltz2.note1.5-8.mp3`

why *C minor* 7 is less stable than *C minor* must be stated explicitly here, since in the context of the minor harmony of *Waltz No. 2*, *C minor* 7 would in fact be the maximally stable chord – assuming an analogous ordering $iv < v < i$ for minor harmony and noting that *C minor* 7 would then have the role of chord degree i as well as the fact that no distinction between a three-note triad such as *C minor* and a four-note tetrachord like *C minor* 7 follows from either Schlenker or Lerdahl.

The justification is as follows. The difference between *C minor* and *C minor* 7 is the presence of the extra note $b\flat$ in the latter. This makes a progression to a $B\flat$ chord plausible – in fact to $B\flat$ 7 – via two semitone transitions, namely from $e\flat$ down to d and from g up to $a\flat$. This progression is more plausible than in the case of the *C minor* triad since that lacks the $b\flat$ note, and moreover $B\flat$ 7 is unstable as it is itself attracted to *E♭ major* (the major parallel of *C minor*) via reverse semitone transitions (d up to $e\flat$ and $a\flat$ down to g). So *C minor* 7 is more likely than *C minor* to transition to an unstable chord which then takes it out of its tonal centre (which could have been assumed to be stable if only it had been given by a triad rather than a tetrachord). Hence *C minor* 7 must be less stable than *C minor*.

So then the rationale for the waves going up and down becomes twofold: it is not only due to the motion between less and more stable harmony, but also to the motion of chord notes moving downwards – i.e. from vii to v as well as from v to iii (viz. equation 3.3).

3.1.3 A transformed musical denotation

Figure 5 depicts bars 13 to 16 of *Waltz No. 2*.[27] It is basically a transformation, or development, of bars 5 to 8 of figure 2. Consequently, it is plausible to consider a denotation for it that is essentially a (simple) transformation of the one given in equation 3.5.

Figure 5: Bars 13 to 16 (last note extended) from Shostakovich's *Waltz No. 2*

Equation 3.6 gives the sound events for the violin and the cello, and equation 3.7 the harmony – split into the chordal and bass voices of the cello.

Equation 3.6.

$$e_{pitch}\colon \begin{bmatrix} e_{violin}\colon [\ F\ ]_3 \frown [\ D\ ]_2 \frown [\ C\ ] \frown [\ B_,\ ]_6 \\ e_{cello}\colon \begin{bmatrix} [\ G,,\ ] \frown \begin{bmatrix} B_, \\ E_, \end{bmatrix}_2 \frown [\ D_{,,}\ ] \frown \begin{bmatrix} F_, \\ D_, \end{bmatrix}_2 \end{bmatrix}_2 \end{bmatrix}$$

[27] https://deneeve.github.io/ac/edu/sgr/sho.waltz2.13-16.mp3

Equation 3.7.

$$e_{harmony}\colon \begin{bmatrix} e_{violin}\colon \begin{bmatrix} IV \end{bmatrix}_3 \frown \begin{bmatrix} II \end{bmatrix}_2 \frown \begin{bmatrix} I \end{bmatrix} \frown \begin{bmatrix} VII \end{bmatrix}_6 \\ e_{cello}\colon \begin{bmatrix} e_{chord}\colon \begin{bmatrix} \begin{bmatrix} z \end{bmatrix} \frown \begin{bmatrix} IV \\ II \end{bmatrix}_2 \frown \begin{bmatrix} z \end{bmatrix} \frown \begin{bmatrix} II \\ VII \end{bmatrix}_2 \end{bmatrix}_2 \\ e_{bass}\colon \begin{bmatrix} \begin{bmatrix} V \end{bmatrix} \frown \begin{bmatrix} z \end{bmatrix}_2 \frown \begin{bmatrix} II \end{bmatrix} \frown \begin{bmatrix} z \end{bmatrix}_2 \end{bmatrix}_2 \end{bmatrix} \end{bmatrix}$$

This time, the harmonic Roman numerals are given in uppercase, because the harmonic function has shifted from a minor first degree (i) to a major fifth degree (V) (in equation 3.2 as well as 3.7, it is the cello's first bass voice note that indicates this). Again, the violin descends, this time passing the first degree and going one note lower to VII. This keeps the basic denotational motion geared towards the oberver. However, note that the VII it now ends on is unstable, in comparison with the i of equation 3.2. At the same time, the chordal motion of the cello runs – like the violin – in parallel to that of equation 3.2. It has the IV on top which descends to II, and the II below that descends to VII – again passing through degree I.

But note that IV the cello plays as its top chord note has been given from the perspective of the base harmonic function I, like the V that the cello plays as its bass note. If the V is considered as root, i.e. as a temporary I, then from that perspective the IV is actually VII, i.e. the cello's top chord note is again a 7^{th} chord and it is again unstable, as in equation 3.2. Similarly, the II it now plays as the first bottom chord note has the function V when considered from this temporary I perspective (which is also true for the cello's second bass note II).

Then since the VII is also equivalent to III from the first degree perspective, the harmonic functionality of the cello is completely analogous to the one in equation 3.2. However, for the violin this is only true for the aspect of descent, as it now starts on IV, which is VII from the first degree perspective, while in equation 3.2 it starts on v.

Equation 3.8 gives a denotation that is essentially a transformation of the denotation given in equation 3.5.

Equation 3.8.

$$e\colon \begin{bmatrix} e_{sun}\colon \begin{bmatrix} e_{cloud}\colon \begin{bmatrix} \mathit{obscuring} \end{bmatrix}_{12} \end{bmatrix} \\ e_{swan}\colon \begin{bmatrix} e_{motion}\colon \begin{bmatrix} \mathit{towards} \end{bmatrix}_{12} \\ e_{mode}\colon \begin{bmatrix} \mathit{flying} \end{bmatrix}_{12} \\ e_{wings}\colon \begin{bmatrix} \begin{bmatrix} \mathit{up} \end{bmatrix} \frown \begin{bmatrix} \mathit{still} \end{bmatrix}_2 \frown \begin{bmatrix} \mathit{down} \end{bmatrix} \frown \begin{bmatrix} \mathit{still} \end{bmatrix}_2 \end{bmatrix}_2 \end{bmatrix} \\ e_{sea}\colon \begin{bmatrix} e_{waves}\colon \begin{bmatrix} \begin{bmatrix} \mathit{still} \end{bmatrix} \frown \begin{bmatrix} \mathit{up} \end{bmatrix} \frown \begin{bmatrix} \mathit{down} \end{bmatrix} \end{bmatrix}_4 \end{bmatrix} \end{bmatrix}$$

The transformed denotation differs from the one in equation 3.5 in two respects. First of all the switch to the degree V harmony with a dominant 7^{th} chord makes the whole situation less stable. This is represented by the addition of $e_{sun}\colon \begin{bmatrix} e_{cloud}\colon \begin{bmatrix} \mathit{obscuring} \end{bmatrix}_{12} \end{bmatrix}$.[28] Secondly, e_{mode} has been altered to be *flying* throughout, rather than reverting to *gliding* for the final two measures.

[28] So it is assumed that $e_{sun}\colon \begin{bmatrix} e_{cloud}\colon \begin{bmatrix} \mathit{clear} \end{bmatrix}_{12} \end{bmatrix}$ is part of the denotation of figure 2.

This is because *gliding* is deemed a more stable position, and this can now no longer be justified to be the case. As noted, the violin now ends on the harmonic function of VII – which is III from the aforementioned temporary first degree perspective. In other words, it does not resolve to the tonic as it does in equation 3.2, rather, it ends a half-note below this and is therefore attracted to the tonic. Hence the segment ends in an unstable position, and therefore the denotation cannot justifiably end (relatively) stable.

3.2 Entailments, equivalences, and falsities

As indicated in section 2.3, event-based semantics help cater for entailments between more specific and more general statements. As a basic example, consider figure 6. The music of this figure[29] can be given the denotation in equation 3.9.

Figure 6: Bars 5 to 8 (violin only) from Shostakovich's *Waltz No. 2*

Equation 3.9.

$$e_{object}\colon \left[\ e_{motion}\colon \left[\ \textit{towards}\ \right]_{12}\ \right]$$

Setting equation 3.5 to denotation $d_{(violin,cello)}$, and equation 3.9 to $d_{(violin)}$, it then holds that $d_{(violin,cello)} \models d_{(violin)}$. So this feature of event semantics arguably caries over to the musical semantics described here, since the violin and cello events are essentially concurrent conjunctions, and e_{swan} is an instantiation of e_{object}.

Note that the situation of equation 3.9 is more underspecified than the one in equation 3.5. And as Schlenker points out in [Sch19], section 6.2, page 69, musical inferences are generally underspecified, so the range of possible denotations of any given piece of music will likely be very diverse. So as an equivalent alternative denotation for the music of figure 2, consider equation 3.10.

Equation 3.10.

$$e\colon \begin{bmatrix} e_{boat}\colon \begin{bmatrix} e_{motion}\colon [\ \textit{towards}\ ]_{12} \\ e_{mode}\colon [\ \textit{rocky}\ ]_{10} \frown [\ \textit{steady}\ ]_{2} \end{bmatrix} \\ e_{sea}\colon \begin{bmatrix} e_{waves}\colon [\ [\ \textit{still}\ ] \frown [\ \textit{up}\ ] \frown [\ \textit{down}\ ]\ ]_{4} \\ e_{swell}\colon [\ [\ \textit{up}\ ] \frown [\ \textit{still}\ ]_{2} \frown [\ \textit{down}\ ] \frown [\ \textit{still}\ ]_{2}\ ]_{2} \end{bmatrix} \end{bmatrix}$$

Here, the e_{object} of equation 3.9 has been instantiated to e_{boat} rather than e_{swan}. Moreover, the action of the swan's wings that are mediated by the cello's bass notes in equation 3.5 have been re-interpreted as the general swell of the sea. In other words, this property of the denotational situation is now part of e_{sea} instead of e_{object} (and the cello now fully mediates e_{sea}).

Again, an entailment can be made. Consider figure 7. The music in this figure[30] may be given the denotation in equation 3.11.

[29] https://deneeve.github.io/ac/edu/sgr/sho.waltz2.vn.5-8.mp3
[30] https://deneeve.github.io/ac/edu/sgr/sho.waltz2.co.5-8.mp3

Figure 7: Bars 5 to 8 (cello only) from Shostakovich's *Waltz No. 2*

Equation 3.11.

$$e_{sea}\colon \left[\begin{array}{l} e_{waves}\colon \left[\; [\; still \;] \frown [\; up \;] \frown [\; down \;] \; \right]_4 \\ e_{swell}\colon \left[\; [\; up \;] \frown [\; still \;]_2 \frown [\; down \;] \frown [\; still \;]_2 \; \right]_2 \end{array} \right]$$

Setting equation 3.10 to $d'_{(violin,cello)}$, and equation 3.11 to $d'_{(cello)}$, it then also holds that $d'_{(violin,cello)} \models d'_{(cello)}$.

Finally, although a diverse range of situations can be true of a given piece of music, there are limitations, and hence situations that are not true, as indicated in section 2.1.3 of this paper (notably equation 2.4). For instance, consider equation 3.12.

Equation 3.12.

$$e\colon \left[\begin{array}{l} e_{plane}\colon \left[\begin{array}{l} e_{motion}\colon [\; towards \;]_{12} \\ e_{mode}\colon [\; flying \;]_{10} \frown [\; splashdown \;]_2 \end{array} \right] \\ e_{sea}\colon \left[\begin{array}{l} e_{waves}\colon \left[\; [\; still \;] \frown [\; up \;] \frown [\; down \;] \; \right]_4 \\ e_{swell}\colon \left[\; [\; up \;] \frown [\; still \;]_2 \frown [\; down \;] \frown [\; still \;]_2 \; \right]_2 \end{array} \right] \end{array} \right]$$

The situation specified here has been altered from the one in equation 3.10 (i.e. the one with the boat), and e_{object} is instantiated to e_{plane}. Other than that, the only change is that $e_{mode}\colon [\; steady \;]_2$ has been altered to $e_{mode}\colon [\; splashdown \;]_2$. But this inference from the last two violin is unlicensed. Note first of all that $e_{boat}\colon [\; e_{mode}\colon [\; steady \;]_2 \;]$ in equation 3.10 is an equivalent alteration of $e_{swan}\colon [\; e_{mode}\colon [\; gliding \;]_2 \;]$ in equation 3.5.

As indicated in section 3.1 (the explanation preceding equation 3.4), the glide is infered from figure 2's violin completing a movement towards the stable harmonic function i, and the same goes for the boat being steady. So a plane making a splashdown is incoherent, since that is a situation which is less stable. This violates the harmonic stability condition (see definition 2.3), so the denotation in equation 3.12 cannot be true of the music in figure 2.

3.3 Informal denotation of *Waltz No. 2*'s opening bars

Given the formal analyses amounting to the event denotations for measures 5-8 and 13-16 of *Waltz No. 2* in equations 3.5 and 3.8, a possible denotation for the first 20 measures of the piece may now be informally sketched as follows.

In the first four bars, only the cello is heard,[31] with the harmony of equation 3.3, and the situation given in the bottom two events of equation 3.4, i.e. the bottom e_{swan} event and the e_{sea} event. This can be interpreted as a swan being seen

[31] See the top of figure 3, and the beginning of `https://deneeve.github.io/ac/edu/sgr/sho.waltz2.mp3`.

flying over a wavy sea, without it yet being clear in what direction the animal is moving, however, its wings flapping is apparent to the observer. Then as the violin enters from measure 5, the denotation of equation 3.5 holds, as it becomes apparent that the bird is moving in the direction of the observer.

Measures 9 to 12 have not been formalised, but these may be interpreted in terms of how the swan is subsequently observed to be moving. See figure 8.[32]

Figure 8: Bars 9 to 12 (plus pick-up) from Shostakovich's *Waltz No. 2*

As pointed out in section 3.1, a lower pitch may imply the source becoming larger, i.e. getting nearer, but conversely, a higher pitch can also mean the source having more energy ([Sch19], section 4.8). So the upward pitch pattern in these measures can be interpreted, for instance, as the animal being thrust forward by a gust of wind.

Similarly, consider measures 17 to 20 in figure 8.[33]

Figure 9: Bars 17 to 20 (plus pick-up) from Shostakovich's *Waltz No. 2*

As indicated in equation 3.8, the preceding four measures have the added denotation of a cloud obscuring the sun. This can be assumed to continue for measures 17 to 19, as the harmonic pattern that justifies it remains the same. Since the piece returns to the tonic in bar 20, the sunshine returns at that point. The similar violin pitch height pattern may then again denote a forward-thrusting gust of wind. But the penultimate note $F\natural$ is dissonant, and indicates a situation of instability ([Sch19], section 1.2, page 38). This can be interpreted as the bird being in a momentary downward sliding position. As the final G (which is also the same as the first note) is (relatively) stable again, the swan regains its posture at this point.

So in all, the first twenty measures of *Waltz No. 2* can denote the following situation. A swan is observed to be flying over a wavy sea while the sun is out. Initially, it is only seen to be moving but not in which direction, until it becomes apparent that the bird is flying in the direction of the observer. The swan flies actively by flapping its wings but then briefly rests them in order to glide. Some gusts of wind subsequently thrust the animal forward. While the

[32] https://deneeve.github.io/ac/edu/sgr/sho.waltz2.9-12.mp3
[33] https://deneeve.github.io/ac/edu/sgr/sho.waltz2.17-20.mp3

bird continues to approach the observer in the same alternation of flight and glide, a cloud obscures the sun. Some more gusts of wind then again thrust the swan forward, and in the end, it slides sideways for a short moment before it recovers its balance.

4 Discussion

The semantics in the preceding section is by no means exhaustive. The aim has been to demonstrate that a more fine-grained representation of (coordinated) musical events following Cooper in [Coo13] can be used to write out denotations in the world as Schlenker does in [Sch19], and make such denotations more detailed in the process – and moreover, to make the case that entailments from more specific to more general events hold as they do in event semantics for language. But obviously some central features of a semantic theory are lacking, notably a definition that states precisely how the constituents of the musical language are to be interpreted.

According to Schlenker, this should come as no surprise. In *ibid*, section 7.1, page 71, the author states that musical form, i.e. syntax, does not by itself determine what events a piece of music is true of. Furthermore, he points out that if it did, the result would be a conjunctive logic in which the order of conjuncts (of musical events) should not matter to the meaning (the denotation in terms of world events). However, in musical semantics order does matter, a point to be revisited further down in section 4.2.

4.1 Formalising transitions

In this paper, the semantics has been specified by a series of transformations from musical pitch events, via harmonic events, to events in the world. For instance, equation 3.1, as a pitch event rendition of figure 2 (bars 5 to 8), was transformed into the harmonic events in equation 3.2, and this in turn into the world events of equation 3.5: the situation of the swan approaching over the sea. The transformation for bars 13 to 16 (figure 5) is perhaps somewhat clearer, i.e. from equation 3.6 (the pithes) via equation 3.7 (the harmony with the cello already split into its chords and bass voices) to equation 3.8 (the situation of the swan approaching with the sun obscured). But how these transformations have been arrived at has been left largely unspecified, except that justifications for various aspects of the transformations have been given from the point of view of music theory in section 3.1.

It should be possible to state rules for the transformation from pitch to harmony events precisely – and also the way pitch events are derived from the score, since the score fully determines pitch events. In the rationales for the transformations given in the text of the aforementioned section, references are made to justifications from auditory cognition as well as from pitch space, and while it may be hard to state these with exhaustive precision in a set of definitions, it seems feasible to state them more accurately than has been done here.

If so, it may be possible to prove the truth of the situation represented by equation 3.5 in section 3.1, or the falsity of the one given by equation 3.12 in section 3.2, with reference to a truth definition. This would likely be close to definition 2.3 (section 2.1.3), but it would have to be extended to coordinated events rather than a series (i.e. as stated in definition 2.2). A further requirement might be an ontology of entities, states, and actions, such that it can be made clear that in terms of relative stability, the difference between $flying$ and $gliding$ from equation 3.5 can be shown to be on a par with *rocky* and *steady* from equation 3.10, and a clause such as *(c)* in definition 2.3 can be invoked for both situations. But clearly, such efforts would have been outside the scope of the current paper, and would moreover not change the basic ideas underpinning Schlenker's musical semantics.

4.2 Dynamic interpretation

As indicated in section 2.1.3, music may be considered as a concatenation of musical events mirroring a concatenation of world events or denotations which the music is true of. As equation 2.4 attests, and assuming that the *luminosity* event can be ignored, Strauss' *Zarathustra* can be said to be true of $p_1 = \textit{min-luminosity} \wedge p_2 = \textit{max-luminosity}$, but not of the reverse $p_2 \wedge p_1$. In Schlenker's approach, a sequence of musical events is a dynamic representation of world events ([Sch19], page 72), and to him, musical semantics differs from logical semantics because in the latter, the order in which conjuncts are combined does not matter for the meaning of the result. But in dynamic semantics, the order in which conjuncts are interpreted does affect the resulting meaning, rendering conjunction non-commutative (cf. [GS88], section 3, page 11).

So the difference between musical and logical semantics may not be as big as Schlenker states. This is significant, since the author's notion of meaning is grounded in the idea of transforming a dynamic representations of musical events into some resulting equally dynamic representations of world events. But the resulting representation, the denotation, is not arrived at dynamically, because the entire musical piece is assumed to be available for semantic analysis, and the question is whether this leads to a psychologically plausible model of musical interpretation, since in reality, listening is a dynamic process and a piece of music reveals itself to the listener in a piecewise fashion.

Schlenker acknowledges this implicitly by refering to so-called deceptive cadences (*ibid*, section 5.3; see also section 2.1.2.1 here). After all, if a cadential movement can be deceptive, then it will appear to be going in a certain direction first, with the composer intentionally setting up a particular expectation in the listener, only to subsequently deviate from that direction in order to end up elsewhere. From a dynamic perspective, the deceptive cadence works because the composer or musician has created a context or state s, by getting the listener to first interpret harmonic mode II and then V, after which the listener interprets mode IV given s. And obviously this would not work with the reverse sequence $[IV, V, II]$. Incidentally the same goes for the perfect cadence $[II, V, I]$, or perhaps more clearly for its final movement $[V, I]$, where it is evident that the reverse $[I, V]$ has a different meaning. After all, the harmonic mode V represents an unstable state, which will plausibly be represented in the listener in some way. Given that the subsequent resolution (mode I) is interpreted against that listening state, the meaning of the segment relates to some notion of 'return', while the meaning of the reverse sequence $[I, V]$ is associated with the (opposing) idea of 'moving away'.

Some inspiration for what such a listening state could be like may be found in the paper *Changing the Context: Dynamics and Discourse* by Jeroen Groenendijk and Martin Stokhof [GS96]. It introduces the field of dynamic semantics, in which the meaning of an utterance is taken to be its ability or potential to change the context of a discourse, and this context is in turn equated with the information that is available to an interpreter of the discourse, which is equated with an information state that represents possible states of affairs at a given point in the discourse. In Groenendijk and Stokhof, information states are sets of possible worlds that contain possible values of items which have been introduced into the discourse (see sections 3.2 and 3.4 for some examples). In short, when an utterance is interpreted, the interpretation depends on the information – i.e. the possibilities – so far, but possibilities may also be added or eliminated. In other words, interpretation is viewed not only as a context-dependent but also as a context-creating process.

Following [GS96], section 1.2, a musical context would be a a listening or information state that represents the set of (world) events that may be considered as possible denotations for the musical segment under consideration. But note that even for a small segment like measures 5 to 8 of *Waltz No. 2*, there is already a large set of possibilities, including the denotations given in equations 3.5 and 3.10. The idea is that these would be among the remainders after the interpretation of these bars concludes and items have been eliminated from an even larger set of contenders. But to be fair, Groenendijk and Stokhof contend that elimination from a larger set is just one way of modeling partiality of information, which has been chosen because it is considered to be technically simpler. An alternative would be to model partiality in terms of partial objects, worlds, or situations (*ibid*, section 3.1, footnote 14). This appears closer to Schlenker's methodology, since various musical events may be considered as incrementally providing information about certain corresponding aspects of world events, in other words, it seems more apt to view denotations as being constructed piecewise rather than arriving at them by having possible candidates eliminated.

In any case, a dynamic interpretation approach would change the core notion of musical meaning. For one, the semantics would no longer be restricted to a given musical piece or segment, but would apply to some appropriate breakdown of its constituent parts. The meaning of a given constituent would then not be its denotation (i.e. situation in the world) or set of possible denotations, but its ability or potential to change the listening (information) state, in other words, the context with respect to which what comes next in the music is to be interpreted – and which will then subsequently further alter (update) the state. So the meaning of a musical utterance becomes a function from listening states to listening states (which can then in turn represent such denotations).

As an example of how this might (broadly) work, consider Strauss' *Zarathustra* again. Its use by Schlenker has been discussed here in section 2.1, but for now, consider [Sch19], section 3.2, page 44, example 6, which associates the music's opening bars with the sunrise from behind a planet imagery from the opening scene from Kubrick's *2001: A Space Odyssey*, and aligns the appropriate images with the music. For clarity, this example is shown in figure 10.[34]

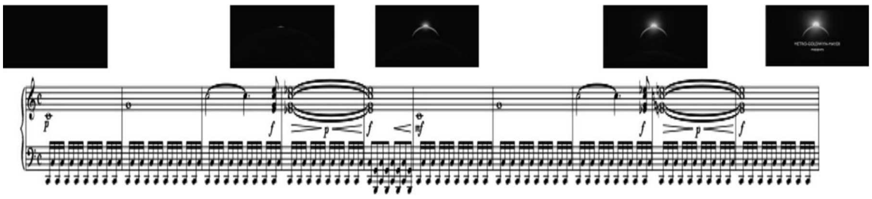

Figure 10: Opening bars from Strauss' *Also Sprach Zarathustra* annotated with imagery from Kubrick's *2001: A Space Odyssey* (example 6 in [Sch19])

The initial listening state is essentially empty, and indeed, the annotated image in figure 10 here appropriately shows an empty space. An assumption that could be made is that the listening state is updated after the completion of a musical idea. The first musical idea consists of the initial three-note sequence $[C, G, c]$.[35] Because it is a crescendo, this sets up the existence of a virtual source which (gradually) increases in energy. So if $s_0 = \emptyset$, then arguably $s_1 = \{\exists x : E(x, t_0) < E(x, t_1)\}$. But alternatively, the state might already be updated at the first note, such that it simply holds that $s_1 = \{\exists x\}$ and at the second note, $s_2 = \{\exists x : E(x, t_0) < E(x, t_1)\}$. Now the meaning of G can be specified as a function from s_1 to s_2: a function that adds energy to x.

The movie shows the planet appearing at the end of the sounding of the third note c, and just before the brief e. But for simplicity, assume the C *minor* chord as the point of the next update (so ignoring the e). The $e\flat$ in C *minor* provides the information that the music is in a minor mode, but more importantly it gives it a (local) tonal centre consisting of the C *minor* triad. This is Schlenker's event $\langle I, 80dB\rangle$ in his equation 23 a (equation 2.1 here). The preceding events being $\langle I, 70dB\rangle$ and $\langle V, 75dB\rangle$, a harmonic movement $[I, V, I]$ in pitch space is now established.[36] This effectively licenses the inference of a movement in physical space, so that a further state can be specified: $s_3 = \{\exists x : E(x, t_0) < E(x, t_1) \wedge loc(x, t_0) \neq loc(x, t_1)\}$.[37] So then the meaning of the C *minor* triad can be stated as a function from s_2 to s_3; one that effectively moves x from some location to another.

In this way, state transitions can be made to encode situations that become increasingly more specific, i.e. partiality of information can be encoded via ad-

[34] In his paper, Schlenker refers to `http://bit.ly/2FH39Ps` for a clip with the score and the music, and to `http://bit.ly/2DfiE3m` for the music plus the movie imagery. However, figure 10 is an annotation of a piano rendition given here: `https://www.8notes.com/scores/7213.asp` - but in fact the semantics are better explained on the basis of the second link of the actual movie opening sequence, i.e. without the piano rhythm, but with the trumpets plus the chords instead.

[35] In ABC notation, see footnote 25.

[36] Strictly speaking this should be notated $[i, v, i]$ as a minor mode has also been established, but Schlenker's notational choice is adhered to for this example.

[37] That x's location at t_1 is different from the one at t_0 is further derived from the increase in energy due to the crescendo (i.e. from auditory cognition).

dition rather than elimination (as in [GS96]). One idea for future work is to use logical formulas such as given above for the states $s_0, \ldots, s_3$ to arrive at denotations for musical pieces (such as in equation 3.5) and, more importantly, constituent musical expressions. This allows one to pinpoint musical meanings at smaller units – even to individual notes as illustrated above, but can also provide more clarity as to what it means for musical semantics to be *"under-specified and abstract"* ([Sch19], section 6.2, page 67). The idea is that while a piece of music unfolds, very general denotations may be gradually filled in until a more specific one can be stated.

So for the example at hand, and given the denotational direction suggested by the motion picture annotations in figure 10, the specific denotation to be eventually arived at here is obviously the *Sun-rise* one from equation 2.4 here (see also footnote 17). This would then have to be obtained by further updating $s_3 = \{\exists x : E(x, t_0) < E(x, t_1) \wedge loc(x, t_0) \neq loc(x, t_1)\}$, plausibly due to the briefly played note e plus the proctractedly sounding subsequent C *minor* signifying some dramatic turn for the virtual source x such that it can be interpreted as a *'shining'* event featuring x, thereby further specifying what is going on with x. For one, it would have to indicate that $loc(x, t_0) \neq loc(x, t_1)$ is actually an event in which x rises, or at least moves in such a way so as to cease to be obscured. How this might be achieved is outside the scope of the current paper, though it may as indictated feature in future work.

4.3 Compositionality

Schlenker states the aim is to develop *"a source-based semantics rather than a compositional semantics"* ([Sch19], section 1.2, page 39). And meanings being inferences (see section 2.1 here), frequently from a sequence of notes at once rather than on a note-for-note basis, it seems clear that his semantics as such is indeed non-compositional. To be sure, it is difficult to see how a note in isolation could have meaning as the primary units in music are arguably not notes but intervals, or perhaps even chords since it is these that provide (pitch space) contexts against which individual notes may obtain the sort of inferential meanings of Schlenker.

However, in the preceding section, a dynamic interpretation analysis was sketched for the begining of *Zarathustra* where actually, the first note was given a meaning, albeit a very basic one: a function $s_0 = \emptyset \to s_1 = \{\exists x\}$, in other words, the introduction of some object.[38] To obtain compositionality, subsequent notes would then be interpreted given the listening state so far, as is done for *Zarathustra*'s opening notes in section 4.2, but note that to arrive at s_3, reference needs to be made to the harmonic function of the notes, but as Schlenker points out, the mode (major or minor) is *"initially underspecified"* ([Sch19], section 6.2, page 65). In fact, even the harmonic root – which turns out to be C – is unknown as the first note sounds.

So it seems inevitable that inferences have to be made from 'idiomatic', i.e. non-compositional groups of notes that constitute the chords or harmony. However,

[38] That there is an event in which x is required to participate (see definition 2.2) has been left implicit in this example. So made explicit, one might instead stipulate that $s1 = \{\exists o \exists e\}$.

this need not form a principled objection to the idea of a compositional musical semantics, since it may be possible for a dynamic approach to also model the harmonic uncertainty, and incrementally resolve this as the music proceeds. This would intuitively involve a possible worlds approach as in [GS96].

By way of illustration, consider figure 11. It shows two instances of the first two bars of *Zarathustra*, plus a third measure, where in the case depicted on the left the harmony moves to *C minor*, as it actually does, and in the right-hand case it settles into *A minor* instead.[39]

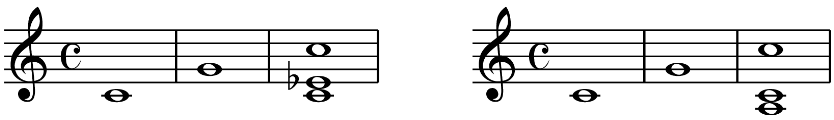

Figure 11: *Zarathustra* opening moving to *C minor* (left), and to *A minor* (right)

So in the present case, unlike in the previous section, the dynamic interpretation process does not serve to model a situation, or the development of a partial denotation, but models the harmonic possibilities that might obtain as subsequent notes are interpreted. For simplicity, assume it is required that the notes in figure 11 be the chord notes root, 3$^{\text{rd}}$, 5$^{\text{th}}$, or 7$^{\text{th}}$. Here, let $s_0 = \mathcal{U}$, i.e not $\emptyset$ but a universal set with respect to the domain at hand, which consists of worlds associated with all major and minor (triad) chords. After all, before the music has started, everything is still possible, so $\mathcal{U}$ represents a state of ignorance.

As the first note sounds, i.e. the low c (in the left as well as in the right-hand case), several worlds remain, so subsequently $s_1 = \{C\ major, C\ minor, A\flat\ major, A\ minor, F\ major, F\ minor, D\flat\ major, D\ minor\}$. All these chords contain c, which for the first two worlds is root, in world three and four it is 3$^{\text{rd}}$, in world five and six 5$^{\text{th}}$, and in world seven and eight it has the role of the 7$^{\text{th}}$.

The second note is g in both cases. Interpreting this eliminates several more worlds (i.e. chords), namely *F major* and *F minor*, since g has the role of the 4$^{\text{th}}$ there, and furthermore *D♭ major* as this would have g feature as a ♭5, and finally *D minor*, since there, g is also 4$^{\text{th}}$. So these are eliminated by virtue of not being the required chord notes, and consequently $s_2 = \{C\ major, C\ minor, A\flat\ major, A\ minor\}$. Now, just like in section 4.2, the meaning of the note g is given by a function $s_1 \to s_2$, except now it is not a denotational (or inferential) meaning, but a harmonic one, as it merely eliminates the subset $\{F\ major,\ F\ minor, D\flat\ major, D\ minor\}$ from s_1, i.e. the worlds in which g is a non-chord note.

The situation for the third bar of the right side of figure 11 is quite straightforward. It contains a c plus two instances of an a. Because an a has the role of a 6$^{\text{th}}$ in *C major* as well as *C minor*, these worlds are eliminated from s_2.

[39] In the third bars of the music in figure 11, unlike in the original, a semibreve (whole note) has been chosen for the high c (plus the resolving chord notes below it). For piano sound renditions, see `https://deneeve.github.io/ac/edu/sgr/str.zara.cmin.mp3` and `https://deneeve.github.io/ac/edu/sgr/str.zara.amin.mp3`, respectively.

$A\flat$ *major* is also removed, since despite the c featuring as a 3[rd], the a is a $\sharp 1$. Because in *A minor*, c is 3[rd] while a is root, this world remains and hence $s_3 = \{A\ minor\}$, and the function $s_2 \to s_3$ by which the listener is left with *A minor* is therefore the harmonic meaning of the notes in this measure.

Bar three in the left-hand side of figure 11 contains two instances of a c plus an $e\flat$, and here things are somewhat more complex. The *C major* world is removed because $e\flat$ has the role of a $\flat 3$, while *A minor* gets eliminated as the same note would be a $\flat 5$. The *C minor* world is retained since (obviously) c is root there and $e\flat$ is 3[rd], but $A\flat$ *major* also remains because c plays the role of 3[rd] there while $e\flat$ features as 5[th]. So at the left of the figure, the meaning of the third-measure notes constitues the function $s_2 \to s_3$ which in this case renders $s_3 = \{C\ minor, A\flat\ major\}$, i.e. the music actually remains harmonically ambiguous at this point.

But figure 12 below shows that this in fact makes sense, as it is perfectly possible for the notes in the third measure to resolve to $A\flat$ *major* in the next bar.[40] In bar four, the low c of the preceding measure is replaced by an $a\flat$, while the $e\flat$ and high c persist. Since it was already clear that the latter two notes fulfill the stated chord-role requirements for notes in the $A\flat$ *major* world, it only needs to be checked whether the addition of the $a\flat$ note satisfies *C minor*. This is not the case, as in the context of *C minor*, $a\flat$ is the 6[th], and therefore the interpretation of bar four eliminates the *C minor* world, so that $s_4 = \{A\flat\ major\}$.

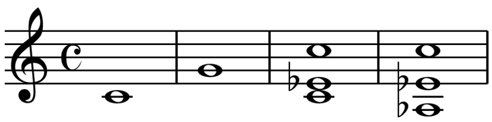

Figure 12: *Zarathustra* opening resolving to $A\flat$ *major*

The above implies that at least in certain cases covered by Schlenker in [Sch19], it is possible to avoid the idiomacy of chords or harmonic functionality spawned by note sequences, by interpreting them dynamically and compositionally, so that musical features that play a central role in Schlenker's semantics may become an integral part of the interpretation process, instead of having to stipulate harmonic or pitch space matters beforehand. To be sure, harmony belongs to the realm of what Schlenker terms 'internal semantics' (*ibid*, appendix I, page 98), but it is crucial for his 'real semantics', i.e. for arriving at a connection between music and a reality external to it. And obviously, the internal semantics precedes the determination of such an external connection.

But the previous section 4.2 suggests that a dynamic approach may similarly cater for a compositional approach to 'external semantics'. However, this has only been suggested for a simple example in the previous section – albeit one which Schlenker treats non-compositionally. It must be said though that Schlenker's aim is to develop a formal theory of musical meaning, and meaning has so far been rather poorly understood – let alone formally – in the domain of music, so it should perhaps come as no surprise that a feature such as compo-

[40] `https://deneeve.github.io/ac/edu/sgr/str.zara.abmajor.mp3`

sitionality is not at the foreground of initial explorations into what is possible when there is not yet much to go on. Moreover, the ambiguity for even a simple example such as in figure 11, where several tonal centres turn out to be demonstrably possible, points to a potentially overwhelming situation if one's aim is to interpret music beyond the simplest examples in a compositional fashion.

Compositionality, however, is a methodological principle, as Groenendijk and Stokhof point out in *Dynamic Predicate Logic* [GS91], section 1. It simply means that the meaning of the whole is a function of the meaning of the parts, in whatever way functions are set up to propagate meanings from smaller to larger chunks of utterances. The present situation with respect to musical semantics may in fact be similar to the one concerning discourse semantics in the 1980s, when several non-compositional systems were introduced to deal with cross-sentential phenomena (most significantly anaphoric relations) until in [GS91] it was demonstrated that the most influential of these, discourse representation theory [Kam81], could be handled compositionally, by recasting meanings as updates on information states instead of as truth conditions.

Obviously, this paper has not demonstrated that the same holds for musical semantics, but, scratching the surface, it has nevertheless been shown that for at least a part of one of Schlenker's examples, it is possible to interpret the music note-by-note, and derive a harmony as well as an event-based denotation.

4.4 Further research issues

If the aim is to develop a formal semantics of music, then especially if it is to be compositional, a few final issues should be mentioned. One of the main motivations for this paper has been the observation that Cooper, in [Coo13], seems to have used a better developed notion of event structure than Schlenker in [Sch19], while at the same time Cooper did not provide musical examples that represent 'complete ideas'. But the question what exactly constitutes a musical idea, or a 'phrase', i.e. effectively a musical analogy to a sentence in language, has only been addressed intuitively in the present paper, and goes no further than that a musical piece or segment ends on (relatively) stable part of the harmony. The notion of phrase is significant, since phrases tend to be repreated and developed, and thus point to repeated and modified events in the semantics. The same holds for phrase parts, such as antecedent-constituent structures (cf. section 2). A (compositional) musical semantics would need the concepts of phrases and constituents to be clear, since these are the things that would potentially be substitutable if they have the same meaning.

Even if in the previous section, harmony was decided via dynamically interpreting note material, it seems nevertheless correct to say that this is not part of the semantics proper – and Schlenker, as mentioned in the previous section, considers harmony at most to be in the realm of what he terms 'internal' semantics, i.e. not part of a reality external to music. And if harmony aids decisions on phrase and constituent boundaries, then plausibly, it is part of the syntax.

Also, in the examples in this section, interpretations have started from a listening state of ignorance, i.e. from an empty (section 4.2) or universal (section 4.3) state. This may not be appropriate, since it ignores what might be in the state prior to the start of the interpretation process. For instance, different persons may interpret the same piece of music differently because of differences in their initial state, or the same person may come to interpret the same piece differently over time. This may on the one hand ignore different ways in which people experience music, but on the other it may in fact be fitting if these different ways are in the possible set of denotations. For instance, if an approaching bird and an approaching car are both valid denotations for a given piece, then plausibly differences between emotional affect associated with these two scenarios might help explain different feelings towards the same music.[41]

In any case, Groenendijk and Stokhof's dynamic semantics may provide some very concrete inspirations. Schlenker likens his semantics to a logic restricted to conjunction – and writes (as mentioned at the beginning of this section)[42] that musical semantics differs from formal semantics of language because the order of conjuncts matters. But as indicated, the order of conjuncts also matters in [GS91]. In section 3 (page 54), Groenendijk and Stokhof state their semantics of conjunction in terms of input and output assignments: $[\![\phi \wedge \psi]\!] = \{\langle g, h\rangle | \exists k : \langle g, k\rangle \in [\![\phi]\!] \textit{ and } \langle k, h\rangle \in [\![\psi]\!]\}$. This casts a model assignment g as an input assignment – i.e. a (listening) state – for the interpretation of ϕ as well as for the whole conjunction, and then partial input assignments k and h for the respective conjuncts, such that h is output state for ψ as well as for the conjunction. This is not how Schlenker describes his semantics, but is arguably the level of description one would want for a (compositional) formal musical semantics.

5 Concluding remarks

In this paper, real-world denotations of musical segments have been given such that these adhere to the fine-grainedness in Cooper [Coo13], while broadly following the methodology of Schlenker [Sch19]. This way, the problem was addressed that while Schlenker provides semantics for complete musical ideas but only gives rough specifications of real-world events as his intended denotations, Cooper uses much more precise event structures for his real-world as well as musical events but does not provide examples that represent complete musical ideas – nor does he focus on the problem of musical semantics. So the contribution of the paper is the use of precise event structures – meaning that events may have subevents and may be specified as occuring serially as well as in parallel – for note-level musical events that are transformed into harmony-level events, in order to obtain equally fine-grained world events.

It has moreover been shown that plausibly, dynamic interpretation can provide a way to arrive at denotations for musical segments such that the semantics can be compositional, unlike in Schlenker's approach. This involved making (some of) the thinking behind how harmonies are understood explicit, such that these can be arrived at by interpreting individual notes instead of having to make assumptions about how an interpreter grasps note sequences. But

[41] See also section 2.1.

[42] Viz. [Sch19], section 7.1, page 72.

future research would have to indicate to what extent a dynamic approach is viable, and how harmonic and denotational interpretation interact.

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.

[GS88] J. Groenendijk and M. Stokhof. Context and Information in Dynamic Semantics. Technical report, Institute for Language, Logic and Information, University of Amsterdam, 1988. https://eprints.illc.uva.nl/1135/1/LP-1988-08.text.pdf.

[GS91] J. Groenendijk and M. Stokhof. Dynamic Predicate Logic. *Linguistics and Philosophy*, 14(1):39–100, 1991.

[GS96] J. Groenendijk and M. Stokhof. Changing the Context: Dynamics and Discourse. In E. Doron and S. Wintner, editors, *Proceedings of the 11th Annual Conference and of the Workshop on Discourse*, pages 104–128. Jerusalem: Israel Association for Theoretical Linguistics, 1996. https://hdl.handle.net/11245/1.123536.

[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.

[Kam81] H. Kamp. A Theory of Truth and Semantic Representation. In J. Groenendijk, T. Janssen, and M. Stokhof, editors, *Formal Methods in the Study of Language*, pages 277–322. Amsterdam: CWI, 1981.

[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. https://doi.org/10.1007/s13164-018-0384-5.